

# Supplemental Data

## Supplemental Figure S1

**A**

```

GmBCH1  MGDGSSSHLLAGEHKHSLFASWRNSIEAIYPSMAAGLPTAAI LKPY-NLVQPP---IPLSKPTT----SUFNPL 72
GmBCH3  -----MAAGLSTAAI LKPY-NLVQPP---IPLSKPTT----SUFNPL 35
GmBCH2  -----MAAGLSAAITMKSL-LRFHQPHLNLPKSIPTT----UPFSM 42
GmBCH4  -----MAAGLSTAAI LKPY-NLQPP---IPLSKPTT----PPUFNPL 35
GmBCH5  -----MAAGLSAAITMKPL-LRFHQPR--LPKPIPTT----UPFSPL 35
CaCRTR-B -----MAAGLVAAGAGQTV-CFRVNSFLT---RKPFSLVADSLTSLPL 39
AtCHY1  -----MAAGLSTAVTFKPLHRSESSSSTDFRLRLPKSLSG---FSPS 43
    
```

```

GmBCH1  RCFHSTILRVRPRRRM-SGFTVCVLTETSK---EIKTVEQE-QEQVIP-----QAVSAGVAEKLARKRSORFT 139
GmBCH3  RCFHSTILRVRPRRRM-SGFTVCVLTETSK---EIKTVEQE-QEQVIP-----QAVSAGVAEKLARKRSORFT 99
GmBCH2  RIFHHTASPRTQ---KVSTFTVCVLMQDPK---GTHMEIEPQEQPPPPPPPAQQVLSPKIAEKLARRESERFT 106
GmBCH4  RCFRSTILRVRPRRRM-SGFTVCVLTETDPK---EIKKDEQE-QEQVNP-----QVVSARVAEKLARKRSORFT 99
GmBCH5  RIFHHTASPIPQ---NFSTFTVCVLMQDPK---GTHMEIQEQ---PPSPPP-QQVLSQKIAEKLARRESERFT 100
CaCRTR-B AQQFSTT-----RRHRRKPRLTVCVLEDEELKAQLVTSBEEAREREKAM---AKRISDARVAEKLARRESERFT 110
AtCHY1  L-----RFRRESVGVVVEERR---NSPIENDERPESTSTNAIDAEYLALRLIAEKLARKRSERST 97
    
```

```

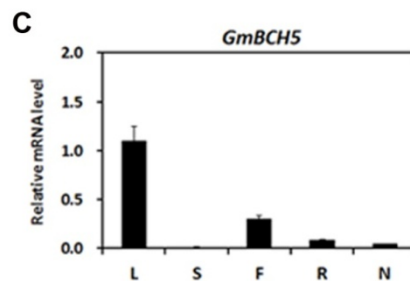
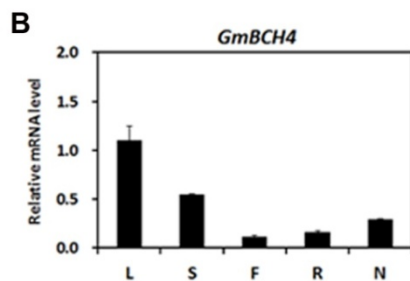
GmBCH1  YLVAAMSSFCITSMAVEAVYRFSWMEGCGDVPWSEMLGTFSLSVGAAVAMEFWARWAHRALWHASLWHEHESH 208
GmBCH3  YLVAAMSSFCITSMAVEAVYRFSWMEGCGDVPWSEMLGTFSLSVGAAVAMEFWARWAHRALWHASLWHEHESH 175
GmBCH2  YLVAAMSSFCITSMAVEAVYRFSWMEGCGEVPWSEMGTFALSVGAAVAMEFWARWAHRALWHASLWHEHESH 182
GmBCH4  YLVAAMSSFCITSMAVEAVYRFSWMEGCGDVPWSEMLGTFSLSVGAAVAMEFWARWAHRALWHASLWHEHESH 175
GmBCH5  YLVAAMSSFCITSMAVEAVYRFSWMEGCGEVPWSEMGTFALSVGAAVAMEFWARWAHRALWHASLWHEHESH 176
CaCRTR-B YLVAAMSSFCITSMAVEAVYRFSWMEGCGEVPWSEMGTFALSVGAAVAMEFWARWAHRALWHASLWHEHESH 182
AtCHY1  YLVAAMSSFCITSMAVEAVYRFSWMEGCGEISLMEGTFALSVGAAVAMEFWARWAHRALWHASLWHEHESH 173
    
```

```

GmBCH1  RPRGPFELNDVEAIIINAVPAIALLSYCFEFKGLVPGLCFAGLGITVFCIAYMEVHDGLVHKRFPVGPPIANVPYF 284
GmBCH3  RPRGPFELNDVEAIIINAVPAIALLSYCFEFKGLVPGLCFAGLGITVFCIAYMEVHDGLVHKRFPVGPPIANVPYF 251
GmBCH2  RPRGPFELNDVEAIIINAVPAIALLSYCFEFKGLVPGLCFAGLGITVFCIAYMEVHDGLVHKRFPVGPPIANVPYL 258
GmBCH4  RPRGPFELNDVEAIIINAVPAIALLSYCFEFKGLVPGLCFAGLGITVFCIAYMEVHDGLVHKRFPVGPPIANVPYF 251
GmBCH5  RPRGPFELNDVEAIIINAVPAIALLSYCFEFKGLVPGLCFAGLGITVFCIAYMEVHDGLVHKRFPVGPPIANVPYL 252
CaCRTR-B RPRGPFELNDVEAIIINAVPAIALLSYCFEFKGLVPGLCFAGLGITVFCIAYMEVHDGLVHKRFPVGPPIANVPYF 258
AtCHY1  RPRGPFELNDVEAIIINAVPAIALLSYCFEFKGLVPGLCFAGLGITVFCIAYMEVHDGLVHKRFPVGPPIANVPYL 249
    
```

```

GmBCH1  RRVAAAHQ LHHSEKFNCPYGLFLGPKVEEVGCEELEKEISRRIRSGS----- 334
GmBCH3  RRVAAAHQ LHHSEKFNCPYGLFLGPKVEEVGCEELEKEISRRIRSGS----- 301
GmBCH2  RRVAASAHQ LHHSEKFNCPYGLFLGPKVEEVGCEELEKEISRRARSYKIANENN---- 314
GmBCH4  RRVAAAHQ LHHSEKFNCPYGLFLGPKVEEVGCEELEKEISRRIRSGS----- 301
GmBCH5  RRVAASAHQ LHHSEKFNCPYGLFLGPKVEEVGCEELEKEISRRARSYKIAREKS---- 308
CaCRTR-B RRVAAAHQ LHHSEKFNCPYGLFLGPKVEEVGCEELEKEISRRIRSGS----- 310
AtCHY1  RRVAASAHQ LHHSEKFNCPYGLFLGPKVEEVGCEELEKEISRRIRSGSYKASGSGSSSS 251
    
```



D

GmZEP1	YATILLCYNLSL--PSTTVFSTRHESVPLNKELELDASLFFVVGYNCGVGRIRKQRKKVMH	58
GmZEP2	-----MAVFSKTHESLMLKELSLDLSFCVSV-----HGSRLTKQRNTLMP	40
GmZEP3	YASILLCNSFN--HSMVAFSKHESLIMYKELSLDLSFCVSV-----YGSRLTKQRKTLMQ	53
Gm09g00260	YAS-----VQLCCTCQVQLQGSIRRYVDYGYF---HHHDKISKHRYYSIRV	45
PA-ZE	YASILLYNSML--LSAAVFSRTHSPIDLNKOFFLDFSLFCIH-IDYHLRSRDRSGQKCLF	57
AtABA1	YGSILPFCYSILPSPSKLDFRTHVFSVSKQFMIDLSLFSG-KPGGVSG-PRS-RRALLG	57
GmZEP1	VKCAVVEAIPGVSPSAKDGNG-----TTPSKKQIRLVAGGGIGGLVFALAARKGFEV	112
GmZEP2	IKASVAEVPVAVRKTVDENEGGGD---GASQKKLRVLVAGGGIGGLVFALAARKGFEV	97
GmZEP3	IKATVAEIPGVVKTIVVSENGGGDGAPOKRRKLRVLVAGGGIGGLVFALAARKGFEV	113
Gm09g00260	VDSGNEMVLSNLE-----EERKLRVLVAGGGIGGLVFALAARKGFEV	89
PA-ZE	EVRRATVASITEVPSAP-----ASTQPKLRVLVAGGGIGGLVFALAARKGFEV	106
AtABA1	VKAATALVYKEEKREA-----VTEKRRKLRVLVAGGGIGGLVFALAARKGFEV	106
GmZEP1	MVFEKDLISAIRGEGQYRGP IQIQSNLAALAEADLSEVADEVMRVGCITGDRINGLVDGVS	172
GmZEP2	VVFEKDLISAIRGEGQYRGP IQIQSNLAALAEADLSEVADEVMRVGCITGDRINGLVDGVS	157
GmZEP3	-----GQYRGP IQIQSNASADLEADLSEVADEVMRVGCITGDRINGLVDGVS	160
Gm09g00260	KVFEKDLISAVRGEGRHRGPTQLLSALAVLEITDQSVARQIWEAGCVTANRNGIADGLS	149
PA-ZE	VVFEKDLISAVRGEGRHRGPTQLLSALAVLEITDQSVARQIWEAGCVTANRNGIADGLS	166
AtABA1	IVFEKDLISAIRGEGQYRGP IQIQSNLAALAEADLSEVADEVMRVGCITGDRINGLVDGVS	166
GmZEP1	GSWYVKDFTFPTAVERGLP VTRVISRMLVQLQILARAVGEDITMNASNVVNEVDDGKQVTV	232
GmZEP2	GSWYVKDFTFPTAVERGLP VTRVISRMLVQLQILARAVGEDITMNASNVVNEVDDGKQVTV	217
GmZEP3	GSWYVKDFTFPTAVERGLP VTRVISRMLVQLQILARAVGEDITMNASNVVNEVDDGKQVTV	220
Gm09g00260	GDWFSVDFTFPTASRKRLEPTIVICRMLVQLQILVNRVGSNIHNRKSRVDFIQEENKVRV	209
PA-ZE	GSWYVKDFTFPTAVERGLP VTRVISRMLVQLQILARAVGEDITMNASNVVNEVDDGKQVTV	226
AtABA1	GSWYVKDFTFPTAVERGLP VTRVISRMLVQLQILARAVGEDITMNASNVVNEVDDGKQVTV	226
GmZEP1	ELENGQRYEGDILVVGADGIWSKVRQLFGLTEAVYSGYTCYTG IADFVPADIESVGYRVE	292
GmZEP2	ELENGQRYEGDILVVGADGIWSKVRKLFQTEAVYSGYTCYTG IADFVPADIESVGYRVE	277
GmZEP3	ELENGQRYEGDILVVGADGIWSKVRKLFQTEAVYSGYTCYTG IADFVPADIESVGYRVE	280
Gm09g00260	ILENGEQHDDLLIGADGIWSKVRKLFQTEAVYSGYTCYTG IADFVPADIESVGYRVE	269
PA-ZE	ILENGQRYEGDILVVGADGIWSKVRQLFGLTEAVYSGYTCYTG IADFVPADIESVGYRVE	286
AtABA1	VLENGQRYEGDILVVGADGIWSKVRQLFGLTEAVYSGYTCYTG IADFVPADIESVGYRVE	286
GmZEP1	LGHKQYFVSDVGGKMQWYAFHKESP GGVDSPNGKKERLEKIFEGWCDN--AVDILLAT	350
GmZEP2	LGHKQYFVSDVGGKMQWYAFHKESP GGVDSPNGKKERLEKIFEGWCDN--VIDILHAT	335
GmZEP3	LGHKQYFVSDVGGKMQWYAFHKESP GGVDSPNGKKERLEKIFEGWCDNRQFDELYHT	340
Gm09g00260	LGLLQYFVSDVGGKMQWYAFHKESP SDPFSPGKKERLEKIFEGWCDN--VIALISEH	327
PA-ZE	LGHKQYFVSDVGGKMQWYAFHKESP GGVDSPNGKKERLEKIFEGWCDN--VIDILLAT	344
AtABA1	LGHKQYFVSDVGGKMQWYAFHKESP GGVDSPNGKKERLEKIFEGWCDN--VIDILHAT	344
GmZEP1	EEBAILRRDIYDRITLFWGKGRVTLGDSVHAMQPNMGQGGCMAIEDSYQLALELDEAW	410
GmZEP2	EEBAILRRDIYDRITLFWGKGRVTLGDSVHAMQPNMGQGGCMAIEDSYQLALELDEAW	395
GmZEP3	SPQHLHGE RAVSPCHVILSNPQSQIWA NEDARLLRPNK-----VQDSYQLALELDEAW	394
Gm09g00260	PEHMLIQRDIDYRDHINHWGIGRVTLLGDAAHPMQPNMGQGGCMAIEDSYQLALELDEAW	387
PA-ZE	EEBAILRRDIYDRITLFWGKGRVTLGDSVHAMQPNMGQGGCMAIEDSYQLALELDEAW	404
AtABA1	EEBAILRRDIYDRITLFWGKGRVTLGDSVHAMQPNMGQGGCMAIEDSYQLALELDEAW	404
GmZEP1	EQSIKSSSDPDI DSSLKSYERERRLRVAIHGMARMAAMMASTYKAYLGVGLGPISEFLTK	470
GmZEP2	QOSIKSSSDPDI DSSLKSYERERRLRVAIHGMARMAAMMASTYKAYLGVGLGPISEFLTK	455
GmZEP3	QRSIKSSSDPDI DSSLKSYERERRLRVAIHGMARMAAMMASTYKAYLGVGLGPISEFLTK	454
Gm09g00260	KHGS DGS---EVISALRRYKKNRPRVRLVITASRMASQMLVNRPMIEFKFVPIHSHVIT	444
PA-ZE	KKSSETITVDVWASLSLSYENSRRLRVAIHGMARMAAMMASTYKAYLGVGLGPISEFLTK	464
AtABA1	KQSVETITVDVWVSSLKRYEERRLRVAIHGMARMAAMMASTYKAYLGVGLGPISEFLTK	464
GmZEP1	FRVHPGRVGGREFVDILMBSMLDVLGGNSKLEGRPLSCLTADKANQIHRWFEEDEA	530
GmZEP2	FRVHPGRVGGREFVDKIMBLMDNVLGGNSKLEGRPVCCRLSDKANQIHRWFEEDEA	515
GmZEP3	FRVHPGRVGGREFVDKIMBLMDNVLGGVAVDITVLSYELYSNE---ISVYLEIFP	511
Gm09g00260	MQIKRPGIHVAQALFKFTFQFVTRMIAQ-----RPSVIFN	473
PA-ZE	FRVHPGRVGGREFVDKIMBLMDNVLGGNSKLEGRSPSCLTADKANQIHRWFEEDEA	524
AtABA1	FRVHPGRVGGREFVDILMBSMLDVLGGNSKLEGRPPSCLTADKANQIHRWFEEDEA	524
GmZEP1	LERAINSEWILLDHCDGTGLSKPISLSRNEMKPTLCSAPMQDNSFSVVTSSPQVSPTH	590
GmZEP2	LERAINSEWILLDHCDGAGPTKPCITQDEMKPCHTSSMQQKHPSSSIIPLPQVSOQH	575
GmZEP3	MHRITGKLVGDLVR-----ILKDTWFECPKQ-----	538
Gm09g00260	-----HCLW-----	474
PA-ZE	LERAINSEWYITLCCQDNDASQLICINRDEKNPTLCSAPHGVSIISIAIPKPVSEM	584
AtABA1	LERITKSEWYITLPHEDDCCVSETLQITKDEDQPIVGESEPDQIFMRIVIPSSQVSKM	584
GmZEP1	ARINYKDGAFPIIDLRSEHGIVITDNECKQVTPPNYPARVRSVDTQFGS-EKRVSRVR	649
GmZEP2	ARINYKDGAFPIIDLRSEHGIVITDNECKRVRTPPNYPARVRSVDTVEFGS-DKASYRVK	634
GmZEP3	-MLVSCDWSILAKQLLTHK-----RPSVIFN-----	564
Gm09g00260	-----RPSVIFN-----	564
PA-ZE	ARISYKDGAFPIIDLRSEHGIVYTDNECKRVRTPPNYPARVRSVDTVEFGS-QVAVSRVR	643
AtABA1	ARVIYKDGAFPIIDLRSEHGIVYTDNECKRVRTPPNYPARVRSVDTVEFGS-DKAAVSRVR	644
GmZEP1	VITSSVPRVSENE---STLALQGV	669
GmZEP2	VITRSASSESEKE---GTKLYQKV	654
GmZEP3	-----	
Gm09g00260	-----	
PA-ZE	VTKS SPSGVEK-----EGILQAA	661
AtABA1	VIRKTPKSTRKNESNNDKLLQTA	667

