Article

Molecular Cloning and Characterization of Carotenoid Pathway Genes and Carotenoid Content in *Ixeris dentata* var. *albiflora*

Chinreddy Subramanyam Reddy ¹†, Sang-Hoon Lee ¹†, Jeong Soo Yun ², Jae Kwang Kim ², Sang Won Lee ¹, Mok Hur ¹, Seong Cheol Koo ¹, Mi Ran Kim ¹, Woo Moon Lee ¹, Jae Ki Jang ¹, Yoonkang Hur ⁴, Yeon Bok Kim ^{1,3*}

- ¹ Department of Herbal Crop Resources, National Institute of Horticultural & Herbal Science, RDA, Eumseong-gun, 27709, Korea; <u>suumani@gmail.com</u>; <u>omega119@naver.com</u>; <u>swlee1004@korea.kr</u>; <u>mok0822@korea.kr</u>; <u>ksch992@korea.kr</u>; <u>milan80623@korea.kr</u>; <u>wmlee65@korea.kr</u>; <u>changjk@korea.kr</u>; <u>yeondarabok@korea.kr</u>
- ² Division of Life Sciences, College of Life Sciences and Bioengineering, Incheon National University, Incheon 22012, Korea; <u>210621036@inu.ac.kr</u>; <u>kjkpj@incheon.ac.kr</u>
- <u>³ Department of Medicinal and Industrial Crops, Korea National College of Agriculture and Fisheries, Jeonju</u> <u>54874, Korea; yeondarabok@korea.kr</u>
- ⁴ Department of Biology, Chungnam National University, 99 Daehak-ro, Yuseong-gu, Daejeon 34134, Korea; ykhur@cnu.ac.kr
- ⁺ This author is contributed equally to this work.
- * Correspondence: E-Mail: yeondarabok@korea.kr; Tel.: +82-63-238-9092; Fax: +82-63-238-9080

IdPSY CbPSY CcPSY CsPSY HaPSY TePSY	MSARLTGVWYSENSEVCSGLGEVETTRVUDPSRLLAKEKGFFRSERFKNAGRKQRCYSYFGELGFLGSRKLDVVSR-VAS MSVRLI-WVYSENSELSNGLGEVDTKLUDASKERSVFKSSRIKNVGKKYKKHCKSCYVADLKFGASGLANGRKSQEDVSRVVAN MSARLI-WVYSENSOVCNGLSFVETTRVIDSGRLFAKEKGVLRSDRRMGNVKYSCFHELGFLGSNKNGRRSDVVSRVVAN MSARLI-WVYSENSEVSGFGFVESVREGNRLDSSKF5FREFLICHGFKVKKQSYKIEKKGNFFVLSSMVAN MSARLM-WVYSENSEVCNQLGFVETTRTTDSSNKKGLFRLKNVGRKHRCYSVENKVG- MSARLI-WVYSENSEVCNQLGFVETTRTTDSSLGSLETEEKRKGNFVLSSMVAN MSARLI-WVYSENSQLSSGLS	80 85 82 75 69 52
IdPSY CbPSY CcPSY CsPSY HaPSY TePSY	SSÖNLAISSEQIVYDVVLKQAALVKEQMRSREDMEVKEDIVLEGILGLINEAYDRCGEVCAEYAKTFYLGILLMTPERRKAIWAIYV FTGLAISSEQIVYDVVLKQAALVKQMKGKEBEEDMEVKEDIVLFGILGLISEAYDRCGEVCAEYAKTFYLGILLMTPERRKAIWAIYV FSGLAVSSEQIVYDVVLKQAALVKEQMRSREEDMEVKEDIVLFGISLISEAYDRCGEVCAEYAKTFYLGILLMTPERRKAIWAIYV FSGLAISSEQIVYDVVLKQAALVKKUKIKISEEDMEVKEDIVLFGISLIVLFGSLSAIDRCGEVCAEYAKTFYLGILLMTPERRKAIWAIYY FSGLAISSEQIVYDVVLKQAALVKKUKIKISEEDVDVKFEIVLFGSLIVLFGSLGEVCAEYAKTFYLGILLMTPERRKAIWAIYV FSGLAISSEQIVYDVVLKQAALVKKUKISEEDVDVKFEIVLFGSLGLINEAYDRCGEVCAEYAKTFYLGILLMTPERRKAIWAIYV FSGLAISSEQIVYDVVLKQAALVKKUKISEE-DVDVKFEIVLFGILGLINEAYDRCGEVCAEYAKTFYLGILLMTPERRKAIWAIYV	167 179 170 162 157 142
IdPSY CbPSY CcPSY CsPSY HaPSY TePSY	MCRRTDELVDGPNASHITEKALDRWESRLEDLEN GREFDMLDAALSDTVSKFEVDI QEFKDMIDGMRMDLRKSRYDNFDELYLYCYYVAG MCRRTDELVDGPNASHITERALDRWESRLEDLFNGREFDMLDAALSDTVSKFEVDI QEFKDMIDGMRMDLRKSRYKTEPDELYLYCYYVAG MCRRTDELVDGPNASHITEKALDRWESRLEDLYKGREFDMLDAALSDTVSKFEVDI QEFKDMIDGMRMDLRKSRYENFDELYLYCYYVAG MCRRTDELVDGPNASHITERALDRWESRLEDLFNGREFDMLDAALSDTVSKFEVDI QEFKDMIDGMRMDLRKSRYENFDELYLYCYYVAG MCRRTDELVDGPNASHITERALDRWESRLEDLFNGREFDMLDAALSDTVSKFEVDI QEFKDMIDGMRMDLRKSRYENFDELYLYCYYVAG MCRRTDELVDGPNASHITERALDRWESRLEDLFNGREFDMLDAALSDTVSKFEVDI QEFKDMIDGMRMDLRKSRYENFDELYLYCYYVAG MCRRTDELVDGPNASHITERALDRWESRLEDLFNGREFDMLDAALSDTVSKFEVDI QEFKDMIDGMRMDLRKSRYENFDELYLYCYYVAG MCRRTDELVDGPNASHITERALDRWESRLEDLFNGREFDMLDAALSDTVSKFEVDI QEFKDMIDGMRMDLRKSRYENFDELYLYCYYVAG	257 269 260 252 247 232
IdPSY CbPSY CcPSY CsPSY HaPSY TePSY	TVGLMSVPIMGIDRDSLATTESVYNAALALGIANQLITNILRDVGEDARRGRVYLPQDELAKAGLSDEDIFAMKVEDKWRVFMKKQIKRAR TVGLMSVPIMGIAPDSQASTESVYNAALALGIANQLITNILRDVGEDARRGRVYLPQDELAQAGLSDEDIFAGKVEDKWRIFFMKRQIKRAR TVGLMSVPIMGIAPESQATTESVYNAALALGIANQLITNILRDVGEDARRGRVYLPQDELAQAGLSDEDIFAGKVENKWRVFMKRQIKRAR TVGLMSVPIMGIAPESQATTESVYNAALALGIANQLITNILRDVGEDARRGRVYLPQDELAQAGLSDEDIFAGKVENKWRVFMKRQIKRAR TVGLMSVPIMGIAPESQATTESVYNAALALGIANQLITNILRDVGEDARRGRVYLPQDELAQAGLSDEDIFAKVENKWRVFMKRQIKRAR TVGLMSVPIMGIAPESNAPTESVYNAALALGIANQLITNILRDVGEDARRGRVYLPQDELAQAGLSDEDIFAKVENKWRFFMKRQIKRAR TVGLMSVPIMGIAPESNAPTESVYNAALALGIANQLITNILRDVGEDARRGRVYLPQDELAQAGLSDEDIFAKVENKWRFFMKRQIKRAR	347 355 322 342 337 322
IdPSY CbPSY CcPSY CsPSY HaPSY TePSY	TFFDQAEEGVTQLSSASRWFVWASLLLYRQILDEIEANDYNNFTRRAYVSKPKRIVALFUAYAKSLVPPSSRKLVPN 424 AFFDAAEGVTQLSSASRWFVWASLLLYRQILDEIEANDYNNFTRRAYVSKPKRIVALFUAYAKSLVPPSRSGILSKTMDV 436 TFFDQAEEGVTQLSSASRWFVWASLLLYRQILDEIEANDYNNFTRRAYVSKPKRIVALFUAYAKSLVPPSRSGLSKTMDV 436 FFFDAEEGVTQLSSASRWFVWASLLLYRQILDEIEANDYNNFTRRAYVSKPKRIVALFUAYAKSLVPPSRLSLLSAKSS- 422 TFFDQAEEGVTQLSSASRWFVWASLLLYRQILDEIEANDYNNFTRRAYVSKPKRIVALFUAYAKSLVPPSRLSLLSAKSS- 414 TFFDQAEEGVTQLSSASRWFVWASLLLYRQILDEIEANDYNNFTRRAYVSKPKRIVALFUAYAKSLVPPSRLSLLSAKSS- 439	

