Supplemental Data for:

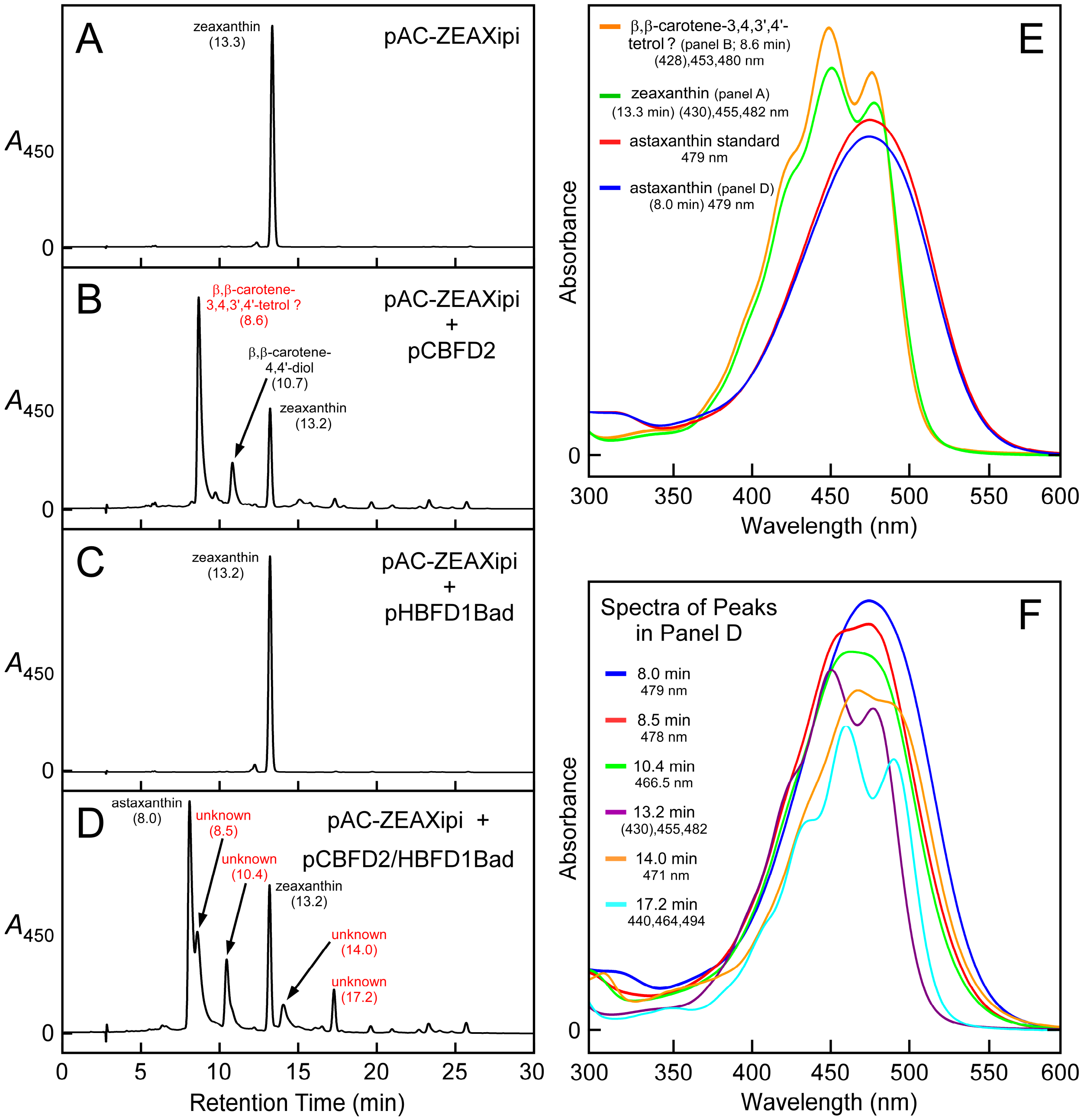
Elucidation of the Pathway to Astaxanthin in the Flowers of *Adonis aestivalis*