**E**

GmVDE1	RLELAKCLSNPSCAAHIAQLQTCNHRPDETECQIKCGDLFENSVVDEFNECAVSRKKCVF	228
GmVDE2	RLELAKCLSNPSCAAHIAQLQTCNHRPDETECQIKCGDLFENSVVDEFNECAVSRKKCVF	208
Gm03g41420	RLELAKCLSNPSCAAHIAQLQTCNHRPDETECQIKCGDLFENSVVDEFNECAVSRKKCVF	168
CsVDE	RLELAKCLSNPSCAAHIAQLQTCNHRPDETECQIKCGDLFENSVVDEFNECAVSRKKCVF	213
NPQ1	RLELAKCLSNPSCAAHIAQLQTCNHRPDETECQIKCGDLFENSVVDEFNECAVSRKKCVF	187
GmVDE1	KKSDVGEFFPAPNPDVLVH SFNLADEFS GKWF IIS GLNP TFDTFDCQLHEFHTE-SNKL VGN	287
GmVDE2	KKSDVGEFFPAPNPDVLVH SFNLADEFS GKWF IIS GLNP TFDTFDCQLHEFHTE-SNKL VGN	267
Gm03g41420	KKSDVGEFFPAPNPDVLVH SFNLADEFS GKWF IIS GLNP TFDTFDCQLHEFHTE-SNKL VGN	227
CsVDE	QKSDVGEFFVPPDPA LLVK SFNLADEFS GKWF IIS GLNP TFDTFDCQLHEFHTE-SNKL VGN	272
NPQ1	RKSDVGEFFPAPNPDVLVH SFNLADEFS GKWF IIS GLNP TFDTFDCQLHEFHTE-SNKL VGN	247
GmVDE1	LSWRINTPDAGEITRSARQRFVQDPSVPGILYNHDNEYLHYODDWYLLS SKIENKPD D Y V	347
GmVDE2	LSWRINTPDAGEITRSARQRFVQDPSVPGILYNHDNEYLHYODDWYLLS SKIENKSD D Y V	327
Gm03g41420	LSWRINTPDAGEITRSARQRFVQDPSVPGILYNHDNEYLHYODDWYLLS SKIENKPD D Y V	287
CsVDE	LSWRINTPDGGEFTRSARQRFVQDPIHPGILYNHDNEYLHYKDDWYLLS SKIENKPD D Y V	332
NPQ1	LSWRINTLDSGEFTRSARQRFVQDENQPGVLYNHDNEYLHYODDWYLLS SKIENKPE D Y I	307
GmVDE1	FVYYRGRNDAWDGYGGAVIYTRSAVLPESI VPELEKAAKSVGRDFSTFI RTDN TCGPEPS	407
GmVDE2	FVYYRGRNDAWDGYGGAVYTRSAVLPESI VPELEKAAKSVGRDFSKFI RTDN TCGPEPS	387
Gm03g41420	FVYYRGRNDAWDGYGGAVIYTRSAVLPESIVPELEKAAKSVGRDFSTFI RTDN TCGPEPS	347
CsVDE	FVYYRGRNDAWDGYGGAVIYTRSAVLPESI VPELEKAAKSVGRDFSKFI RTDN TCGPEPS	392
NPQ1	FVYYRGRNDAWDGYGGAVYTRSSVLPESI VPELEKAAKSVGRDFSTFI RTDN TCGPEPA	367
GmVDE1	LVERIEKKVEEGEETIIKREVEQLEEEVEK----VGKTEATLFOQLAEGFKVFEDEEENFL	463
GmVDE2	LVERIEKKVEEGEETIIKREVEQLEEEVEK----VGKTEATLFOQLAEGFKMFEDEEENFL	443
Gm03g41420	LVERIEKKVEEGEETIIKREVEQLEEEVEK----VGKTEATLFOQLAEGFKVFEDEEENFL	403
CsVDE	LVERIEKKVEEGEETIIKREVEQLEEEVEK----VGKTEATLFOQLAEGFKELQKDEEENFL	448
NPQ1	LVERIEKIVEEGEETIIKREVEQLEEEVEKVEKVGRTETLFOQLAEGFKELQKDEEENFV	427
GmVDE1	RGLSKEEMELLDGLKMEAGEVEKLFGRALPVRKLR	498
GmVDE2	RGLSKEEMELLDGLKMEAGEVEKLFGRALPVRKLR	478
Gm03g41420	RGLSKEEMELLDGLKMEAGEVEKLFGRALPVRKLR	438
CsVDE	RGLSKEEMDLLSELKMEASEVENLFGRALPVKKLR	483
NPQ1	RGLSKEEMELLDGLKMEAGEVEKLFGRALPVRKLR	462

F

GmNCED1a	KTIDPRVQLAGNEAPVPEHPVQHSLSLVVIGTIPDRAUNGVYLRNGANPDELPKKCGHLEFDGD	140
GmNCED 1b	KTIDPRVQLAGNEAPVPEHPVQHSLSLVVIGTIPDRAUNGVYLRNGANPDELPKKCGHLEFDGD	172
NCED 2	KTIDPRVQLAGNEAPVPEHPVQHSLSLVVIGTIPDRAUNGVYLRNGANPDELPKKCGHLEFDGD	168
vp14	STADPAVQLAGNEAPVGERPPVHSLPVSGRIPPEIDGVTARNGANPDELPKKCGHLEFDGD	174
→		
GmNCED 1a	GMVHGVKENDGPNAS--YACRFETEORLMOERKLGDPVFPKKAICGLHCHSGIARLMLFYARG	199
GmNCED 1b	GMVHGVKENDGPNAS--YACRFETEORLMOERKLGDPVFPKKAICGLHCHSGIARLMLFYARG	231
NCED 2	GMVHGVKENDGPNAS--YACRFETEORLMOESQLGDPVFPKKAICGLHCHSGIARLMLFYARG	227
vp14	GMVHGLRDRNGPRAESYACRFETEORLMOERKLGDPVFPKKAICGLHCHSGIARLMLFYARG	234
←		
GmNCED 1a	LCCIVDHRSCRCGSANAGLVFENGILLAMSEDDLPYELEITPARGDLETICRYSEIDQLNSS	259
GmNCED 1b	LCCIVDHRSCRCGSANAGLVFENGILLAMSEDDLPYELEITPARGDLETICRYSEIDQLNSS	291
NCED 2	LCCIVDHRSCRCGSANAGLVFENGILLAMSEDDLPYELEITPARGDLETICRYSEIDQLNSS	287
vp14	ACGLVDPSTACTGVANAGLVFENGILLAMSEDDLPYHVVVDIDGDLLETVGRYDFICQLGCA	294
GmNCED 1a	MIADPKVDDPISGELFELSYDVT S-KYLNKYEFE SPDEEKSPDVEIPLDAPMTMDFFAITE	318
GmNCED 1b	MIADPKVDDPISGELFELSYDVT S-KYLNKYEFE SPDEEKSPDVEIPLDAPMTMDFFAITE	350
NCED 2	MIADPKVDDPISGELFELSYDVT S-PCVNYENE SPK-EVKSPEVEIPLDAPMTMDFFAITE	346
vp14	MIADPKVDDPATGELFELSYDVTIRPYLNKYEFE RPD-GAKSDDVEIPLDAPMTMDFFAITE	353
GmNCED 1a	NFVWIPDQVVVFKLGEMVCGGSPVVIYDCEKNSREGVLPKRYASDASSIWWVDSPIETEREFE	378
GmNCED 1b	NFVWIPDQVVVFKLGEMVCGGSPVVIYDCEKNSREGVLPKRYASDASSIWWVDSPIETEREFE	410
NCED 2	NFVWIPDQVVVFKLGEMVCGGSPVVIYDCEKNSREGVLPKRYASDASSIWWVDSPIETEREFE	406
vp14	NFVWVDPDQVVVFKLGEMVCGGSPVVIYDCEKNSREGVLPKRYASDASSIWWVDSPIETEREFE	413
GmNCED 1a	INRWEEEDTDEVVVIGSCMTPPDSIFKITEESIKSVLTEIRLNRSEKSRRRVUVE---E	435
GmNCED 1b	INRWEEEDTDEVVVIGSCMTPPDSIFNDRERIKSVLTEIRLNRSEKSRRRVUVE---E	467
NCED 2	ENRWEEEDTDEVVVIGSCMTPPDSIFNDRSENLITLITEIRLNRSEKSRRRVUVE---N	463
vp14	INRWEEEDTGEVVVIGSCMTPPDSIFNESDERIBSVLTEIRLNRSEKSRRRVUVE---SQQ	473
GmNCED 1a	INLEAGMVRNRKRLGRNTRERYLQVAEPWPKVSGVAKVILSCEVKKRMEYGERREGGEPFE	495
GmNCED 1b	INLEAGMVRNRKRLGRNTRERYLQVAEPWPKVSGVAKVILSCEVKKRMEYGERREGGEPFE	527
NCED 2	INLEAGMVRNRKRLGRNTRERYLQVAEPWPKVSGVAKVILSCEVKKRMEYGERREGGEPFE	523
vp14	INLEAGMVRNRNLLGRETTRERYLQVAEPWPKVSGVAKVILSCEVKKRMEYGERREGGEPFE	533