Supplementary Figure Legends

Supplementary Figure S1: Deduced amino acid multiple sequence alignments of the IdPSY with other PSYs. Identical amino acids are indicated by a black background, and similar residues are shaded with a gray background. The aspartate rich domain (DXXXD, where X represents any Amino Acid) was underlined and bold underline indicates trans-isoprenyl diphosphate synthase (trans-IPPS-HH) motif. CbPSY: *Chrysanthemum boreale* (AGU91437.1); CcPSY: *Cynara cardunculus var. scolymus*(KVI03037.1); CsPSY: *Camellia sinensis* (AJB84620.1); HaPSY: *Helianthus annuus* (CAC19567.1); TePSY: *Tagetes erecta* (AAM45379.1): IdPSY: *I. dentata* (MF611777).



Supplementary Figure S2: Deduced amino acid multiple sequence alignments of the IdPDS with other PDSs. Identical amino acids are indicated by a black background and similar residues are shaded with a gray background. The arrowed black underline indicates the dinucleotide-binding domain and another underline represents carotene binding domain. CmPDS: *Chrysanthemum x morifolium* (BAE79552.1); CsPDS: *Camellia sinensis* (AHB32104.1); DkPDS: *Diospyros kaki* (ACY78343.1); HaPDS: *Helianthus annuus* (AHA36971.1); TePDS: *Tagetes erecta* (AAG10426.1); IdPDS: *I. dentata* (MF611774).