Francis X. Cunningham, Jr. and Elisabeth Gantt
Department of Cell Biology and Molecular Genetics
University of Maryland, College Park, MD 20742 USA

Supplemental Figure 1 (below). The Product of the *hbfd1* cDNA Does Not Modify Carotenoid 3-Hydroxy-β-Rings, the Activity of an *Erwinia herbicola* CrtZ-type β-Ring 3-Hydroxylase Enzyme (encoded by the *crtZ* gene in plasmid pAC-ZEAXipi) Impairs the Conversion of β-Carotene into Astaxanthin, and the Products of the *A. aestivalis cbfd2* cDNA and the *E. herbicola crtZ* Gene Together Convert β-Carotene into a Compound with a Retention Time, Absorption Spectrum, and Molecular Mass Indicative of a Tetrahydroxy-Carotenoid, Probably β,β-Carotene-3,4,3',4'-Tetrol. Shown are HPLC elution profiles for extracts of *E. coli* wherein cells contained the following plasmids: **(A),** pAC-ZEAXipi; **(B),** pAC-ZEAXipi and pCBFD2; **(C)**, pAC-ZEAXipi and pHBFD1Bad; **(D)**, pAC-ZEAXipi and pCBFD1/HBFD1Bad. Panel **(E)** displays absorption spectra (in the HPLC mobile phase) for the major peaks in (B) and (D) and for a synthetic astaxanthin standard [peak in panel (E) of Figure 3 of the manuscript). Panel **(F)** shows absorption spectra for selected peaks of panel (D). Numbers in parentheses below the carotenoid names in panels (A), (B), (C) and (D) are HPLC retention times in minutes. Labels for the absorption spectra in (E) and (F) refer to the HPLC retention times for the corresponding peaks in (B) and (D), and, in the case of the synthetic astaxanthin standard, in panel (E) of Figure 3 of the manuscript. Absorption maxima are listed below these retention times, with peak "shoulders" enclosed in parentheses. Spectra were recorded in the HPLC mobile phase just as the individual pigments eluted from the HPLC column. Those peak identifications indicated with red text and followed by question marks in panels (B) and (D) are speculative: they are consistent with the absorption spectra, the HPLC retention times, and the known catalytic capabilities of the carotenoid pathway enzymes that are present in the *E. coli* cells, but the requisite standards were not available for comparison. The plasmid pAC-ZEAXipi (Cunningham and Gantt, 2007) contains carotenoid pathway genes from *E. herbicola* that lead to the synthesis and accumulation of zeaxanthin (3,3'-dihydroxy-β,β-carotene) in *E. coli*. Extracts for which the elution profiles are shown in (C) and (D) were from cultures grown in liquid media containing arabinose (0.2%, w/v) in order to induce production of the HBFD1 polypeptide. **Note:** because considerable amounts of both zeaxanthin (β,β-carotene-3,3'-diol) and isozeaxanthin (β,β-carotene-4,4'-diol) were present in extracts of *E. coli* cultures containing pAC-ZEAXipi and pCBFD2 [panel (B)], it is not clear whether 3,4-dihydroxy-β-rings are made from 3-hydroxy-β-rings [with hydroxylation of the number 4 carbon then catalyzed by the product of the *cbfd2* (*AdKeto2*) cDNA] or from 4-hydroxy-β-rings (with hydroxylation of the number 3 carbon then catalyzed by the CrtZ enzyme) or *via* both of these routes.

Reference for Supplemental Figure 1:

Cunningham, F.X. Jr., and Gantt, E. (2007) A portfolio of plasmids for identification and analysis of carotenoid pathway enzymes: *Adonis aestivalis* as a case study. Photosynth. Res. **92:** 245-259.