**G**

GmCCD7a	-----YQAPPIHNT---BTYIPPPIRSP-PVHQTPPSIPSTIKPRAI	38
GmCCD7b	-----YQAPPIHNT---BTYIPPPIRSP-PVHQTPPSIPSTIKPRAV	38
Gm01g35910	-----MVPMP-----ILVTTTPPPATSAPSPILLNAVSTLKEEKPTQ	37
SlCCD7	MDLQFVSLPPNSKTKAKMOAACHNINNIPEKLLPPAKLSTVAMSPSQTLPQSHVARRAI	60
AtCCD7	-----VSLP-----IPEKFTPEPKLSPPIHQQTPPELAPERRAAITSI	36
GmCCD7a	SISAPNIPPIAVPIPIPIVEPDDSNAAAYDYQELFISORSEATRPVTEI STVEGATPDAFPSC	98
GmCCD7b	SISAPNTPVAVPIPIPIVEPDDSNAAAYDYQELFISORSETSPQVTEI STVEGVTPDAFPSC	98
Gm01g35910	TLKTPTTTTQKATIKPSPSSSTKTKTPQQRVELKRTNSWVPALMNAFDDIINNFIDPP	97
SlCCD7	TTTTSPTEHEVYTP---EIDDTVFAAYDYQELFVTSORSEATEPVSIRVVEGSIPTDFPSC	116
AtCCD7	SDPTTGLGRFTIID---ESTSSAFRDYQSLVTSORSETIEPVIKPIEGSIIVNFPSC	92
GmCCD7a	IIYLLTGPGLIADHGGSTVHPLDGHGYLRAFTEDNNTVTKNVKYMAYKYIKTDAHVEEHDPKT	158
GmCCD7b	IIYLLTGPGLIADHGGSTVHPLDGHGYLRAFTEDN-ATKNVKYMKYIKTDAHVEEHDPKT	157
Gm01g35910	LKPSIDPRHVLSONFADVHSLPPT---ECEVEGSLPP---SLDGAYTRNGPNPQ-----	146
SlCCD7	IIYLLTGPGLIADHGGSTVHPLDGHGYLRAFTEDNSTGQ-VKFMARYIETEAQTEERDPVS	175
AtCCD7	IIYLLAGPGLITDDHGGSTVHPLDGHGYLRAFTEDNKRK-ATFTAKYIKTEAKKEEHDPT	151
GmCCD7a	NKWFTHRGPFSVLKGGKVGNTKVMKNVANTSVLKWGEKLLCWEGGEPYEIQAGHIDT	218
GmCCD7b	DKWFTHRGPFSVLKGGKVGNTKVMKNVANTSVLKWGEKLLCWEGGEPYEIQAGHIDT	217
Gm01g35910	----FLRGRVYHLEDCDGLHALRISHGKKPLCSRYVVKYKYMENADGFPILPNVFSG	202
SlCCD7	GKWFTHRGPFSVLKGGKVGNTKVMKNVANTSVLQWCELEFCWEGGDDPEIDSKHMT	235
AtCCD7	DIWRFTHRGPFSVLKGGKRGNTKVMKNVANTSVLKWAGRILLCWEGGEPYETESGSLDT	211
GmCCD7a	IGRNMMD-GADMENHDQSKGG---GGGVVVEVAANLLKPILYGVFKMPPRRLLSHYKVDS	275
GmCCD7b	IGRNMMD-GADLENHDESKDG---VDVVEVAANLLKPILYGVFKMPPRRLLSHYKVDS	272
Gm01g35910	LNSLVAASAARGSLAARFLTG-----QFNPAANGIGLANTSLAFGNRDFALGESD-	252
SlCCD7	LGLKLELKH-NSDQVLEDKKIS---HSDLELVAAQLLKPILYGVFKMSPRRLLSHYKIDT	290
AtCCD7	VGRNVENNNGCESDDESDDRDLSGHDIDTADAILLKPILQGVFKMPPRRLLSHYKVDG	271
GmCCD7a	RRNRLLTVSNCAEDMLLPRSNFTFSEYDSNENVVQKQVFRIPDHLMIHDWAFDTTHYIVF	335
GmCCD7b	RRNRLLTVSNCAEDMLLPRSNFTFSEYDSNENVVQKQVFRIPDHLMIHDWAFDTTHYIVF	332
Gm01g35910	-LPYVNVTPDGDIEDLGR---HDEEGKLTFSMTAHPKIDP-DTAECAFRYGPPVPPF	305
SlCCD7	RRNRLLTVSNCAEDMLLPRSNFTFSEYDSNENVVQKQVFRIPDHLMIHDWAFDTTHYIVF	350
AtCCD7	RRNRLLTVSNCAEDMLLPRSNFTFSEYDSNENVVQKQVFRIPDHLMIHDWAFDTTHYIVF	331
GmCCD7a	ANRIKLDVLSMAAVYGMSPMVSALRVNPSKSTSPIYLLPRFPDKNN-----GKERDWR	389
GmCCD7b	ANRIKLDVLSMAAVYGMSPMVSALRVNPSKSTSPIYLLPRFPDKNN-----DKERDWR	387
Gm01g35910	LTIFYRFVGGKKHEDVPIFSMLTSPFTHDFAIKKYALFCDIQGLN-----PLDMI	357
SlCCD7	GNRIKLDIPGSMFAVCGHSPMVSALSHPNSKSTSPIYLLPRFRNN-----NVERDWR	402
AtCCD7	ANRVKLPITGSLAAMCGMSPMVSALSHPNSKSTSPIYLLPRFSDKYS-----RGERDWR	385
GmCCD7a	VPVEAPSQLMILLHVGNAFEVR-HPHGHLDIKIQAAACSYQWFNFSKIFGYDWQSKKLDPS	448
GmCCD7b	VPVEAPSQLMILLHVGNAFEVR-YPHRHLDIKIQAAACSYQWFNFSKIFGYDWQSKKLDPS	446
Gm01g35910	SGGSPVGSVASKVPRIGILPRDAKDESMKWFEPGFEN---IHALNAWEDEGRIVVL	413
SlCCD7	KPVEAPSQLMILLHVGNAFEVIDEQGHLLNIOIQASCSYQWFNFKIFGYDWQSKKLDPS	462
AtCCD7	VPVEVSSQLMILLHVGNAFEVR-EDHGHLKIQIQASCSYQWFNFKIFGYDWQSKKLDPS	444
GmCCD7a	IMNVKGG-SKLLPHLVQVSTKLDSDYNCQECDDVKKPK-KKKKSDFPATNPTSGKKNKY	506
GmCCD7b	IMNVKGG-NELLPHLVQVSTKLDSDYNCQECDDVKKPK-KKKKSDFPATNPTSGKKNKY	504
Gm01g35910	VAPNILSMEHATERMELVHAMVEKVRIELDGTITRQPVVSARMDFAVINPAVGGKKNRF	473
SlCCD7	MMNVEEGEEKLLPHLVQVSTKLDKSGNCTKCSVNDLMPENKARDPFAVMEPESGKKNRY	522
AtCCD7	VMNENRGDDKLLPHLVQVSTKLDSTGNCNQCDDVKKPK-GWKKPSDFPATNPTSGKKNKY	503
GmCCD7a	IYAATTLGSRKTLPCFPFDIVVKIDLETESVAQIWTAGSRRFTGEPVFEKGY----Q	561
GmCCD7b	IYAATTLGSRKTLPCFPFDIVVKIDLES-DNSAQIWTAGSRRFTGEPVFEKGD----E	558
Gm01g35910	VYAAVGDPMPKISGVVKLDVSKGEE--RRDCTVGRMYEGCCYGEPPFVAREEGG---E	528
SlCCD7	IYAATCTGSRQALPHFPFDIVVKIN--AVDKSVQKWSAGRRFTGEPVFEKGTN----K	576
AtCCD7	MYSAASSGRSELPHFPFDIVVKID--LDSNLQRTVSTGARRFTGEPVFEKHSVEEGEE	561
GmCCD7a	EDDGYLLVVEYAVSMNRCYLVTLLDPKRIADNALTARTEIE--SHLNFPLGFHGFVA-N	618
GmCCD7b	EDDGYLLVVEYAVSMNRCYLVTLLDPKRIADNALTARTEIE--SHLNFPLGFHGFVA-N	615
Gm01g35910	EDDGYLLVSYVHDERKGESRFLVMDAK--SPELDVVAAVRLPR-----RVPYGFHGLFVKE	581
SlCCD7	EDDGYLLVVEYAVSITRCYLVTLLDAQTIGEKNEVVARLEVE--RHLNFPLGFHGFVWPTN	634
AtCCD7	EDDGYIVVVEYAVSVERCYLVTLLAKIGESDAVVSVMKVIYIAKINYLICVHSFYEDRN	621
GmCCD7a	-----	
GmCCD7b	-----	
Gm01g35910	SELRKVSL-----	590
SlCCD7	SSLANLQKIESKCKNSWSMMKDNMVKLGQ	663
AtCCD7	IAFHLSHK-----	629