IdZDS	-MATSSTSSTSSLCEPAUSAAGTRNSP-TTDTFLRCRRSROLTNIKAHKSWURSDIDDRDWSDMRMMAPKGGPPPPPPHWRGPKGYAU	88
CmZDS	MARTSSTSST.CEPAUSAA GTRH SAP-TNDTFLECNEHEN RO-INRI.KEWURSDIDK DUSDMERNA PKGLEPPEPERKI KVAT T	85
DCZDS	- AATSSTYPPANSRPDSAGTSLSRCRPLAOLR-THRUMWURSDIEKNUSDMSTNAPKGLEPPEPEHYRGPKIKVATT	77
HaZDS	-MATSS-SSTASLCFPANSAAGTRSSHTTTSTLLRCRRSROLTNKVRKAVIRSDTDRDVSDMRWAPKCJPPPPPPPPPKKKAALI	88
LCZDS	-MATSSAYFCCPANSTTGKKHVAPNGSAGFLVFGCPRLSNRWYRKSWIRADIDSMVSDMGWNAPKGUAPPPPPPHHYRGPKUKVATI	85
NSZDS	- MATSSAVLCCPATSATGKKH TOPNGSAGFLVFRGPRLSNEWURKSVIRADIOSMVSDMSTNA PKGLPPPEPEHVRGPKLKVAT I	85
		00
IdZDS	GAGMACMSWAVDIMADOGHDVDIYYDSRTDIICGKVGSEVDKRGNHIBMGMHVEEGGVANMERIMAKKVGADKNMAVKDHWHWEVARGDIGDI	17
CmZDS	GAGLAGMSTAVELLDOGHEVDIYESRTFIGGKVGSFVDKGGNHIEMGLHVFFGCYNNLFRLLKKVGAEKNLLVKDHTHTFVNRGGEIGEL	17
DCZDS	GAGLAGMSTAVELLDOGHEVD I YES REFIGKVGSFVD KRGNH I EMGLHVFFGCYNNLFRLLKKVGAEKNLLVKDHTHTFVNRGGENGEL	16
HaZDS	GAGDAGMSTAVEDDOGHEVD I VESSTEL GEKVGSEVD KOGNH I EMGMEVFEGGYNNDERMAKKVGAEKNDAVKDEMETEVNKGGELGED	17
LCZDS	GAGLAGMSTAVELLDÖGHEVD I YESRTFI GGKVGSFVD RRGNH I EMGLHVFFGCYNNLFRLLKKVGAEKNLLVKDHTHTFVNRGGEI GEL	17
NSZDS	GAGLAGMSTAVELLDOGHEVDIYESREFIGGKVGSFVDRRGNHIEMGLHVFFGCYNNLFRLLKKVGAEKNLLVKDHTHTFVNRGGEIGEL	17
	dinucleotide-binding domain	
IdZDS	DFRPPVGAPLHGINAFLTTNOLNTYDKARNAVALALSPVVRALVDPDGAMTOIRNLDNISFSEWFMSKGGTRKSIORMWDPVAYALGFID	26
CmZDS	DFRFPVGAPLHGINAFLTTNOLKTYDKARNAVALALSPVVRALVDEDGAMTOIRSLDNVSFSEWFMSRGGTRASIORMWDEVAYALGFID	26
DCZDS	DFRFPVGAPLHGINAFLTTNOLTTYDKARNALALALSPVVRALVDFDGAMRDIRNLDNISFSEWFLSKGGTRKSIORMWDPVAYALGFID	25
HaZDS	DFRFPVGAPLHGINAFLTTNOLKTYDKARNAVALALSPVVRALVDFDGAMTOIRNLDNISFSEWFMSKGGTRISIORMWDPVAYALGFID	26
LCZDS	DFRFPVGAPLHGINAFLTTNOLKTYDKARNAVALALSPVVRALVDEDGALOOIRDLDSVSFSEWFMSKGGTRASIORMWDEVAYALGFID	26
NSZDS	DFRFPVGAPLHGINAFLTTNOLTTYDKARNAVALALSPVVRALVDFDGALOOIRDLDSVSFSEWFMSKGGTRTSIORMWDFVAYALGFID	26
IdZDS	CDNISARCMLTIFSLFATKTEASLLRMLKGSPDVYLSGPIRDYIEKGGRFHLRWGCREILYDKSANGDTYVTGLAMSKATOKKIVKADV	35
CmZDS	CDNISARCMLTIF <mark>S</mark> LFATKTEASLLRMLKGSPDVYLSGPIRDYI <mark>IE</mark> KGGRFHLRWGCRE <mark>I</mark> LY <mark>EK</mark> SDNGDTYVTGLAMSKATOKKIVKADA	35
DCZDS	CDNMSARCMLTIF <mark>SLFATKTEASLLRMLKGSPDVYLSGPIRDYTTO</mark> KGGRFHLRWGCRE <mark>TLYEKSSDGOT</mark> YISGIAMSKATOKKVVKADA	34
HaZDS	CDNISARCMLTIF <mark>S</mark> LFATKTEASLLRMLKGSPDVYLSGPIRDYI <mark>IE</mark> KGGRFHLRWGCRE <mark>ILYEKS</mark> ANGDTYVTGLAMSKATOKKIVKADA	35
LCZDS	CDNISARCMLTIFALFATKTEASLLRMLKGSPDVYLSGPTKKYIIDKGGRFHLRWGCREVLYETSSDGSMYVSGLAMSKATOKKTIKADA	35
NsZDS	CDNISARCMLTIFALFATKTEASLLRMLKGSPDVYLSGPTKKYILDKGGRFHLRWGCREVLYETSSDCSMYVSGLAMSKATOKKIVKADA	35
IdZDS	Y <mark>iaacdvegikrllesnwreweffdn</mark> iyklvgvevvtvqlryngwytemqdlers <mark>rqsrkae</mark> gldnlly <mark>b</mark> edadfscfadlal <u>a</u> spedyy	44
CmZDS	Y <mark>T</mark> AACDVPGIKRLLP <mark>SS</mark> WR <mark>DW</mark> EFFD <mark>D</mark> IYKLVGVPVVTVQLRYNGWVTEMQDIER <mark>A</mark> RQ <mark>SRKA<mark>H</mark>GLDNLLY<mark>H</mark>PDADFSCFADLAL<mark>H</mark>SPEDYY</mark>	44
DCZDS	Y <mark>y</mark> aacdvpgikrllp <mark>so</mark> wr <mark>ew</mark> effdn <mark>iyklvgvpvvtvolryngwtemodlers</mark> ro <mark>mrhaa</mark> gldnllyspdadfscfadlal <mark>a</mark> spedyy	43
HaZDS	Y <mark>T</mark> AACDVPGIKRLLP <mark>SN</mark> WR <mark>EW</mark> EFFD <mark>N</mark> IYKLVGVPVVTVQLRYNGWVTELQDLER <mark>S</mark> RQ <mark>HRQ</mark> AAGLDNLLY <mark>H</mark> PDADFSCFADLAL <mark>A</mark> SPEDYY	44
LCZDS	Y <mark>y</mark> aacdvpgikrlvp <mark>ok</mark> wreleffdn <mark>iyklvgvpvvtvqlryngwtelqdlers</mark> ro <mark>lkra</mark> rgldnllyr <mark>pdadfscfadlale</mark> s pedyy	44
NsZDS	Y <mark>y</mark> aacdvpgikrlvp <mark>ok</mark> wr <mark>el</mark> effd <mark>n</mark> iyklvgvpvvtvQlryngwtelodler <mark>s</mark> Ro <mark>lkr</mark> angldnllynpddfscfadlal <mark>a</mark> s pedyy	44
IdZDS	ID <mark>GQGSLLQCVLTPGDPYMPLPNBEIIS</mark> RV <mark>QK</mark> QVL <mark>S</mark> LPPSSQGLEVTWSSVVKI <mark>G</mark> QSLYREGPGKDPFRPDQ <mark>K</mark> TPVKNFFLAGSYTKQDY	53
CmZDS	<mark>ID</mark> GQGSLLQCVLTPGDPYMPLPN <mark>BEIIR</mark> RV <mark>HE</mark> QVL <mark>V</mark> LFPSSQGLEVTWSSVVKI <mark>R</mark> QSLYREGPGKDPFRPDQ <mark>K</mark> TPVKNFFLAGSYTKQDY	53
DCZDS	LE <mark>GQGSLLQCVLTPGDPYMPLPN<mark>G</mark>EIIERV<mark>TK</mark>QVLALFPSSQGLEVTWSSVVKI<mark>G</mark>QSLYREGPGKDPFRPDQ<mark>R</mark>TPVENFFLAGSYTKQDY</mark>	52
HaZDS	IE <mark>GQGSLLQCVLTPGDPYMPLPN</mark> EEII <mark>S</mark> RV <mark>SK</mark> QVLALFPSSQGLEVTWSSVVKI <mark>G</mark> QSLYREGPGKDPFRPDQ <mark>K</mark> TPV <mark>K</mark> NFFLAGSYTKQDY	53
LCZDS	IE <mark>GQGSLLQCVLTPGDPYMPLPND</mark> EII <mark>KRVSK</mark> QVL <mark>S</mark> LFPSSQGLEVTWSSVVKI <mark>G</mark> QSLYREGPGKDPFRPDQKTPVENFFLAGSYTKQDY	53
NsZDS	<u>ie</u> gqsllqcvltpgdpymplln <mark>deiir</mark> rv <mark>sk</mark> qvl <mark>a</mark> lfpssqglevtwssvvki <mark>g</mark> qslyregpgkdpfredqrtpp <mark>brfpdgsytkqdy</mark>	53
IdZDS	IDSMSGAWASERO AFHSDPIKRHLSRISS 567	
CmZDS	HISMERANDSEXQASHFICDACEEDALERKQLAAVESVGTIGVDEATAV 584	
DcZDS	IDSMEGAVINSGRQASAYICDAGEDHAALQKKIGVIESNTPTG-ABISHV 575	
HaZDS	HISMERAWISERQASAFICDACEEDAALRKVLAAIQSIDNVGVDBISAV 587	
LCZDS	MISMICAWISCZQASAY ICNACEQUALITIKNIA SAESNEISKGVSLSDEISKV 588	
NtZDS	IDSMEGAVISGRQASAYVCDAGEK <mark>LVA</mark> FRKKIAAAE <mark>SNEHS</mark> EDVSVSDEHSLV 588	
	Carotene binding domain	

Supplementary Figure S3: Deduced amino acid sequence multiple alignments of the IdZDS with other ZDSs. Identical amino acids are indicated by a black background, and similar residues are shaded with a gray background. The arrowed black underline indicates the dinucleotide-binding domain and another underline represents carotene binding domain. CmZDS: *Chrysanthemum x morifolium* (BAE79555.1); DcZDS: *Daucus carota subsp. sativus* (NP_001316091.1); HaZDS: *Helianthus annuus* (AHA36972.1); LcZDS: *Lycium chinense* (AIZ50712.1); NtZDS: *Nicotiana tabacum* (AEG73891.1); IdZDS: *I. dentata* (MF611772).