A
zeaxanthin (13.3)
pAC-ZEAXipi
A_{450}
0
B
β,β-carotene-3,4,3',4'-tetrol ? (8.6)
β,β-carotene-4,4'-diol (10.7)
zeaxanthin (13.2)
pAC-ZEAXipi + pCBFD2
A_{450}
0
C
zeaxanthin (13.2)
pAC-ZEAXipi + pHBFD1Bad
A_{450}
0
D
astaxanthin (8.0)
unknown (8.5)
unknown (10.4)
zeaxanthin (13.2)
unknown (14.0)
unknown (17.2)
pAC-ZEAXipi + pCBFD2/HBFD1Bad
A_{450}
0
0
5
10
15
20
25
30
Retention Time (min)
E
β,β-carotene-3,4,3',4'-tetrol ? (panel B; 8.6 min) (428),453,480 nm
zeaxanthin (panel A) (13.3 min) (430),455,482 nm
astaxanthin standard 479 nm
astaxanthin (panel D) (8.0 min) 479 nm
Absorbance
0
300
350
400
450
500
550
600
Wavelength (nm)
F
Spectra of Peaks in Panel D
8.0 min 479 nm
8.5 min 478 nm
10.4 min 466.5 nm
13.2 min (430),455,482
14.0 min 471 nm
17.2 min 440,464,494
Absorbance
0
300
350
400
450
500
550
600
Wavelength (nm)

Supplemental Figure 2 (below). Three SALK T-DNA Homozygous Lines with Insertions Within or Immediately Upstream of the Coding Region of *A. thaliana* Gene *At1g50450*, Which Encodes a Polypeptide Similar in Sequence to the Product of the *A. aestivalis hbfd1* and *hbfd2* cDNAs, do not Differ from Wild-Type *A. thaliana* Plants in the Content and Composition of the Carotenoids Within Their Leaves. Shown are HPLC elution profiles for extracts of young leaves of: **(A)**, the wild-type parent strain (Col-0; ABRC Stock number CS60000). **(B)**, Salk_137455C mutant [insertion in 3rd of 6 exons of *At1g67080* (*aba4*)]. **(C)**, Salk_019364C mutant [insertion in the 3rd of 11 exons of *At5g57030* (*lut2*), which encodes a lycopene ε-ring cyclase enzyme (LCYe)] **(D)** Salk_060961C mutant (insertion in the 5’ UTR of *At1g50450*, near to the initiation codon). **(E)**, Salk_150621C mutant (insertion in the 1st exon of *At1g50450*). **(F)**, Salk_070593C mutant (insertion in the 12th of 13 exons of *At1g50450*). **(G)**, Absorption spectra, in HPLC mobile phase, of selected peaks in panel (A). Absorption maxima for each pigment are in parentheses. (**H)**, Absorption spectra, in HPLC mobile phase, of selected peaks in panel **(C).** Absorption maxima for each pigment are in parentheses. The absorption spectra displayed in panels (G) and (H) support the identifications of the corresponding peaks in panels (A) and (C). Elution profiles for the *aba4* and *lut2* mutants are shown here because they assist in the identification of several of the carotenoid peaks in the elution profiles of the wild-type and *At1g50450* insertion mutants. The *aba4* mutant [panel (B)] lacks neoxanthin, as reported earlier for other mutants with insertions in gene *At1g67080* (North et al., 2007). The *lut2* mutant [panel (C)], lacks carotenoids with ε-rings (lutein and α-carotene) and has increased amounts, relative to chlorophyll *a*, of most β-ring carotenoids including violaxanthin, antheraxanthin, zeaxanthin, and β-carotene, as reported earlier for another *lut2* mutant (Pogson et al., 1996). Extraction of carotenoid pigments and HPLC analysis were generally as described in Norris et al., (1995), but with some minor modifications. Young leaves from short day plants were collected near the end of the dark period, briefly ground with a Teflon pestle in a 1.5 mL microfuge tube, and then extracted first with absolute acetone and then with ethyl acetate. The extracts were transferred to and combined in a second microfuge tube, water was then added, the tube was centrifuged briefly, and the epiphase containing the chlorophyll and carotenoid pigments was collected and evaporated under a stream of nitrogen. The pigments were then resolubilized in a small volume of the initial HPLC mobile phase. The HPLC separation utilized a gradient of 4-50% mobile phase B in A over the course of 35 min. The *A. thaliana* wild-type, as well as each of the mutants, was of the Col-0 ecotype. Seeds for the wild-type and mutant lines were obtained from the ABRC at The Ohio State University. The SALK T-DNA Insertion Lines are described in Alonso et al., (2003). **Abbreviations:** ABRC, Arabidopsis Biological Resource Center; UTR, untranslated region.