**H**

GmCCD8a	--MLVPD---PHINISLRNKG-IYKGR-----NINLRDLTKAVASPSLPVLPPTTE	49
GmCCD8b	--MLVPD---PHNNISLRNKGFIYKQH-----GINLRDFTKAVASPSLPVLPPTTE	50
Gm08g10190	LNSSTCAKQPFPHPLSSSFSSSSSLTRDLGFPTRFQKKPIRKVVDCALHSPVVIHFPPNQF	65
PhCCD8	MASFAS---RKYCNKILPERFDHGRKDEPHFGKIKINEKTKKMDLKLVTNVSQLTV	57
AtCCD8	MASLIITKAMSHHHVLSSTRITLYSDNSIGDQQIKTKPQVPHRLFARRIFGVRAVIN	60
→		
GmCCD8a	ERP---GITHD--CHHHIVAWTSIPQERWEGELVQGGIPLWLEGGTYLRNGPQGMWHIG	102
GmCCD8b	ERK---ETTADDRHHHHIVAWTSIPQERWEGELVQGGIPLWLEGGTYLRNGPQGMWHIE	106
Gm08g10190	YKNPLNQFPFKTTDPVQVIAGNFAPVPEHPVHSLPVIQTTPDAINGVYLRNGANPLFKP	125
PhCCD8	I VS---PLDQEVVKEKLLAAWTSVQERWEGELVVEGEP L W L S G T Y L R N G P G L W H V G	113
AtCCD8	SAAPSPLPEKEKVEGERRCHVAVTSVQERWEGELVQGGIPLWLEGGTYLRNGPGLWHIG	120
←		
GmCCD8a	DYNFRHLFDGYATLVRLGFRNGRLVAGHROIESEAYRAAKKMKKLCYREFSEVPK----A	158
GmCCD8b	DYNFRHLFDGYATLVRLGFRNGRLVAGHROIESEAYRAAKKMKKLCYREFSEVPK----A	162
Gm08g10190	KAGH-HLFDGDMVHAVKENDG-TASYACRFTEQRLMQERKLGKPVFPKAIQELHGHSG	183
PhCCD8	DYNFRHLFDGYATLVRLGFRNGRLVAGHROIESEAYRAAKKMKKLCYREFSEVPK----P	169
AtCCD8	DHDFRHLFDGYSTLVKLVQEDGGRLFAAHLLESDAYRAAKKMKKLCYREFSEVPKSVIIN	180
GmCCD8a	ANFLAVVGGELASLFSGASLTDNANTGVVVKLGDGRVVCLETETQKGSIVINPETLLETIGKFE	218
GmCCD8b	ANFLAVVGGELASLFSGASLTDNANTGVVVKLGDGRVVCLETETQKGSIVINPETLLETIGKFE	222
Gm08g10190	--IARIMLFYARGLCGIVDHRRGAGAANAGLVFENGKLLAMS EDDFPYELRITASGDLETI	242
PhCCD8	ENFLSIVTDMKALFSGASLTDNANTGVVVKLGDGRVVCLETETQKGSIVINPETLLETIGKFE	229
AtCCD8	KNPFSIVTEIVRLFSGASLTDNANTGVVVKLGDGRVVCLETETQKGSIVVDHETLLETIGKFE	240
GmCCD8a	YSDSLGG-LIHSAPHIVTDEEFTLLLPDLVRAGYLVVRMEPGTNERRVIGRVNCRGGP-A	277
GmCCD8b	YSDSLGG-LIHSAPHIVTDEEFTLLLPDLVRAGYLVVRMEPGTNERRVIGRVNCRGGP-A	280
Gm08g10190	GRYSFHGQLNSMKVDPVSGELFTLSYDVTSKLLKYFHFSPEDERKSPDIEIPLDAPT--	300
PhCCD8	YSDSLGG-LIHSAPHIVTDEEFTLLLPDLVRAGYLVVRMEPGTNERRVIGRVNCRGGP-A	287
AtCCD8	YDDVLSDDHMTQSAPHIVTDEEFTLLLPDLVRAGYLVVRMEPGTNERRVIGRVNCRGGP-A	300
GmCCD8a	PGWVHSFPPVTOHYVIVPEMPLRYCAQNLLKAEP TPLYKFEWHPESRAF MHVMCKTSGKIV	337
GmCCD8b	PGWVHSFPPVTOHYVIVPEMPLRYCAQNLLKAEP TPLYKFEWHPESRAF MHVMCKTSGKIV	340
Gm08g10190	--MTHDFAITENFVVIPDQVVFVVLKLGEMIKGSSPVIIDGEEKSRFGILPKYASDAS-SI-	357
PhCCD8	P G W V H S F P P V T O H Y V I V P E M P L R Y C A Q N L L K A E P T P L Y K F E W H P E S R A F M H V M C K T S G K I V	347
AtCCD8	P G W V H S F A V T E N Y V V I P E M P L R Y S V K N L L R A E P T P L Y K F E W C P Q D G A F I H V M S K T G E V V	360
GmCCD8a	ASVEVPDFVTFHFINAYEE-QDE DGRVTAIADCCENSDTILDKRLRLQLNRSNGEDV	396
GmCCD8b	ASVEVPDFVTFHFINAYEE-QDE DGRVTAIADCCENSDTILDKRLRLQLNRSNGEDV	399
Gm08g10190	VWVDSIDTEFDFHFNAWE---RDKDEVVVISSCMTPE-----PDSIFNDREE	401
PhCCD8	ASVEVPDFVTFHFINAYEE-KDE DGRVTAIADCCENSDTILDKRLRLQLNRSNGEDV	406
AtCCD8	ASVEVPDFVTFHFINAYEEDKNGDGKATVIADCCENADTILDKRLRLQLNRSNGEDV	420
GmCCD8a	LPDARVGRFRIPLDGSPYGLDAALEPNEHGRMDMCSINPNYLGGKYRYAYACGAORPC	456
GmCCD8b	LPDARVGRFRIPLDGSPYGLDAALEPNEHGRMDMCSINPNYLGGKYRYAYACGAORPC	459
Gm08g10190	RLKSVLTVLNMRSKGARRRVLVEEMN----LEAGMVNRKRIGRKRFRAYLC-IAEPW	453
PhCCD8	LPDARVGRFRIPFDGSPYGLDAALDPNEHGRMDMCSINPNYLGGKYRYAYACGAORPC	466
AtCCD8	LPDARVGRFRIPLDGSKYGLLETAVEAEKHGRMDMCSINPNYLGGKYRYAYACGAORPC	480
GmCCD8a	NFPNILLTKIDFELKAKNWHEEGAVPSEPPFFVARPGATEDDGVVISIVSEKNGEGYFALV	516
GmCCD8b	NFPNILLTKIDFELKAKNWHEEGAVPSEPPFFVARPGATEDDGVVISIVSEKNGEGYFALL	519
Gm08g10190	PKVSGVAKVDLESGEVVKREYGRRRFGGEPFFLPRGNGNEDEGYVMAFVHDETWSLQI	513
PhCCD8	NFPNILLTKIDLFEKAKNWHEEGAVPSEPPFFVARPGATEDDGVVISISDKNGEGYFALI	526
AtCCD8	NFPNALLSKVDLVEKVKNWHEHGMIPSEPPFFVARPGATEDDGVVISIVSEKNGEGYFALL	540
GmCCD8a	LDGSTFEEIARAKFPYGLPYGLHGCWVPKE	546
GmCCD8b	LDGSTFEEIARAKFPYGLPYGLHGCWVPKE	549
Gm08g10190	LNALD--LKEAIVMLPSRVPYGFHGTVE	541
PhCCD8	LDGSTFEEIARAKFPYGLPYGLHGCWVPKK	556
AtCCD8	LDGSTFEEIARAKFPYGLPYGLHGCWVPKD	570

I

GmCCD1a	MGD DGGKMSGEGGIVKVEPKPSNGFTSKVVDLLEKLVVKELMDSSLTHHYLTGTFAPVP-	59
GmCCD1b	MGD DGGKLSGEGGIVKVEPKPSNGFTSKVVDLLEKLVVKELMDSSLPHHYLTGTFAPVS-	59
MtCCD1	--MESEKIG--GGIVKVEPKPSNGFTSKVVDLLEKLVVKELMDSSLPHHYLTGTFAPVKD	56
RtCCD1	---MREKLSDGSSIISVHPRPKSGFSSKLDLLERLVVKELMDSSLPHHYLTGTFAPVTRD	57
GmCCD1a	ETPPDKDLPVKGYLPDCLNGEFVVRVGNPKKESPVAGYHCSGICFSNFDLLNCRMTHGLRI	119
GmCCD1b	ETPPDKDLPVKGYLPDCLNGEFVVRVGNPKKESPVAGYH-----FDGD--GMIHGLRI	110
MtCCD1	ETPPDKDLPIKGYLPDCLNGEFVVRVGNPKKESPVAGYH-----FDGD--GMIHGLRI	107
RtCCD1	ETPPDKDLPVHGYLPDCLNGEFVVRVGNPKKESPVAGYH-----FDGD--GMIHGVRI	108
GmCCD1a	KDGKATVSRVYKTSRLKQEEYFGGSKFMKIGDKGLFGLLMVNIHMLRTKQKVLDAVSYG	179
GmCCD1b	KDGKATVSRVYKTSRLKQEEYFGGSKFMKIGDKGLFGLLMVNIHMLRTKQKVLDAVSYG	170
MtCCD1	KDGKATVSRVYKTSRLKQEEYFGGSKFMKIGDKGLFGLLMVNIHMLRAKQKVLDAVSYG	167
RtCCD1	KDGKATVSRVYKTSRLKQEEYFGGAKFMKIGDKGLFGLLMVNIHMLRTKQKVLDAVSYG	168
GmCCD1a	TGTANTALVYHHGKLLALSERDKPYRIKIMTR-DWATPSLLIQ-----	221
GmCCD1b	TGTANTALVYHHGKLLALSERDKPYRIKVEEDGDLQTLGMLDYDKRLEGHSFTAMPKVDFP	230
MtCCD1	HGTANTALVYHHGKLLALSERDKPYRIKVEEDGDLQTLGMLDYDKRLEGINFTAMPKVDFP	227
RtCCD1	NGTANTALVYHHGKLLALSERDKPYRIKVEEDGDLQTLGMLDYDKRLEHSFTAMPKVDFP	228
GmCCD1a	KGDVY-----FYRVISKDGKGMDDPVPITVSDPIMHDFAITERYRIFLDLPLIF	270
GmCCD1b	TGEMFTFGYRHTPPYITFYRVISKDGKGMDDPVPITVSDPIMHDFAITERYRIFLDLPLIF	290
MtCCD1	TGEMFTFGYSHTPPYITFYRVISKDGKGMDDPVPITVSDPIMHDFAITERYRIFLDLPLIF	287
RtCCD1	TGEMFTFGYSHTPPYITFYRVISKDGKGMDDPVPITVSDPIMHDFAITERYRIFLDLPLIF	288
GmCCD1a	RPKEMVKKIKTLIESFDS TKKARF GVLPRY AKDEKLRWFE LPNCF IEHNANAWEE EDEVV	330
GmCCD1b	RPKEMVKKIKTLIESFDS TKKARF GVLPRY AKDEKLRWFE LPNCF IEHNANAWEE EDEVV	350
MtCCD1	RPKEMVKKIKTLIESFDS TKKARF GVLPRY AKDEKLRWFE LPNCF IEHNANAWEE EDEVV	347
RtCCD1	RPKEMVKEKIKTLIESFDS TKKARF GVLPRY AKDELIRWFE LPNCF IEHNANAWEE EDEVV	348
GmCCD1a	LI TCRL INPNLD VGGTAKKLENF S NEL YEMRFNMKTGEASQKLSASAVDFPRVNESY	390
GmCCD1b	LI TCRL INPNLD VGGTAKKLENF S NEL YEMRFNMKTGEASQKLSASAVDFPRVNESY	410
MtCCD1	LI TCRL INPNLD VGGTAKKLENF S NEL YEMRFNMKTGEASQKLSASAVDFPRVNESY	407
RtCCD1	LI TCRL INPDLVSGVAKKLENF S NEL YEMRFNMKTGEASQKLSASAVDFPRVNECY	408
GmCCD1a	TGRKORYVYGTLDSTAKV TGI IKFDLHAE PDHGKT KLEV GGNVVG YDLGPGRYGSEAV	450
GmCCD1b	TGRKORYVYGTLDSTAKV TGI IKFDLHAE PDHGKT KLEV GGNVVG YDLGPGRYGSEAV	470
MtCCD1	TGRKORYVYGTLDSTAKV TGI IKFDLHAE PDHGKT KLEV GGNVVG YDLGPGRYGSEAV	467
RtCCD1	TGRKORYVYGTLDSTAKV TGI IKFDLHAE PDHGKT KLEV GGNVVG YDLGPGRYGSEAV	468
GmCCD1a	YVPRVP GTD SEEDDGYL ICFVHDENTGKS FVHVINAKTMS RD PVAVVEL PHRVPYGFHAF	510
GmCCD1b	YVPRVP GTD SEEDDGYL ICFVHDENTGKS FVHVINAKTMS RD PVAVVEL PHRVPYGFHAF	530
MtCCD1	YVPRVP GTD SEEDDGYL ICFVHDENTGKS FVHVINAKTMS RD PVAVVEL PHRVPYGFHAF	527
RtCCD1	YVPRET ---REEDDGYL ICFVHDENTGKS CVTVIDAKTMS REPVAVVEL PHRVPYGFHAF	525
GmCCD1a	FVTE EQLQE-----TGKL-----	523
GmCCD1b	FVTE VCVTTTFIYSYIAFL PLELSSITLAKLQCLFSNHFLRVIV	574
MtCCD1	FVTE EQLQE-----QAKL-----	540
RtCCD1	FVTE EQLQE-----QLLI-----	538



**Supplemental Figure S1. Comparison of the amino acid sequences of  $\beta$ -carotene hydroxylases, zeaxanthin epoxidases (ZEPs), violaxanthin de-epoxidases (VDEs), 9-cis-epoxycarotenoid dioxygenase 1s (NCED1s), carotenoid cleavage dioxygenases 7s, carotenoid cleavage dioxygenases 8s, and carotenoid cleavage dioxygenases 1s.**

(A) Sequence alignment of  $\beta$ -carotene hydroxylases. The conserved histidine motifs are indicated above the sequence. Identical amino acid residues are shaded black and conserved amino acid residues, gray. Dashes indicate gaps introduced to optimize alignment. . The ORFs of GmBCH4 and GmBCH5 have a high degree of similarity to functionally confirmed  $\beta$ -carotene hydroxylases, e. g., 79 % and 73 % to CaCRTR-B from *Capsicum annuum* (Simkin et al., 2008), respectively. A solid line above amino acid sequence indicates the region of *GmBCH(1+3)* probe for *in situ* hybridization (Fig. 3, D to E) and a double line above amino acid sequence indicates the region of *GmBCH2* probe for *in situ* hybridization (Fig. 3, I to M). GenBank accession numbers: *Glycine max GmBCH1* (AY575953); *Glycine max GmBCH2* (BT093388); *Glycine max GmBCH3* (BT098487); *Glycine max GmBCH4* (JF970190); *Glycine max GmBCH5* (JF970191); *Coffea arabica CaCRTR-B* (ABA43903); *Arabidopsis thaliana AtCHY1* (NP\_194300).