IdLCYB CmLCYB RkLCYB TeLCYB ToLCYB VvLCYB	- MDTLLRTHSSFEFLHELNRFAGNLTNICSS SQIHEPRLSPKKSHSKSGRGGCVKAS-SSALLELVPETKKENIDFELFLDPSRG MMDTLLKIHNSFEFVHESKRFAGNVALSSKRPYFKWGYNNYNKDCVKASGSSALLELVPETKKENIDFELFUMUTYKG -MDTLLKTHNKLEFLHELHGJAGNLINNINGLNQSKSGFQDFRFGPKKSQFKLGQKYCVKAS-SSALLELVPETKKENIDFELFYDFYF -MDTLLKTHNKLEFLHELHGJAGNLINNINGLNQSKSGFQDFRFGPKKSQFKLGQKYCVKAS-SSALLELVPETKKENIDFELFYDPYF -MDTLLKTHNKLEFLHELHGJAGNLINNISLSSKPQIHEPRLSPKKSHLKWSHGCVKAS-SSALLELVPETKKENIDFELFYDPYF -MDTLLKTHNKLEFLHELHGJAEKLGNLTFFRLQNQEFRFGPKKSNLKWGRNGCVKAS-SSALLELVPETKKENIDFELFYDPYF -MDTLLKTHNKLEFLHELHGJAEKLGNLTFFRLQNQEFRFGPKKSNLKWGRNGCVKAS-SSALLELVPETKKENIDFELFYMT SCONSERVED FFLHELHGJAEKLGNLTFFRLQNQEFRFGFKKSNLKWGRNGCVKAS-SSALLELVFE CONSERVED FFLHELHGJAEKLGNLTFFRLQNQEFRFGFKKSNLKWGRNGCVKAS-SSALLELVFE CONSERVED FFLHELHGJAEKLGNLTFFRLQNQEFRFGFKKSNLKWGRNGCVKAS-SSALLELVFE CONSERVED FFLHELHGJAEKLGNLTFFRLQNQEFRFGFKKSNLKWGRNGCVKAS-SSALLELVFE CONSERVED FFLHELHGJAEKLGNLTFFRLQNQEFRFGFKKSNLKWGRNGCVKAS-SSALLELVFE CONSERVED FFLHELHGJAEKLGNLTFFRLQNQEFRFGFKKSNLKKWGRNGCVKAS-SSALLELVFE CONSERVED FFLHELHGJAEKLGNLTFFRLQNQEFRFGFKKSNLKKWGRNGCVKAS-SSALLELVFE CONSERVED FFLHELHGJAEKLGNLTFFRLQNQEFRFGFKKSNLKWGRNGCVKAS-SSALLELVFE CONSERVED FFLHELHGJAEKLGNLTFFRLQNQEFRFGFKKSNLKWGRNGCVKAS-SSALLELVFE CONSERVED FFLHELHGJAEKLGNLTFFRLQNQEFRFGFKKSNLKWGRNGCVKAS-SSALLELVFE CONSERVED FFLHELHGJAEKLGNLTFFRLQNQEFRFGFKKSNLKWGRNGCVKAS-SSALLELVFE CONSERVED FFLHELHGJAEKLGNLTFFRLQNQEFRFGFKKSNLKWGRNGCVKAS-SSALLELVFE CONSERVED FFLHELHGJAEKLGNLTFFRLQNGFFLHEFFLHELFF	85 79 85 85 85
IdLCYB CmLCYB RkLCYB TeLCYB ToLCYB VvLCYB	LVVDLVVVGGGF GLAVAQQVSDAGLTVCSIDFSPKLIWPNNYGVWDEFEAMDLLDCLDTTWSSAVVYIDENSTKNLGREYGRVNRKQL VVDDLVVVGGGF GLAVAQQVSDAGLTVCSIDFSPKLIWPNNYGVWDEFEAMDLLDCLDTTWSSAVVYDENSTRSINRFYARVNRKQL LVVDLATVGGGFAGLAVAQQVSDAGLTVCSIDFSPKLIWPNNYGVWDEFEAMDLLDCLDTTWSSAVVYDERSTKSINRFYARVNRKQL VVVDLVVVGGGF GLAVAQQVSDAGLTVCSIDFSPKLIWPNNYGVWDEFEAMDLLDCLDTTWSSAVVYDERSTKSINRFYARVNRKQL LVVDLVVGGGF GLAVAQQVSDAGLTVCSIDFSPKLIWPNNYGVWDEFEAMDLLDCLDTTWSSAVVYDERSTKSINRFYARVNRKQL LVVDLVVGGGF GLAVAQQVSDAGLTVCSIDFSPKLIWPNNYGVWDEFEAMDLLDCLDTTWSSAVVYDERSTKNLGRFYGRVNRKQL LVVDLVVGGGF GLAVAQQVSDAGLTVCSIDFSPKLIWPNNYGVWDEFEAMDLLDCLDTTWSSAVVYIDENSTKNLGRFYGRVNRKQL LVVDLAVVGGGF GLAVAQQVSDAGLTVCSIDFSPKLIWPNNYGVWDEFEAMDLLDCLDTTWSGAVVFIDDHSTKNLGRFYGRVNRKQL LVVDLAVVGGGFAGLAVAQQVSDAGLTVCSIDFSPKLIWPNNYGVWDEFEAMDLLDCLDTTWSGAVVFIDDHSTKNLGRFYGRVNRKQL LVVDLAVVGGGFAGLAVAQQVSDAGLTVCSIDFSPKLIWPNNYGWVDEFEAMDLLDCLDTTWSGAVVFIDDHSTKNLGRFYGRVNRKQL DFnucleotide binding site	175 169 175 178 175 175
IdLCYB CmLCYB RkLCYB TeLCYB ToLCYB VvLCYB	KSKMIKKCIANGVKFHQAKVIKVIHEETKSLLICNDGVTIQATTVLDATGFSRSLVQYDKFYNFGYQVAYGILAEVEEHPFDVDKMIFMD KTKMICKCISNGVKFHQAKVIKVHEEFKSLLICNDGVTIQATTVLDATGFSRSLVQYDKFYNFGYQVAYGILAEVEEHPFDVDKMIFMD KSRMMCKCILNGVKFHQAKVIKVIHEEKSLLICNDGVTIQATVLDATGFSRSLVQYDKFYNFGYQVAYGILAEVEEHPFDVDKMIFMD KSKMIKKCISNGVKFHQAKVIKVIHEETKSLLICNDGVTIQATTVLDATGFSRSLVQYDKFYNFGYQVAYGILAEVEEHFFDVDKMIFMD KSKMIKKCISNGVKFHQAKVIKVIHEETKSLLICNDGVTIQATTVLDATGFSRSLVQYDKFYNFGYQVAYGILAEVEEHFFDVDKMIFMD KSKMIKKCISNGVKFHQAKVIKVIHEETKSLLICNDGVTIQATTVLDATGFSRSLVQYDKFYNFGYQVAYGILAEVEEHFFDVDKMIFMD	265 259 265 268 265 265
IdLCYB CmLCYB RkLCYB TeLCYB ToLCYB VvLCYB	WRDSHLKNNPBIKDRNSKIPTFLYAMPFSSNRIFLEETSLVARPGLQDEDIQERMVCRLKHLGIKVKSIEEDERCVIPMGGPLFVLPQRV WRDSHLGKNTBIKBRNSKIPTFLYAMPFSSNRIFLEETSLVARPGLKMEDIQERMVCRLKHLGIKVKSIEEDERCVIPMGGPLFVLPQRV WRDSHLDNNRBIKBRNSRIPTFLYAMPFSSNRIFLEETSLVARPGLKMEDIQERMEARLRHLGIKVKSIEEDERCVIPMGGPLFVLPQRV WRDSHLDNNPBIKARNSRIPTFLYAMPFSSNRIFLEETSLVARPGLKMEDIQERMAYRLKHLGIKVKSIEEDERCVIPMGGPLFVLPQRV WRDSHLMNNPBIKARNSRIPTFLYAMPFSSNRIFLEETSLVARPGLKFEDIQERMVCRLKHLGIKVKSIEEDERCVIPMGGPLFVLPQRV WRDSHLMNNPBIKARNSRIPTFLYAMFSSNRIFLEETSLVARPGLKFEDIQERMVCRLKHLGIKVKSIEEDERCVIPMGGPLFVLPQRV WRDSHLMNNMBIKNRSRIFTFLYAMFSSNRIFLEETSLVARPGLKFEDIQERMVCRLKHLGIKVKSIEEDERCVIPMGGPLFVLPQRV WRDSHLMNNMBIKORSRIFTFLYAMFSSNRIFLEETSLVARFGLKFEDIQERMVCRLKHLGIKVKSIEEDERCVIPMGGPLFVLPQRV	355 349 355 358 355 355
IdLCYB CmLCYB RkLCYB TeLCYB ToLCYB VvLCYB	BLCY CAD IGIGGTAGMVHPSTGYMVARTLAAAPIVAKSIIQYLSSEKMVAGTDLSAGIWRDLWPIERRROREFFCFGMDILLKLDLEGTRRFF IGIGGTAGMVHPSTGYMVARTLAAAPIVAKSIIQYLNSEKSIFGNDLSAEVWKNLWPIERRROREFFCFGMDILLKLDLEGTRRFF IGIGGTAGMVHPSTGYMVARTLAAAPIVAKSIIQYLNNEKSMVADVTGDDLAAGIWRELWPIERRROREFFCFGMDILLKLDLEGTRRFF IGIGGTAGMVHPSTGYMVARTLAAAPIVAKSIIQYLNNEKSMVADVTGDDLAAGIWRELWPIERRROREFFCFGMDILLKLDLEGTRRFF IGIGGTAGMVHPSTGYMVARTLAAAPIVAKSIIQYLNSEKAASGTDLSAEIWRDLWPIERRROREFFCFGMDILLKLDLEGTRRFF VGIGGTAGMVHPSTGYMVARTLAAAPIVAKSIIQYLSEKSFFC NELSSEVWRDLWPIERRROREFFCFGMDILLKLDLEGTRRFF VGIGGTAGMVHPSTGYMVARTLAAAPIVANSIVQYLGSDRSFFC NELSSEVWRDLWPIERRROREFFCFGMDILLKLDL2GTRRFF Charged region	441 435 441 448 441 441
IdLCYB CmLCYB RkLCYB TeLCYB ToLCYB VvLCYB	DAFFDLEPRYWHGFLSSRLFLPELLTFGLSLFGHASNTCRIEIMAKGTLPLATMINNLVODRE 504 DAFFDLEPRYWHGFLSSRLFLPELLTFGLSLFAHAPNTSRIEIMAKGTLPLANMINNLVODRE 498 DAFFDLEPRYWHGFLSSRLFLPELLTFGLSLFSHASNTCRIEIMCKGTLPLVRMINNLVODRE 504 DAFFDLEPRYWHGFLSSRLFLPELLTFGLSLFGHASNTCRIEIMAKGTLPLANMICNLVRDRE 511 DAFFDLEPRYWHGFLSSRLFLPELLTFGLSLFGHASNTCRIEIMAKGTLPLANMICNLVRDRE 504 DAFFDLEPRYWHGFLSSRLFLPELLTFGLSLFGHASNTCRIEIMAKGTLPLANMINNLVODRE 504	