Expression of *GmBCH4* (B) and *GmBCH5* (C), respectively, in soybean tissues including 27-day-old nodules. Transcript levels were determined by real time RT-PCR and normalized. Data are representative of three independent experiments. Bars represent standard deviations (n=3). L, leaf; S, stem; F, flower; R, root; N, Nodule. GenBank accession numbers: *Glycine max GmBCH4* (JF970190); *Glycine max GmBCH5* (JF970191).

(D) Amino acid sequence alignment of zeaxanthin epoxidases (ZEPs). Identical amino acid residues are shaded black and conserved amino acid residues are shaded gray. Dashes indicate gaps for optimizing alignment. Primers used in real-time PCR are indicated by arrows. GenBank accession numbers: *Glycine max GmZEP1* (HM366149);



*Glycine max GmZEP2* (HQ875555); *Glycine max GmZEP3* (HQ875556); *Prunus armeniaca PA-ZE* (AAD42899); *Arabidopsis thaliana AtABA1* (NP\_851285).

(E) Amino acid sequence alignment of violaxanthin de-epoxidases (VDEs). Partial amino acid sequences of VDEs are aligned. Identical amino acid residues are shaded black and conserved amino acid residues are shaded gray. Dashes indicate gaps for optimizing alignment. Primers used in real-time PCR are indicated by arrows. GenBank accession numbers: *Glycine max GmVDE1* (HQ875559); *Glycine max GmVDE2* (HQ875560); *Citrus sinensis CsVDE* (ADH82116); *Arabidopsis thaliana NPQ1* (NP\_172331).

(F) Amino acid sequence alignment of 9-cis-epoxycarotenoid dioxygenase 1s (NCED1s). Partial amino acid sequences of NCED1s are aligned. Identical amino acid residues are shaded black and conserved amino acid residues are shaded gray. Dashes indicate gaps for optimizing alignment. Primers used in real-time PCR are indicated by arrows. GenBank accession numbers: *Glycine max GmNCED1a* (HQ875557); *Glycine max GmNCED1b* (HQ875558); *Phaseolus vulgaris NCED2* (AAY82457); *Zea mays viviparous14 (vp14)* (NP\_001105902).

(G) Amino acid sequence alignment of carotenoid cleavage dioxygenases 7s. Identical amino acid residues are shaded black and conserved amino acid residues are shaded gray. Dashes indicate gaps for optimizing alignment. Primers used in real-time PCR are indicated by arrows. GenBank accession numbers: *Glycine max GmCCD7a* (HM366150); *Glycine max GmCCD7b* (JF894312); *Solanum lycopersicum SICCD7* (ACY39882); *Petunia x hybrida RhCCD7* (ACY01408); *Arabidopsis thaliana AtCCD7* (NP\_182026).

(H) Amino acid sequence alignment of carotenoid cleavage dioxygenases 8s. Identical amino acid residues are shaded black and conserved amino acid residues are shaded gray. Dashes indicate gaps for optimizing alignment. Primers used in real-time PCR are indicated by arrows. GenBank accession numbers: *Glycine max GmCCD8a*

(HM366151); *Glycine max GmCCD8b* (JF894313); *Petunia x hybrida PhCCD8* (AAW33596); *Zea mays ZmCCD8* (ACR33785); *Arabidopsis thaliana AtCCD8* (NP\_195007).

(I) Amino acid sequence alignment of carotenoid cleavage dioxygenases 1s. Identical amino acid residues are shaded black and conserved amino acid residues are shaded gray. Dashes indicate gaps for optimizing alignment. Primers used in real-time PCR are indicated by arrows. GenBank accession numbers: *Glycine max GmCCD1a* (FK009136); *Glycine max GmCCD1b* (JF894311); *Medicago truncatula MtCCD1* (CAR57918); *Arabidopsis thaliana AtCCD1* (NP\_191911); *Coffea arabica CaCCD1* (ABA43904).

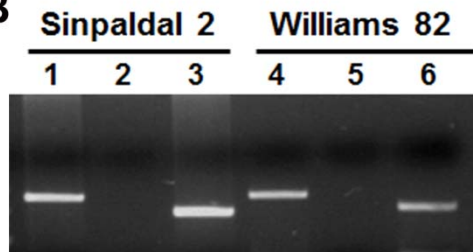
## Supplemental Figure S2

**A**

```

GAGTTTAGGT CCAGCGTCCG TGGGGGGGGT ACGACGAGAA ATTCTCTCGA GGTCAACAAGT -373
GTTTCATCTTC GTTGTTTCTC TTCGTAACAA AAGATGAAGA GCTTACACCG CAATTAGCCC -313
CACCATGTTT ACCGTCACCA ACACATGAT TACTTCCTGA ACCTCCTCCG GCACAGCTCA -253
GTTTACACGC CCTGTCTGGC CATTGAGCAC CTGAAACTTT ACCCGTGCAA GGGTTCATTG -193
GTGATTTTTTC GGTAAGCATT TTAGTAGATG GAGGCAGCAC ATATAATTTT CTGCACCACC -133
GTGTTGTGAT GGCCTTAGGT ATTTCAACAA CGGAGACTGC ACCACTTCGA GTAACGGTAG -73
GCAACAGGGA CGAGATTCGT TGCCATCAGC TGTGTAACGC TGTACACGTC CAAATACAAG -13
GGCATTTCGCT TCATGGGGGA TAGAATCGAA TTCCCGCGGC CGCCATGGCG GCCGGGAGCA +48
    
```

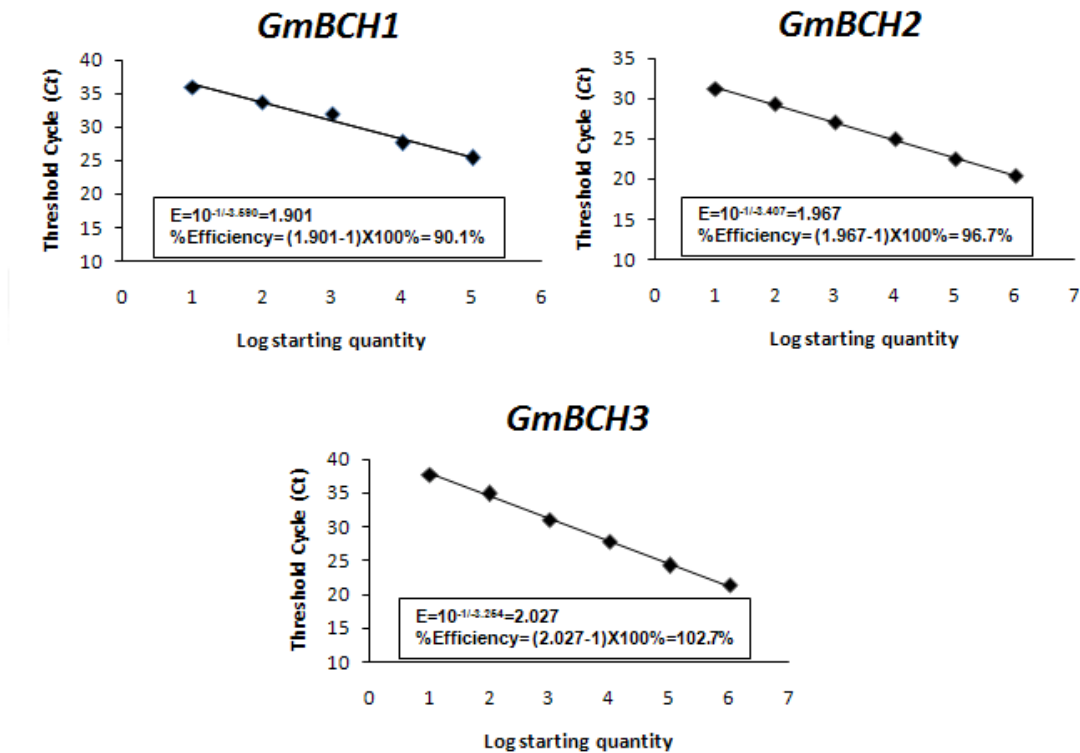
**B**



**Supplemental Figure S2. The upstream region of *GmBCH1* isolated by DNA walking.**

The 5' region of *GmBCH1* was isolated by using a DNA Walking SpeedUP<sup>TM</sup> Premix Kit (SEEGENE) according to manufacturer's recommendation. The underlined is the position of primer used in the DNA Walking and the start codon is indicated by shaded gray. (B) Genomic PCR analysis of *GmBCH1*. Genomic DNAs of the cultivar Sinpaldal 2 (lanes 1 to 3) and the cultivar Williams 82 (lanes 4 to 6) were used, respectively. Genomic PCR of *GmBCH3* was loaded in lanes 1 and 4 as control and primers used are the same indicated in Supplemental Fig. S4. Lanes 2 and 5 show genomic PCR product with *GmBCH1* forward primer (shown in red color) and *GmBCH3*-specific reverse primer, which serve as negative control. Lanes 3 and 6 indicate the existence of upstream region of *GmBCH1*. Primers used in genomic PCR of *GmBCH1* are indicated in red (forward primer) and orange (reverse primer) color in panel A.

## Supplemental Figure S3



**Supplemental Figure S3. Assessment of PCR efficiency of primer sets for *GmBCHs*.**

The efficiency of the primer sets used for *GmBCHs* was tested by plotting the threshold cycle (*Ct*) at each concentration of the samples. The slope of a linear-regression trendline indicates the primer efficiency. PCR efficiency of primers was calculated as  $[10(-1/S)-1] \times 100\%$ , where *S* is the slope of the regression line [Hu R, Fan C, Li H, Zhang Q, Fu Y-F (2009) Evaluation of putative reference genes for gene expression normalization in soybean by quantitative realtime RT-PCR. *BMC Mol Biol* 10: 93].



**Suppelemental Figure S4. Assessment of probes for *in situ* hybridization.**

(A and B) Sequence alignment of *GmBCH(1+3)* probe with *GmBCH2* genomic DNA (A) and *GmBCH2* probe with *GmBCH(1+3)* genomic DNA (B). (C) Whole-mount *in situ* hybridization of soybean leaves with *GmBCH(1+3)* sense probe, *GmBCH2* sense probe, *GmBCH(1+3)* antisense probe, and *GmBCH2* antisense probe. (D) The intensity of the images in panel C for *GmBCH(1+3)* and *GmBCH2* expression was measured using Image J (National Institutes of Health, Bethesda, MD, USA) as a gray scale value and shown as the numbers obtained by subtracting background gray values (Ubuka and Bentley, 2009).

## Supplemental Figure S5

### A. 5' upstream region of *GmBCH2*

```

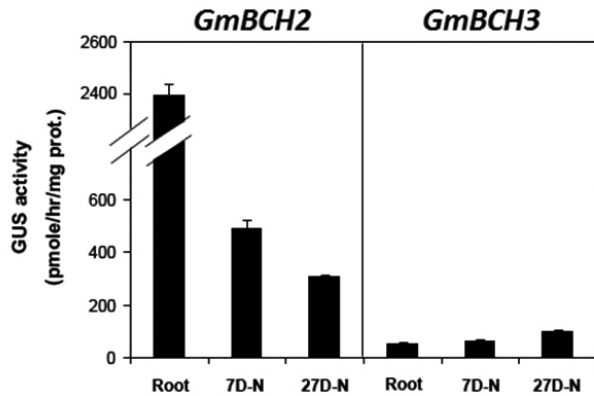
TCTTGCATTG CTATACTCTC GTTGATTAAAT TTAATTTATT TCGCTTACGA AAGCGAGTCA TGAAGTTATA -1490
AAATAC TCAA CT TACAAATA TGATT ACATT ACAAA TTTGA CTACATCCAC TGAATAAGCAT GCACACCGGA -1420
TCAT TT TAAT AT AT TATCGT TAACT GAAAT TTA TTAATAT A TAAATGAA ATATCA TC TC TATATAATA -1350
TCAAAT TTAG TAAA TAAAAA TTCAC AC TCG TTT TATTTAG TAAAATTTGT ATAATT TTTT TTATCAATTA -1280
TATA TA TATA TATA TATAT ATATA TTAAT AACAT TCTTCTTAAA TTAAT GAAC AAAATAACTC -1210
GACAGAAAT TT TAATGTCA CTCAT TAATG TAT GACCTAC AATAGATGA TTAATG TAAA TTAAGTTTGA -1140
GTAT TG TAAA TAAA AAAATGA TTAAG GGAGA ATA TT TACT AATAATAGTT ACTCAGATT TT TTAATAAA -1070
GATT AA TTAC CAAT AAAAAT TATAAAT ATT TTAACA TCAAT GTTATAGTGA AAAAAA TATT AT TAAATATT -1000
TTTT AA TATA TATA TATATA TATAT AT TGA TAT TTATATA T TATTAAT AA AATAAATAT AT ATTTGTACA -930
ATTT CTCTTC AT AT ATCAA TGAAC AC TTT GAC TT GGTGA TATGCTAGTG TACTTTGGGT GAAC TGGTGG -860
GGTGT TATC CC TC ATCAT C AATCC CAACA TCA TT TGTGA GCGTTACATT GGTAGAATT GT TAAATATA -790
CAAT GC TACT TAGA AAAAT C GAAC AAAAT TAA CCATACA TTTTACAT AT TTTAAA TTAG TCACACCAA -720
GTAT AT TATT TG TT TTCAT T TTTTA TAACT TAAACA ACTA TACTCTCTAA TATAGTATAT CATTTCTTAC -650
TTTT AACATA AA TG GAAGA TATTGAGCCA TAT TT AGATT TCTAAAAT TT TGA AAAAGTA AACACATAAA -580
TTCT CT TCTA GAAA AAAAAA CAATT TT TTC ACA GAAATGA AAGAATATCA AACCATATTA ATATGTGTGT -510
AATT TT TAAA AGAG AACGTC TAGAC TC TTC ACT TAACACA CCTTATGCA TTA AAA TTTA TAGAAAATTT -440
AACT AA TTTT AT TA TTTAT TTTAA TGATA TTG TT ATAAA ATATCATT TA TTTATT TTTA ATAACTCTAT -370
TTTT AT TATG AT AT TAAAT T TAAAT AAAAA CCA TT TAAA AATAACTT TA TTTTAT TAAA AATTAATAAA -300
ATTT AA TAA T TCAACTATA TTTT TAATA AGT CT TAAAG T GATTATT TT TGCATT TTTT AT TTCTCTGT -230
TATA GC TAAA AATGACTCCT ATACT GT GGT AAT CT TCCAC GAGCTTTT TC ATTTTG TGTT CACACTGGCC -160
CCAC AT AAAA AAAC CACCGA AGAAA CA AAT TCC ACACATA ACAA AAAAGT TGTGGTCTCT GT TCATCTCC -90
CTTT TT AAA C CT CC ACCT T CAAA C T T T CACAC AAACATCTCT CTCTCTCTCA CTCTCTCTCA -20
TCTC TC TCTC TC TC TCTCTC TAAACA CAC AAA CA CTACA C ATCACAATC TCACAT TGTT CTCAAGTTTC -1
AACACT TTCT AG TT CTACC
    
```

### B. 5' upstream region of *GmBCH3*

```

GGTTGGACAT CTGCAGGCCA AGACTTTACT TAAAT AATGT G TTTTTCATC AAAATAAGAC TGGAACCTGT -1527
CAAAAC TTGA CACT ACCCAA CCTAT TT CTA TTTCTAGATT AACAGTTAGG TTCTTTAGTG CTAGTTGTTA -1457
GGTA TTAGAA TT GC GCTGG A TGAAA GCGTT CAC CGAGCAT CATTTATT TT TTAATT AAGC CAAAAGCTAT -1387
TTGATGAAAT GGTT TTAATG TAAAT TT CTC AAA CAATAAA AAAAAAAT TA GTGTAGGAAAT AAAATTACAG -1317
TAAG TAAAAG CTGT AAGAACT CGGAAA AAAA TGA TG GCGCC AAAAGTTT TT TTCTTAGGAT AATGACCCA -1247
AAGCTT AGAG AC CT ATCTAT TATGAAAAGT CCT CAATCGA ATAGTCCAAT TTTTAAA AACT ACATTCTATC -1177
ACTT CA CACT AAC AAAAAA TTTTAAA CAA CTT TT TTTTG GGAAGAGAAT TTTCTTACTT TGCATGATTA -1107
AATCAA TAA ACAG TACTAT CACTGCAAAA TAAAT TTAGA ATCCTTCT TT TTAATT CAAT AAAAATAGTC -1037
TATC AC TGT TT TT TTAT A TTATA AAA GTA TTC AT GTCAA GAGTAGCAAG AAATGC TTA A GAAATCCTAA -967
CAAGTC CTT TT GT ATTAC TAGGG TGACT CAAGA GACTC ACTCTTAC TG ATTGAA TATA AT TTAGTCTA -897
ATAT CAAA GA TT CAGTTGT TAATA TCAAA GAC CT AATTA TCGTACTT TG GACAA TCCAC TCATTATAGT -827
TGGT TG GCGT AC TT GGA TGC ATAGT CCAACA ATC TC TCAAG TATTGAACAT GAAA GTCGTG GGCAC TTTAT -757
GGTT AAAAAA AT CC TCAAT C GATAT TT AAG TTA TT GTTAT ACTTGACCAA CCAAGT TAAC TAAACCATCC -687
AGAA TC TCA TGT GATGAC TTTAAA GAC AAGAA CATTG TGTATTTAG TGGCAT AATC TTC AATTAAT -617
TTTTAT AAA CT TT TT CAC A TGACC CGTGT AGA TC AAGGC AGATTAAC TT TATCTA TAAA AT TTGATTGA -547
GTTCAA GATC AAAGGCCCAA TTATC TTACT TTGCTAACTT ATTTATTA TG ATCTAT CAAC ATACTTAGAT -477
AAAT GA TCCA TAGT CCGCTCA CACAT TA GAC CTA AAA GACA TAAA GTCGTT GGTATGATTA GTCGATTTT -407
TCTGAT GAGT TTAT GCTTAA TTAGT TG TCG TAT CT GAATG TATTTTGA TT TTTGTAGTTA AAATATTGCT -337
AAAA GA TTT CT AC ATTA A TTTTACC TTT TA TTTCT G TTGAGTCAT AATGTT TTGC TAAATAAAAT -267
GTTA CAATGG GT TT AATGT A ATTAAGA CAA GTG TT GGGTG T GAGATGTA GAGGTT CCA CAGCTATAGT -197
TGTT CCACGA GGTT TAAAT T ATTTT TA TTG AAG TGAACCG ACCAC CAC CTATTGTGCA CAGTACAAA -127
AATT TT GGTG TATAAATATA TGCAACC TCC CTC TC TTCAC AATCATCA CA TCCCTTACTC GCAGGCCAAG -57
ACAAACACTC TC TC TTTGCC TCTTGCCCA ATT CGATCGA AGCTATCTAC CCTTCC -1
    
```

### C

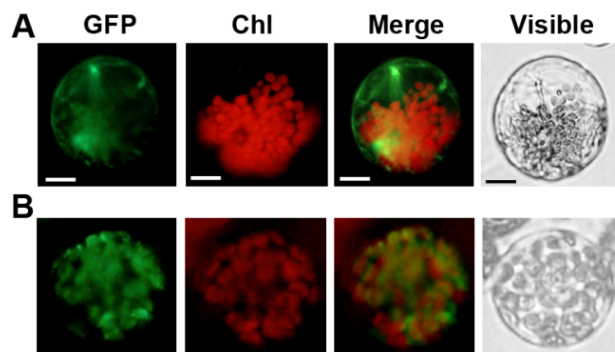




**Suppelemental Figure S5. Expression of p*GmBCH2-GUS* and p*GmBCH3-GUS* in transgenic soybean nodules measured by fluorometric assay of GUS.**

For *GmBCH2-GUS*, a 1,629-bp of 5' upstream region of *GmBCH2* was used (A). For *GmBCH3-GUS*, a 1,596-bp of 5' upstream region of *GmBCH3* was used (B). Primers used in cloning upstream regions are indicated by arrows. (C) Fluorometric GUS activity in transgenic nodules was measured according to Oh et al., (2001). GUS activities in transgenic nodules expressing *GmBCH2-GUS* and *GmBCH3-GUS* are shown as the means and standard deviations of the means of three independent experiments. *GmBCH1* promoter-*GUS* was not examined as the sequence of the *GmBCH1* promoter was not available.

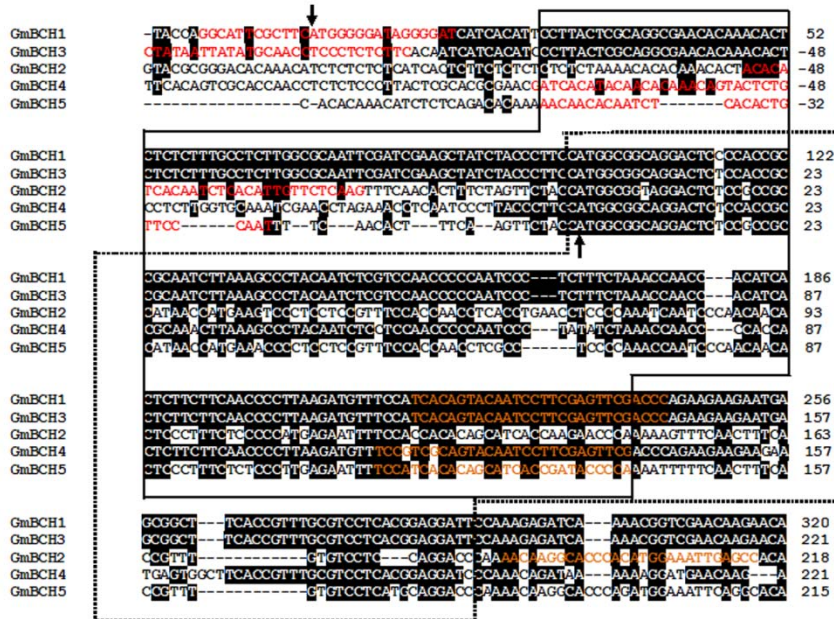
## Supplemental Figure S6



### Supplemental Figure S6. Subcellular localization of N-terminally deleted GmBCH2 and a fusion of GmBCH1 with the transit peptide of GmBCH2.