Supplementary Figure S4: Deduced amino acid multiple sequence alignment of IdLCYB with various LCYbs. Identical amino acids are indicated by a black background, and similar residues are shaded with a gray background. Conserved LCYB region, Di-nucleotide binding site, LCY's specific motif, Cyclase motif βLCY catalytic active domains and charged regions were underlined. CmLCYB: *Chrysanthemum x morifolium* (BAE79544.1); RkLCYB: *Rhododendron kiusianum x Rhododendron indicum* (BAS69436.1); TeLCYB: *Tagetes erecta* (ARR74226.1); ToLCYB: *Taraxacum officinale* (BAE78471.1); VvLCYB: *Vitis vinefera* (AFP28799.1); IdLCYB: *I. dentata* (MF611775).

IdCHXB EgCHXB2	RMPFLCCKFTSRTSOFFSSTRSFDFIFELR-RTFRLTVCFVAGDQKLETQIVENNDTNNN MAAGISMATTSRUIVHSRHSPPLGEKETSLFTFTSKLFHSLQKFHAVLRFKRRKSNLTVCFVLEEEKLADWRIQDSGEGSG MAMPELSFTSSSOM-TYPEHSD2LGENDUFFISSFSENLCTILESE-RESEVUCEVLCEKINTOFFKINAOLETCEE	59 81 78
SPCHYB2		75
SPCHAB2		70
SUCHAB2	MARAN SASSISSI TITANI TITANI TITANI TITANI TITANI TITANI ANG TINGKI MASA NATSI TVO MEDAA	76
VVCHAB2	MUVERSVATSERSFALGAN-AFROMATE FFTSLEFESTRARENTERCK-ARCONCERSENCE	10
IdCHXB	ER ARKKERFWIL AA MS GINSMA MAVYVRESWOMECCE	149
EaCHXB2	DDEGSGSEISAAAAGLAEKVARKRSERFTYTVAAVMSTLGITSMAVMAVYYREAWOMEGGEVEFTALSVGAAVGMET	164
NSCHXB2	IELKIEEOISATRLAEKUARKKSERFTYLLAAVMSSEGITSLAIMAVYYRESWOMEGGEMPLTEMEGTFALSVGAAVGMER	159
SpCHXB2	FEKKIEECILATRLAEKLARKKSERFTYLVAAIMSSFGITSMAVVAVYYRFSWOMEGGBVPVTEMLGTFALSVGAAVGMEF	156
StCHXB2	FEKKIEEOISATRLAEKLARKKSERFTYLVAAVMSSEGITSMAVMAVYYRESWOMEGGBVPLWEMLGTFALSVGAAVGMER	159
VvCHXB2	ETLASOISAARVAEKLARKRSERLTYTVAAVMSSECITSMAVMAVYTRESWOMEGEBVELSEMEGAAVGMER	155
	нхххн нххнн	
IdCHXB	MARWAHEAMWHASIAWHMHESHHKERREDEBIANDWFATINAWEATAMANYEEFHKCI IRECCEGAGIGUTWECMAYMEWHDGAVHKRROVC	239
EgCHXB2	WARWAHWALWHASIWHMHESHHKPREGPFELNDVFATINAGPATALLSYGFFHKGLVPGLCFGAGLGTTVFGMAYMFVHDGLVHRRFPVG	254
NSCHXB2	WARWAHRALWHD SIWHMHESHHK PREGPFEMNDYFA I YNAWPA I ALLNYGFFHKGUI PGLCFGAGLGI TYFGMAYMFYHDGLYHKR FLYG	249
SDCHXB2	WARWAHKALWHASIWHMHESHHK PREGPFELNDVFATTNAWPATALLNYGFFHKGTTPGLCFGAGLGTTVFGMAYMFVHDGLVHKR PPVG	246
StCHXB2	WARWAHKALWHASIWHMHESHHKPREGPFELNDVFATUNAWPATALLNYGFFHKGTPPGLCFGAGLGTTVFGMAYMFVHDGLVHKRFPVG	249
VVCHXB2	WARWAHKALWHASLWHMHESHHKPREGPFEINDVFATUNAUPATALLSYGFFIKGAVPGLCFGAGLGTVFGMAYMFVHDGLVHRRFEVG	245
	HXXH	210
IdCHXB	ELENVEYLERVANTOVTCDHHSSTISETNLSIL273	
EgCHXB2	PURDUPYPRRVASATOLIHISEKEDGUPYGIELGPKELEEVGELEEVGELEENENTNIRETKISKKL- 314	
NeCHYB2	EVANUENTERVATES LUBSEKENGVEVGTET CERETERVG	
SDCHYB2		
SPCHAB2		
UTCHARZ		

Supplementary Figure S5: Deduced amino acid multiple sequence alignment of IdCHXB with various CHXBs. Identical amino acids are indicated by a black background, and similar residues are shaded with a gray background. Spatially conserved four histidine domains believed to be participated in iron ion adhesion while hydroxylation are (HXXXH) and HXXH) where x indicated for any amino acid) marked with text. EgCHXB: *Erythranthe guttata (XP_012858684.1); NsCHXB: Nicotiana sylvestris (XP_009780858.1); SpCHXB: Solanum pennellii (XP_015080043.1); StCHXB: Solanum tuberosum (XP_006360197.1); VvCHXB: Vitis vinifera (XP_002273581.1); IdCHXB: <i>I. dentata* (MF614116).