(A) The 5'-region corresponding to a putative transit peptide was removed from GmBCH2, and the modified GmBCH2 was fused with GFP. (B) The N-terminal region of GmBCH2 containing the putative transit peptide was fused with GmBCH1-GFP fusion construct. The resulting constructs were introduced into *Arabidopsis* protoplasts. After 16 h, transfected protoplasts were examined by fluorescence microscopy. The red signal shows the autofluorescence of chloroplasts (Chl). Scale bar = 10  $\mu$ m.

Supplemental Figure S7

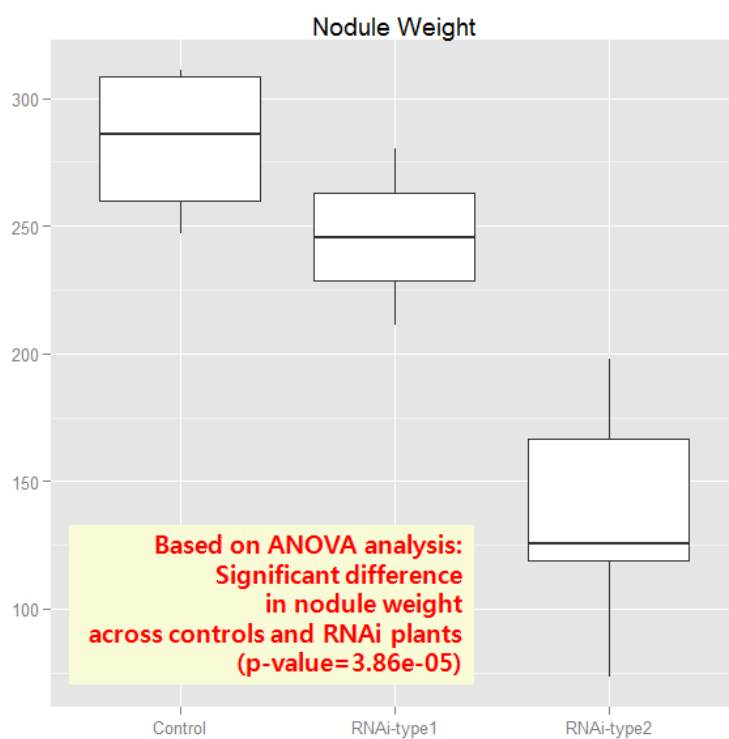


Supplemental Figure S7. DNA sequences used in preparing for *GmBCH(1+3)*-RNAi and *GmBCH2*-RNAi constructs.

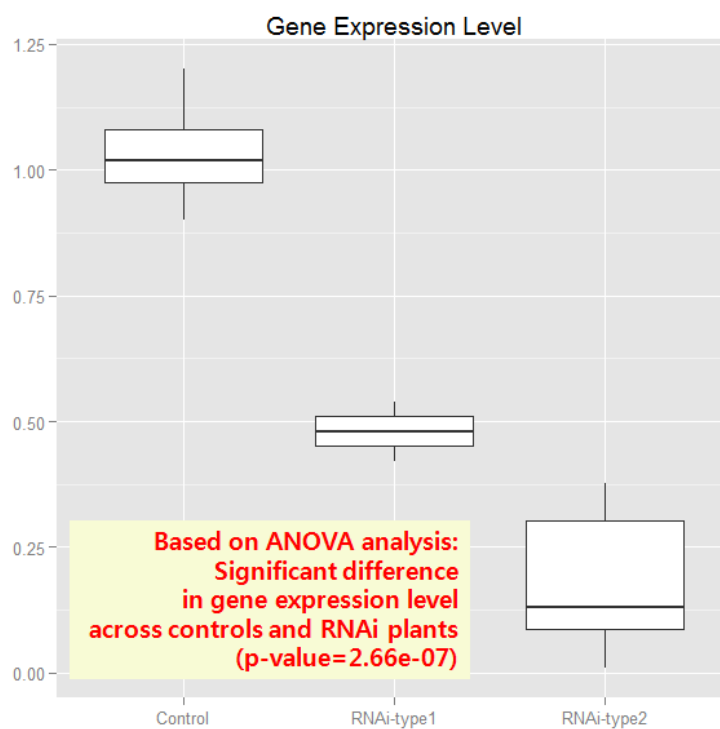
DNA sequence alignment of GmBCHs. Boxed areas with a solid line or a dashed line indicate DNA regions used in preparing for *GmBCH(1+3)*-RNAi or *GmBCH2*-RNAi constructs, respectively. Identical DNA sequences in both *GmBCH1* and *GmBCH3* are shaded black. Primers used in real-time PCR are indicated; forward primers for each GmBCHs in red color and reverse in orange color. The exact positions of forward and reverse primers are as follows; *GmBCH1*, -12 to 16 and 216 to 243; *GmBCH3*, -117 to -89 and 117 to 144; *GmBCH2*, -52 to -24 and 188 to 215; *GmBCH4*, -75 to -48 and 113 to 140; *GmBCH5*, -51 to -24 and 113 to 140. Initiation codons are indicated by either downwards arrow (for *GmBCH1*) or upwards arrow (for *GmBCH2*, *GmBCH3*, *GmBCH4* and *GmBCH5*).

## Supplemental Figure S8

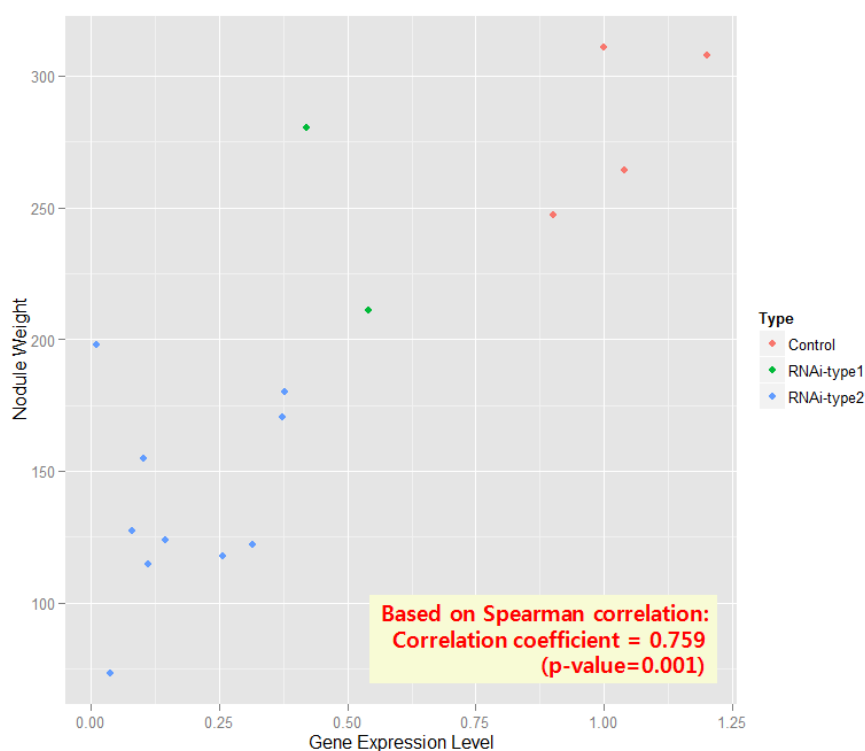
**A**



**B**



C

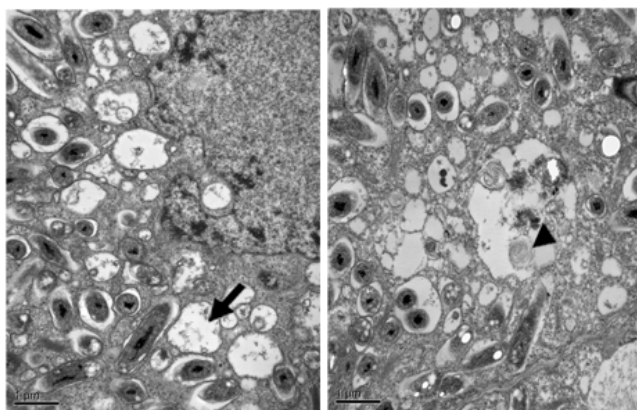


**Supplemental Figure S8. Statistical analyses of nodule weight and gene expression in *GmBCH(1+3)*-RNAi nodules.**

(A) Distribution of nodule weight between three plant groups. We compared nodule weights of those plants, and broken transgenic plants into two groups (say RNAi-type1 = transgenic plants with nodule weight similar to controls, including #2; RNAi-type2 = transgenic plants with nodule weight significantly smaller than controls, including #20, #21). In order to verify the statistical significance of this grouping (i.e., control vs RNAi-type1 vs RNAi-type2), we conducted analysis of variance (ANOVA) with Tukey's multiple comparison procedure, and found significant difference in nodule weight between control and RNAi-type2 (p-value = 0.00005) as well as between RNAi-type1 and RNAi-type2 (p-value = 0.00636), but not between control and RNAi-type1 (p-value = 0.5). (B) Distribution of gene expression level between three plant groups. We conducted ANOVA with Tukey's multiple comparison procedure for difference in gene

expression levels between three groups, and found that significant difference between each pair of groups: control and RNAi-type1 (p-value = 0.00083), control and RNAi-type2 (p-value = 0.0000002), RNAi-type1 and RNAi-type2 (p-value = 0.0294). Based on the results of two ANOVAs, we verified the significance of our grouping and characterized two groups of transgenic plants as follows: RNAi-type1 = no significant decrease in nodule weight with considerably reduced gene expression; RNAi-type2 = significant decrease in nodule weight with completely reduced gene expression. (C) Correlation between nodule weight and gene expression level. Using all of the 16 plants in three groups, we examined the correlation between gene expression levels and nodule weights based on Spearman rank correlation coefficient. Strong positive correlation of 0.76 (p-value = 0.001) was observed. Note that we used Spearman rank correlation, a popular correlation measure that can deal with small sample size and requires no normality assumption.

## Supplemental Figure S9



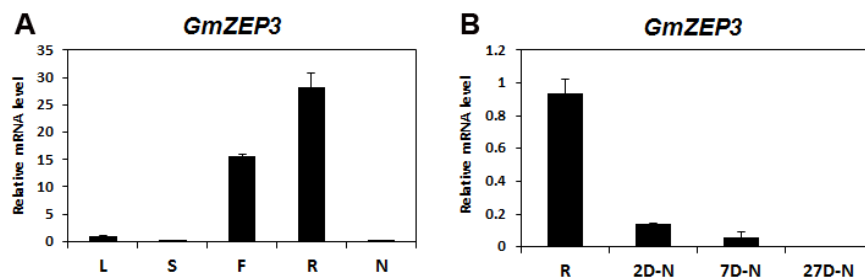
*GmBCH(1+3)*-RNAi

### Supplemental Figure S9. Electron microscopic analysis of *GmBCH(1+3)*-RNAi nodules.

*GmBCH(1+3)*-RNAi nodules often contained empty vesicles without bacteroids (arrow) and in severe case, bacteroids without being enclosed by symbiosome membranes (arrowhead).



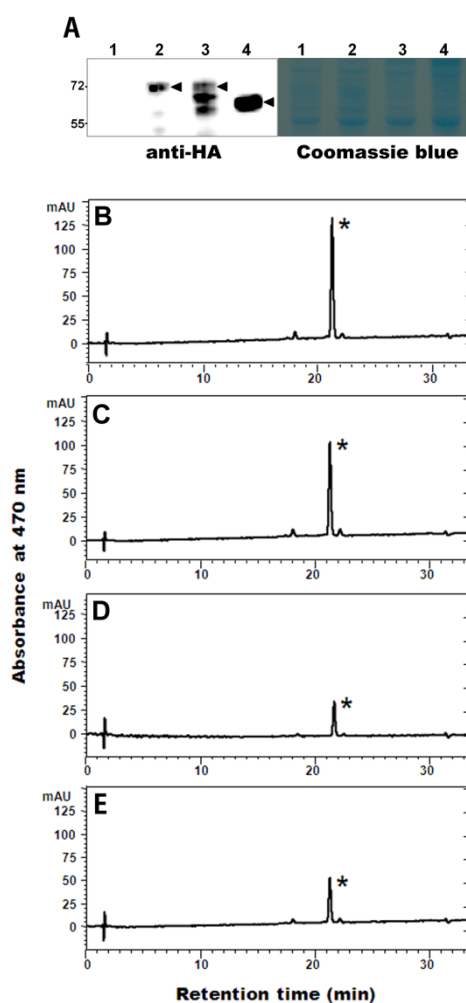
## Supplemental Figure S10



### Supplemental Figure S10. Expression of *GmZEP3* in *Glycine max*.

(A) Expression of *GmZEP3* in different tissues. RNA was extracted from different tissues including 27-day-old nodules of *Glycine max*. (B) Expression of *GmZEP3* during nodule development. RNAs were extracted from roots, 2-, 7-, and 27-day-old nodules, and transcript levels were determined by real time RT-PCR and normalized. Data are from three independent experiments. L, leaf; S, stem; F, flower; R, root, N, Nodule.