IdCHXE CsCHXE DcCHXE HuCHXE MnCHXE SiCHXE	MPSSLSLSS-FSLHANPSHFPHPLYKPTIPYPSTNLPKSLSIKSSIDKPEKPTNNTTKQSSWVSPDWLTAPTRSLSNTQ-SD MPYHSISSLSLLP-IPIRQNLSKHHPPFHQPPHSLPLSIKSSLDNKPPKSNQGSWVSPDWLTSLTKSITLSKD MHSCVALSSITFPSLPPPRLTINIPLQSFSVKSSIEKSPATKPKP-TSPSKSTSWVSPDWLTSLTKSLTIGS-ND MSLSLSFSSSLSILSPPLLKPTSIFATTTPLHSLSVKSSIDKDPASTSNFKRNTTKPSSWVSPDWLTSPTKSLTIGR-DE MSSSSLSLSS-LSSTTHHRHRRQILAPSFCTKNSTNNKPPTSKLTSWVSPDWLTSLTKSVTILGPKD	21 79 72 73 79 66
IdCHXE CsCHXE DcCHXE HuCHXE MnCHXE SiCHXE	DSNIFIANARUEDVSDLLGGALFLPLFKWMNEYGFIYRLAAGFRNFVIVSDEAIAKHVLRNYGTKYAKGLVAEVSEFLFGSGFAIAEGSU DSNIFIANARUEDVSELLGGALFLPLFKWMNEYGFIYRLAAGFRNFVVVSDEAIAKHVLRNYG-KYAKGLVAEVSEFLFGSGFAIAEG DSNIFIASARUEDVSELLGGALFLPLFKWMNMYGFIYRLAAGFRDFVVVSDEAIAKHVLRNYG-KYAKGLVAEVSEFLFGSGFAIAEG DSGIFIASAQUEDVSELLGGALFLPLFKWMNEYGFIYRLAAGFRNFVVVSDEAMAKHVLRNYG-KYAKGLVSEVSEFLFGSGFAIAEGSL G SAR DVSDLLGGALFLPLFKWMNEYGFIYRLAAGFRNFVVVSDEAMAKHVLRNYG-KYAKGLVAEVSEFLFGSGFAIAEGSL DSNIFIASARUEDVSELLGGALFLPLFKWMNEYGFIYRLAAGFRNFVVVSDEAMAKHVLRNYG-KYAKGLVSEVSEFLFGSGFAIAEGSL G SAR DVSDLLGGALFLPLFKWMNEYGFIYRLAAGFRNFVVVSDEATAKHVLRNYG-KYAKGLVAEVSEFLFGSGFAIAEGSL DSNIFIASARUEDVSDLLGGALFLPLFKWMNEYGFIYRLAAGFRNFVVVSDEATAKHVLRNYGKYAKGLVAEVSEFLFGSGFAIAEGSL	111 168 161 162 168 156
IdCHXE CsCHXE DcCHXE HuCHXE MnCHXE SiCHXE	NTARRRAVVPSLHKKYLSVIVDRVFCKCSERFVEKLKSYARNDTAVNMEQQFSQLTLDVIGJAVFNYNFDSLTADSPVIESVYTALKEAE MTVRRRAVVPSLHKKYLSVIVDRVFCKCAERLVEKLKTNALNGSAVNMEENFSQLTLDVIGJALFNYNFDSLTDSPVIDAVYTS MTARRRAVVPSLHKKYLSVIVDRVFCKCAERLVEKLEISALNGSAVNMEEQFSQLTLDVIGJSVFNYNFDSLTDSSPVIDAVYTALKEAE MTVRRRAVVPSLHKKYLSVIVDRVFCKCAERLVEKLEISALNGSAVNMEEKFSQLTLDVIGJSVFNYNFDSLTSDSPVIDAVYTALKEAE MTARRRAVVPSLHKKYLSVIVDRVFCKCAERLVEKLQASAQDGTAVNMEEKFSQLTLDVIGJSVFNYNFDSLTSDSPVIDAVYTALKEAE MTARRRAVVPSLHKKYLSVIVDRVFCKCAERLVEKLQASAQDGTAVNMEEKFSQLTLDVIGJSVFNYNFDSLTSDSPVIDAVYTALKEAE MTARRRAVVPSLHKKYLSVIVDRVFCKCAERLVEKLQASAQDGTAVNMEENFSQLTLDVIGJSVFNYNFDSLTSDSPVIDAVYTALKEAE MTARRRAVPSLHKKYLSVIVDRVFCKCAERLVEKLQASAQDGTAVNMEENFSQLTLDVIGJSVFNYNFDSLTSDSPVIDAVYTALKEAE	201 258 251 252 258 246
IdCHXE CsCHXE DcCHXE HuCHXE MnCHXE SiCHXE	ARSTDILPYWKIKALCKI I PRQIKAEQAVTVIR ^B TVEELILKCKEI VEKEGEKIDDEDYWNDADPSILRFLLASREEVSSEQLRDDLLSM ARSTDILPYWKINALCKI I PRQIKAEKAVTLIRKTVEELINKCKEI VESEGERINEE YWNEAD PSILRFLLASREEVSSI QLRDDLLSM ARSTDILPYWKIEALCKI I PRQVKAEKAVTVIRTVEELIEKCKKI VDIEGERISEE YWNEAD PSILRFLLASREEVSSI QLRDDLLSM LRSTDILPYWKIEALCKI I PRQVKAEQAVTVIRTVEELIIKCKEI VESEGERIDE BYWNDAD PSILRFLLASREEVSSI QLRDDLLSM ARSTDILPYWKIKALCKI I PRQVKAEQSVIVIRTVEELIIKCKEI VESEGERIDE BYWNDAD PSILRFLLASREEVSSI QLRDDLLSM LRSTDILPYWKIKALCKI I PRQVKAEQSVIVIRTVEELIAKCKEI VESEGERIDE BYWNDAD PSILRFLLASREEVSSI QLRDDLLSM LRSTDILPYWKIKALCKI I PRQVKAEQSVIVIRTVEELIAKCKEI VESEGERIDE BYNNDAD PSILRFLLASREEVSSI QLRDDLLSM LRSTDILPYWKINALCKVI PRQIKAEKAVTI I RQTVEELIAKCKEI VESEGERIDE BYNDAD PSILRFLLASREEVSSI QLRDDLLSM LRSTDILPYWKINALCKVI PRQIKAEKAVTI I RQTVEELIAKCKEI VESEGENINE BEYNNDAD PSILRFLLASREEVSSI QLRDDLLSM LRSTDILPYWKINALCKVI PRQIKAEKAVTI I RQTVEELIAKCKEI VESEGENINE BEYNDAD PSILRFLLASREEVSSI QLRDDLLSM	291 348 341 342 348 336
IdCHXE CsCHXE DcCHXE HuCHXE MnCHXE SiCHXE	LVAGHETTG- LVAGHETTG- LVAGHETTGSVLTWTAYLLSKDPSSLSKAQEEVDRVLQGRSFTYEDIKNLKGLTRCINESURLYPHPPVLLRRAQVADVLPGNYKVNEGQ LVAGHETTGSVLTWTYLLSKDSFILLKAQEEVDRVLGGRPFAYEDIKDLKGLTRCITESURLYPHPPVLLRRAQVDDILPGNYKVKAGQ LVAGHETTGSVLTWTYLLSKDSFILLKAQEEVDRVLQGRSPSYEDVKDLKGLTRCITESURLYPHPPVLLRRAQVDDILPGNYKVKAGQ LVAGHETTGSVLTWTAYLLSKESSALLKAQEEVDRVLQGRSPSYEDVKDLKGLTRCITESURLYPHPPVLLRRAQVDDULPGNYKVKAGQ LVAGHETTGSVLTWTAYLLSKESSALLKAQEEVDRVLQGRSPSYEDVKDLKGLTRCITESURLYPHPPVLLRRAQVDVLFGNYKVNAGQ LVAGHETTGSVLTWTAYLLSKESSALLKAQEEVDRVLQGRLPTYEDIENLKGLTRCITESURLYPHPPVLLRRAQVDVLFGNYKVNAGQ LVAGHETTGSVLTWTAYLLSKESSALLKAQEEVDRVLQGRLPTYEDIENLKGLTRCITESURLYPHPVLLRRAQVDVLFGNYKVNAGQ LVAGHETTGSVLTWTAYLLSKESSALLKAQEEVDRVLQGRLPTYEDIENLKGLTRCITESURLYPHPVLLRRAQVADVLFGNYKVNAGQ	300 438 431 432 438 426
IdCHXE CsCHXE DcCHXE HuCHXE MnCHXE SiCHXE	DIMISVYNIHESPOVMDRAEVEVPERFDIDGEMENEINIDERFIPFSGGPRKCIGDQFALLEAIVALAIEVQHMDBELVEDQNISMITGA DIMISVYNIHESSKVMDRAESFIPERFDIDGEVENESNADYRFIPFSGGPRKOVGDQFALLEAIVALAIEVQHMDBELVEDQVIMITGA DIMISVYNIHESSVVMERAESFIPERFDIGSEVENENADYRFIPFSGGPRKOVGDQFALLEAIVALAIFIQHINFELVEDQNISMITGA DIMISVYNIHESSKVMERAESFVPERFDIGGEVENESNADERFIPFSGGPRKOVGDQFALLEAIVALAIFIQHFIDELVEDQNISMITGA DIMISVYNIHESSKVMERAESFVPERFDIGGEVENESNADERFIPFSGGPRKOVGDQFALLEAIVALAIFIQHFIDELVEDQNISMITGA DIMISVYNIHESSKVMERAESFVPERFDIGGEVENENNDERFIPFSGGPRKOVGDQFALLEAIVALAIFIQHFIDELVEDQNISMITGA	300 528 521 522 528 516
IdCHXE CsCHXE DcCHXE HuCHXE MnCHXE SiCHXE	THEBNGLYMEVSQRQTKSAFAASSSSSR THEPINGLYMEVSQR	300 557 548 550 618 542
IdCHXE CsCHXE DcCHXE HuCHXE		

MnCHXE WIFATEQLSCAAVGIK 634 SiCHXE ----- 542

Supplementary Figure S6: Deduced amino acid multiple sequence alignments of the IdCHXE with other CHXE. Identical amino acids are indicated by a black background, and similar residues are shaded with a gray background. The sequence (AGHE) indicates center of α helix which can bind to heme and the sequence (EGEN) indicates salt bridges were underlined. The arrowed underline indicates heme binding region. CsCHXE: *Camellia sinensis* (AJB84623.1); DcCHXB: *Daucus carota subsp. sativus* (NP_001316100.1); HuCHXE: *Herrania umbratica* (XP_021289420.1); MnCHXE: *Morus notabilis* (XP_010098846.1); SiCHXE: *Sesamum indicum* (XP_011094564.1); IdCHXE: *I. dentata* (MF611776).

IdZEP CbZEP CsZEP LsZEP RcZEP VvZEP	NANCHVFCTSINETTTIHSKUHLPIFLHSIHYKHHQFRSKETGSGRÜFGVSSNIKAVLABSPP-QAAEQ-GGAKKÖNV NATSHVYCSFTITLHSKUNLPVPFSSIHYKHHLKSKENGSGKKITNFGKIKALVTPTPP-RKTEQSGGEKEKKI NTSTVFYT-SLNESTTLFSRUHFPIPISRDFSLDLLHPVNSNYG-FRTKENGRMKRMTKPKVSVSBAPPERSSAAEVDGNSKKL NANCNVFCSSINETATLHSKUHLPIFLHSIHSKHHQFRSGSSSSSSIGKÖTGVSSNVKAVLABSPP-QAAER-GGEKKÖNV VASSAFFCNSINESTSVFSRUHFSFPIFSTSTVPFSSFAQYNFH-FKTKKSDHQNKRFTQVKAVVTBSPTVAESNGKLS-EQKKL MASAVFYS-SVQESIFSRUHFPIPISKDSFEPFGHSINYKHY-FRSNPCGQKKRVAQVKATLABATPAPSAPSLPSKRV	77 75 82 80 83 77
IdZEP CbZEP CsZEP LsZEP RcZEP VvZEP	RULVAGGGIGGLVFALAAKRKGFEVVVFERDLSAIRGEGOYRGFIQIQSNALAALEAIDEDVADEVMKAGCIIGORINGLVDGVSGNWYI RULVAGGGIGGLVFALAAKRKGFEVVVFEKDLSAIRGEGOYRGFIQIQSNALAALEAIDHDVADEVMKAGCIIGORINGLVDGISGNWYI RULVAGGGIGGLVFALAAKRKGFDVMVFEKDLSAIRGEGOYRGFIQIQSNALAALEAIDHEVADEVMKIGCIIGDRINGLVDGISGNWYC RULVAGGGIGGLVFALAAKRKGFEVVVFERDLSAIRGEGOYRGFIQIQSNALAALEAIDHEVADEVMKIGCIIGORINGLVDGVSGNWYC RULVAGGGIGGLVFALAAKRKGFEVVVFERDLSAIRGEGOYRGFIQIQSNALAALEAIDHEVADEVMKIGCIIGORINGLVDGVSGNWYC RULVAGGGIGGLVFALAAKRKGFEVVVFERDLSAIRGEGOYRGFIQIQSNALAALEAIDHEVADEVMKIGCIIGORINGLVDGVSGNWYC RULVAGGGIGGLVFALAAKRKGFEVVVFERDLSAIRGEGOYRGFIQIQSNALAALEAIDHEVADEVMKIGCIIGDRINGLVDGVSGNWYC RULVAGGGIGGLVFALAAKRKGFEVVVFERDLSAIRGEGOYRGFIQIQSNALAALEAIDHEVABEVMRAGCIIGDRINGLVDGVSGWYC RULVAGGGIGGLVFALAAKRKGFEVVVFERDLSAIRGEGOYRGFIQIQSNALAALEAIDHEVABEVMRAGCIIGDRINGLVDGVSGWYC	167 165 172 170 173 167
IdZEP CbZEP CsZEP LsZEP RcZEP VvZEP	KFDTFTPAVERGLEVTRVISRMTLQKILADAVGEDIILNCSNVVSFEDHGEKVSVTLENGERFEGDLLVGADGIWSKVRKNLFGEKDVTY KFDTFTPAVERGLEVTRVISRMTLQKILADAVGDEIILNCSNVNFEDHGDKVSVVLENGERFEGDLLVGADGIWSKVRKNLFGEKDVTY KFDTFTPAVERGLEVTRVISRMTLQCILACAVGEDVIMDSNVVNFEDDGCKVTVILENGQRYEGDLLVGADGIWSKVRKNLFGEKDVTY KFDTFTPAVERGLEVTRVISRMTLQCILACAVGEDVIMDSNVVNFEDDGCKVSVTLESGERFEGDLLVGADGIWSKVRKNLFGEKDVTY KFDTFTPAVERGLEVTRVISRMTLQCILACAVGEDVIMDSNVNFEDDGEKVSVTLESGERFEGDLLVGADGIWSKVRKNLFGEKDVTY KFDTFTPAVERGLEVTRVISRMTLQCILACAVGEDVIMNASNVINEQDNEEKVSVTLESGERFEGDLLVGADGIWSKVRKNLFGEKDVTY KFDTFTFAAERGLEVTRVISRMTLQQILACAVGEDVIMNASNVINEQDNEEKVTVILENGQQFEC-LVGADGIWSKVRKNLFGEKDXTY KFDTFTFAAERGLEVTRVISRMTLQQILARAVGEDIIMCSNVVDFEDDGCKVTVILENGQFEGDLLIGADGIWSKVRKNLFGEKEATY	257 255 262 260 261 257