## Supplemental Figure S11



**Supplemental Figure S11. Functional assay of GmCCD7 and GmCCD8 in  $\beta$ -carotene-accumulating *E. coli*.**

Expression of *GmCCD7*, *GmCCD8*, and *AtCCD7* in *E. coli* strains that carry pACCAR16 $\Delta$ crtX and accumulate  $\beta$ -carotene. The  $\beta$ -carotene-accumulating *E. coli* strain with empty vector alone served as negative control. (A) Proteins from  $\beta$ -carotene-accumulating *E. coli* cells expressing *HA-GmCCD7* (lane 2), *HA-AtCCD7* (lane 3) and *HA-GmCCD8* (lane 4) or with empty vector alone (lane 1) were isolated (right panel) and immunoblotted with HA antibody (left panel). Proteins of the expected sizes corresponding to *GmCCD7* (70 kDa), *AtCCD7* (72 kDa) and *GmCCD8* (60 kDa) were

detected. Immunodetected bands were indicated by arrowheads. HPLC analysis of carotenoids extracted from  $\beta$ -carotene-accumulating *E. coli* cells expressing *HA-GmCCD7* (C), *HA-AtCCD7* (D) and *HA-GmCCD8* (E) or with empty vector alone (B). The  $\beta$ -carotene peak is indicated by an asterisk (\*). mAU, milli-absorbance units.

**Supplemental Table S1.** RNA-seq expression data for soybean genes in various tissues. The expression levels were shown according to Severin et al. (2010) (A) and Libault et al. (2010) (B)

(A) RNA-seq expression data for soybean genes from Severin et al. (2010)

Gene name in the paper	Gene model #	young_leaf	flower	root	nodule
GmBCH2	Glyma16g29790	45	12	8	7
GmBCH3	Glyma10g37560	2	32	7	4
GmBCH4	Glyma20g30230	2	2	2	1
GmBCH5	Glyma09g24530	15	5	0	0
GmZEP1	Glyma17g20020	46	18	0	0
GmZEP2	Glyma11g05960	1	6	3	1
GmZEP3	Glyma01g39310	0	0	0	0
GmVDE1	Glyma03g41420	9	7	0	0
GmVDE2	Glyma19g44010	14	4	1	0
GmNCED1a	Glyma05g27250	0	1	0	0
GmNCED1b	Glyma08g10190	0	129	0	3
GmCCD1a	Glyma12g36530	2	2	1	0
GmCCD1b	Glyma13g27220	15	19	12	11
GmCCD7a	Glyma11g16370	0	0	4	0
GmCCD7b	Glyma01g14170	0	0	13	0
GmCCD8a	Glyma06g09000	0	0	33	1
GmCCD8b	Glyma04g08910	0	0	19	3

(B) RNA-seq expression data for soybean genes from Libault et al. (2010)

Gene name in the paper	Gene model #	leaf	flower	root	nodule
GmBCH2	Glyma16g29790	33	14	17	9
GmBCH3	Glyma10g37560	14	40	8	2
GmBCH4	Glyma20g30230	15	0	3	1
GmBCH5	Glyma09g24530	2	3	1	0
GmZEP1	Glyma17g20020	886	273	0	0
GmZEP2	Glyma11g05960	11	6	5	2
GmZEP3	Glyma01g39310	0	1	0	0
GmVDE1	Glyma03g41420	77	47	1	1
GmVDE2	Glyma19g44010	58	36	3	0
GmNCED1a	Glyma05g27250	0	1	0	0
GmNCED1b	Glyma08g10190	1	469	0	1

GmCCD1a	Glyma12g36530	16	8	0	1
GmCCD1b	Glyma13g27220	212	79	35	34
GmCCD7a	Glyma11g16370	0	0	7	0
GmCCD7b	Glyma01g14170	0	0	3	0
GmCCD8a	Glyma06g09000	0	0	17	0
GmCCD8b	Glyma04g08910	0	0	9	9

**Supplemental Table S2.** Primers used for gene cloning (A) and gene expression by real time RT-PCR (B)

(A) Primers for gene cloning

Gene	Primer sequence (5'-3')	Purpose
<i>GmBCH2</i>	5'-RACE : GTCTACCAGGCATTCGCTTCAT Gene-specific R: GTTAAAGAAACCATAGGAAAGGAGAGCG	Isolation of the 5' region of <i>GmBCH2</i>
	F: CTAGTCTAGACATGGCGGTAGGACTCTC R: CATTCTGGATCCGTATTAGTTGTTCTCA	Isolation of the full-length <i>GmBCH2</i> cDNA
	F: CTAGTCTAGACATGGCGGTAGGACTCTC R: TGATTGGATCCAGTTGTTCTCATTAGC	Construct for subcellular localization
	F: CTA CTCTAGAAATGGCGGCAGGACTCTCC R: TTGGCGGATCCA TGA ACCGGACCTGATGC	Cloning of N-terminal end- deleted <i>GmBCH2</i>
	F: CTATCTAG ACC ATGGCGGTAGGACTCTC R: AAGTTTCTAGATTTTGGGTTCTGGTGA	Cloning of N-terminal region of <i>GmBCH2</i> fused to <i>GmBCH1</i>
	Sense F: CTCTCTAGAACACACAAAACACTACACAT Sense R: GGTC AAGCTTGAGGACACAAAACGGTGAA Antisense F: TCTCTCGAGAACACACAAAACACTACACA Antisense R: TTTGGTACCTGCATGAGGACACAAAACGGT	Construct for <i>GmBCH2</i> -RNAi
	F: CTCACAAGCTTTCTTGGATTGCTATACTCTCGTTG R: ACCGCCATGGTAGA ACTAGAAAAGTGTGAAACTTG	Isolation of <i>GmBCH2</i> promoter
	F1: GTAAAGTTTGATACTTTCACTCCTGCG R1: GAAACAAAAGTACTGTTTGTGTCCCAAG	Isolation of a part of <i>GmZEP</i>
	5'-RACE: GTCTACCAGGCATTCGCTTCAT R2: ACTGCAGG AGTGAATGTATCAAAAATTGA	Isolation of the 5' region of <i>GmZEP</i>
	F2: TGCCTGCTGACATTGAAACTGTTGGATA 3'-RACE: TCGTAGTCGCAGCATTACACAG	Isolation of the 3' region of <i>GmZEP</i>
<i>GmBCH1</i>	F: GGCATTCTAGACATGGGGGATAGGGGAT R: GAACCCTAGGAGTACTTGGCCTGG	Isolation of the full-length <i>GmBCH1</i> cDNA
	F: GGCATTCTAGACATGGGGGATAGGGGAT R: TTGGCGGATCCATGAACCGACCTGATG	Construct for subcellular localization
	Sense F: GATCTCTAGATTCCCTTACTCGCAGGCCG Sense R: ACTCGAAGCTTTGTACTGTGATGGAAACA Antisense F: GATCATCTCGAGCCTTACTCGCAGGCCGAA Antisense R: ACT CGAAGGTACCTACTGTGATGGAAACA	Construct for <i>GmBCH(1+3)</i> -RNAi
	F: CTATCTAGACATGGCGGCAGGACTCTCC R: GAACCCTAGGAGTACTTGGCCTGG	Isolation of the full-length <i>GmBCH3</i> cDNA
<i>GmBCH3</i>	F: CTACTTAGAATGGCGGCAGGACTCTCC R: TTGGCGGATCCATGAACCGACCTGATGC-3	Construct for subcellular localization
	F: TAACAAGCTTGGTTGGACATGTGCAGGCCAAGA R: CCGCCATGGAAGGGTAGATGCTTCGATCGAAT	Isolation of <i>GmBCH3</i> promoter
	F: ACTCAAGCTTAATGCAAGCCAAACCCATT R: TTGGGATCCAGCTAATTAGCTGCCCG	Isolation of the full-length <i>GmCCD7a</i> cDNA
<i>GmCCD8a</i>	F: CATGTCTAGATATGCTTGTTCCTGACATGCCTCATAACA R: TAGGGATCCCTACTCTTTTGGAAACCAGCATCCATGC	Isolation of the full-length <i>GmCCD8a</i> cDNA
<i>AtCCD7</i>	F: GTTTC TAGAGATGTCTCTCCCTATCCCGCCGAAATTC R: AGAAGGATCCAGCTTGTGCTTTGCCAATCAT	Isolation of the full-length <i>AtCCD7</i> cDNA

(B) Primers for gene expression by real time RT-PCR

Gene	Primer sequence (5'-3')	Efficiency*
<i>GmBCH1</i>	F: GGCATTCTAGACATGGGGGATAGGGGAT R: GGGTCGAACTCGAAGGATTGTACTGTGA	90.1%
<i>GmBCH2</i>	F: ACACATCACAATCTCACATTGTTCTCAAG R: GGCTCAATTTCCATGTGGGTGCCTTGTT	96.7%
<i>GmBCH3</i>	F: CTATAATTATATGCAACCTCCCTCTCTTC R: GGGTCGAACTCGAAGGATTGTACTGTGA	102.7%

<i>GmBCH4</i>	F: GATCACATACAACACAAACAGTACTCTG R: CGAACTCGAAGGATTGTA CTGCGACGGA	89.5%
<i>GmBCH5</i>	F: AACAAACACAATCTCACACTGTTCCCAAT R: TGGGGTATCGGTGATGCTGTGTGATGGA	90.1%
<i>GmZEP(1+2)</i>	F: GTAAAGTTTGATACTTTCACCTGCG R: GAAACAAAGTACTGTTTGITGCCAAG	100.4%
<i>GmZEP3</i>	F: GCCGTTGATATTACTTATCTTGTAAGT R: ATCACGCTAGGGCGCTTATGAGTACATA	98.1%
<i>GmVDE(1+2)</i>	F: GCTATCAAGATGACTGGTATATTTTGTC R: CACCTCTTCTTCTAACTGTTCAACTTCC	91.0%
<i>GmNCED(1a+1b)</i>	F: ACCACCTCTTCGACGGCGACGGAATGGT R: ATGGCGAGGAGTTTTCCGTTGAAGAAGA	89.2%
<i>GmCCD(1a+1b)</i>	F: GGAACCACATTAGACAGCATTGCAAAAAGT R: TCAAGTATCCATCATCTTCTCAGAATC	92.3%
<i>GmCCD(7a+7b)</i>	F: ATCACGGGTCCACGGTGCACCCTCTA R: GTGTTAGCCCAATTTTCATGACCTTAGT	96.0%
<i>GmCCD(8a+8b)</i>	F: AGGAATGGTCCGGGCATGTGGCACATA R: TTCGCCGCCTTAGGCACCTCGGAGAACT	92.7%
<i>GmLbc</i>	F: ATGGGTGCTTTCACTGAGAAGCAAGAGG R: TGCTGCTGCCAATTCATCATAGGCTACT	102.7%
<i>Ubiquitin</i>	F: GGGTTTTAAGCTCGTTGT R: GGACACATTGAGTTCAAC	104.2%
<i>GmELF1b</i>	F: GTTGAAAAGCCAGGGGACA R: TCTTACCCCTTGAGCGTGG	89.7%
<i>GmActin2/7</i>	F: CTTCCCTCAGCACCTTCCAA R: GGTCCAGCTTTCACACTCCAT	90.5%

\*The efficiency of the primer sets used for *GmBCHs* was tested by plotting the threshold cycle (*C<sub>t</sub>*) at each concentration of the samples. The slope of the linear-regression trendline indicates the primer efficiency. PCR efficiency of primer sets was calculated as  $[10^{(-1/S)} - 1] \times 100\%$ , where S is the slope of the regression line (Hu et al., 2009).