IdZEP CbZEP CsZEP LsZEP RcZEP VvZEP	SGYTCYTGIADFIFEDINTVGYRVFLGHKQYFVSSDVG ^G GKMQWYAFHNEPAGG ^S DKPNGKKERLLQIFG ^G GWCDNVDLHLATDEBAILR SGYTCYTGIADFIFEDINSVGYRVFLGHKQYFVSSDVG ^G GKMQWYAFHNEPAGG ^S DKPNGKKERLLEIFG ^G WCDNVDLHLATDEBAILR SGYTCYTGIADFIFEDIDTVGYRVFLGHKQYFVSSDVG ^G GKMQWYAFHNEPFGGVD ^G STGKKQRLLKIFE ^G WCDNVDLHLATDEBAILR SGYTCYTGIADFIFEDIDTVGYRVFLGHKQYFVSSDVG ^G GKMQWYAFHNEPFGGVD ^G STGKKQRLLKIFE ^G WCDNVDLHLATDEBAILR SGYTCYTGIADFIFEDIDTVGYRVFLGHKQYFVSSDVG ^G GKMQWYAFHNEPFGGVD ^G SPGGKERLLKIFE ^G WCDNVDLHLATDEBAILR SGYTCYTGIADFV ^F 4DIES ^V GYRVFLGHKQYFVSSDVG ^G GKMQWYAFHNEPFGGVD ^G SPGGKERLLKIFE ^G WCDNVIDLHATDEDAILR SGYTCYTGIADFV ^F 4DIES ^V GYRVFLGHKQYFVSSDVG ^A GKMQWYAFHNEPFGGVD ^G SPGGKE ^R LLKIFE ^G WCDNVIDLHATDEDAILR	347 345 352 350 351 347
IdZEP CbZEP CsZEP LsZEP RcZEP VvZEP	RDIEDRMERETWGRGRVTLLGDSVHAMQPNLGQGGCMAIEDSVQLALELDKAWRQSSESGAPIDIASSIRRYEDARRVRVAVIHGDARMA RDIEDRIERETWGRGRITLLGDSVHAMQPNLGQGGCMAIEDSVQLALELDKAWIQSIKSGAPIDIQSSIRRYEDARRVRVAVIHGDARMA RDIYDRAFTENWGRGRVTLLGDSVHAMQPNLGQGGCMAIEDSVQLAMELDKAWERSVESGSPIDVVSSLRSYEDARRVRVAVIHGDARMA RDIEDRIERETWGRGRVTLLGDSVHAMQPNLGQGGCMAIEDSVQLAMELDKAWSRSIESGARVDIAISJRRYEDARRVRVAVIHGDARMA RDIEDRIERETWGRGRVTLLGDSVHAMQPNLGQGGCMAIEDSVQLADELDKAWSRSIESGARVDIAISJRRYEDARRVRVAVIHGDARMA RDIYDREPPTWGRGRVILLGDSIHAMQPNMGQGGCMAIEDSVQLADELDKAWRQSIESGTPVDVVSSLRSYERTRVRVAVIHGDARMA RDIYDREPPTWGRGRVILLGDSVHAMQPNMGQGGCMAIEDSVQLADELDKAWEQSIKSGTPIDVVSSLRSYERTRVRVAIHGMARMA	437 435 442 440 441 437
IdZEP CbZEP CsZEP LsZEP RcZEP VvZEP	AIMASTYKAYLGVGLGPLSFLTNFRIPHPGRVGGRFFIDIGMETMLSWVLGGNGSNIEGRPQOCRLTDKANDELONWEKDDDALERALTG AIMASTYKAYLGVGLGPLSFLTNFRIPHPGRVGGRFFIDIGMETMLSWVLGGNGSKLEGRPQSCRLTDKANDELONWERDDDALERALTG AIMASTYKAYLGVGLGPLSFLTKERIPHPGRVGGRFFIDLAMETMLSWVLGGNGSNLDGRSLSCRLSDKANDGLONWERDDDALERAVNG AIMASTYKAYLGVGLGPLSFLTNFRIPHPGRVGGRFFITIGMETMLSWVLGGNGSNLDGRSLSCRLSDKANDGLONWERDDDALERAVNG AIMASTYKAYLGVGLGPLSFLTNFRIPHPGRVGGRFFITIGMETMLSWVLGGNGSNLEGRPQOCRLTDKANDELONWERDDDALERAVNG AIMASTYKAYLGVGLGPLSFLTNFRIPHPGRVGGRFFITIGMETMLSWVLGGNGSSKLEGRPLSCRLSDKANDGLONWERDDDALERALNG AIMASTYKAYLGVGLGPLSFLTNFRIPHPGRVGGRFFIDIAMETMLSWVLGGNSSKLEGRPLSCRLSDKANDGLONWERDDDALERALNG AIMASTYKAYLGVGLGPLSFLTNKIRIPHPGRVGGRFFIDIAMETMLSWVLGGNSSKLEGRPLSCRLSDKANDGLONWEDDDALERALNG	527 525 532 530 531 527
IdZEP CbZEP CsZEP LsZEP RcZEP VvZEP	SWF17LL	544 615 622 620 620 615
Idzep Cbzep Cszep Lszep Rczep Vvzep	Y	

Supplementary Figure S7: Deduced amino acid sequences multiple alignments of the IdZEP with other ZEPs. Identical amino acids are indicated by a black background, and similar residues are shaded with a gray background. The solid black colour underline indicates the short motifs of lipocalin family proteins and Arrowed red underline represents FHA domain. CbZEP: Chrysanthemum boreale (AGU91434.1); CsZEP: Camellia sinensis (AJB84624.1); LsZEP: Lactuca Sativa (BAE72089.1); RcZEP: Ricinus communis (XP_002523587.1); VvZEP: Vitis vinifera (NP_001268202.1); IdZEP: I. dentata (MF611773).