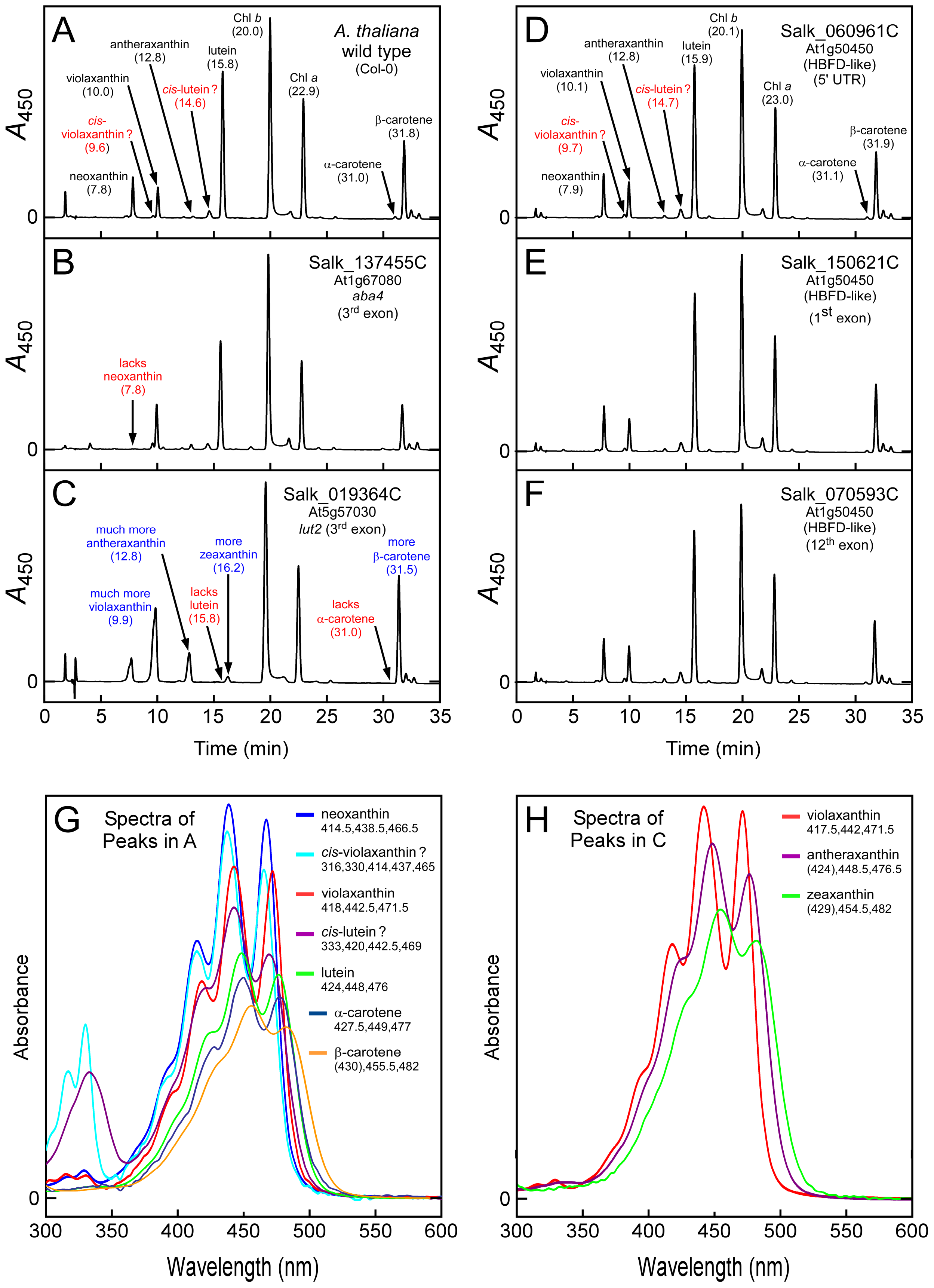
References for Supplemental Figure 2:

Alonso, J.M., et al., (2003) Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. Science **301:** 653-657.

Norris, S.R., Barrette, T.R., and DellaPenna, D. (1995) Genetic dissection of carotenoid synthesis in Arabidopsis defines plastoquinone as an essential component of phytoene desaturation. Plant Cell. **7:** 2139-2149.

North, H.M., De Almeida, A., Boutin, J.P., Frey, A., To, A., Botran, L., Sotta, B., and Marion-Poll, A. (2007) The Arabidopsis ABA-deficient mutant *aba4* demonstrates that the major route for stress-induced ABA accumulation is *via* neoxanthin isomers. Plant J. **50:** 810-824.

Pogson, B., McDonald, K.A., Truong, M., Britton, G., and DellaPenna, D. (1996) Arabidopsis carotenoid mutants demonstrate that lutein is not essential for photosynthesis in higher plants. Plant Cell **8:** 1627-1639.


A
A. thaliana
wild type
(Col-0)
neoxanthin (7.8)
cis-violaxanthin ? (9.6)
violaxanthin (10.0)
antheraxanthin (12.8)
cis-lutein ? (14.6)
lutein (15.8)
Chl b (20.0)
Chl a (22.9)
α-carotene (31.0)
β-carotene (31.8)
B
Salk_137455C
At1g67080
aba4
(3rd exon)
lacks neoxanthin (7.8)
C
Salk_019364C
At5g57030
lut2 (3rd exon)
much more antheraxanthin (12.8)
much more violaxanthin (9.9)
lacks lutein (15.8)
more zeaxanthin (16.2)
lacks α-carotene (31.0)
more β-carotene (31.5)
D
Salk_060961C
At1g50450
(HBFD-like)
(5' UTR)
neoxanthin (7.9)
cis-violaxanthin ? (9.7)
violaxanthin (10.1)
antheraxanthin (12.8)
cis-lutein ? (14.7)
lutein (15.9)
Chl b (20.1)
Chl a (23.0)
α-carotene (31.1)
β-carotene (31.9)
E
Salk_150621C
At1g50450
(HBFD-like)
(1st exon)
F
Salk_070593C
At1g50450
(HBFD-like)
(12th exon)
A450
0
0 5 10 15 20 25 30 35
Time (min)
G
Spectra of Peaks in A
neoxanthin 414.5,438.5,466.5
cis-violaxanthin ? 316,330,414,437,465
violaxanthin 418,442.5,471.5
cis-lutein ? 333,420,442.5,469
lutein 424,448,476
α-carotene 427.5,449,477
β-carotene (430),455.5,482
H
Spectra of Peaks in C
violaxanthin 417.5,442,471.5
antheraxanthin (424),448.5,476.5
zeaxanthin (429),454.5,482
Absorbance
0
300 350 400 450 500 550 600
Wavelength (nm)

Supplemental Figure 3 (below). Alignment of Plant and Green Algal CHYb-Type Carotenoid β-Ring 3-Hydroxylase Enzymes and Polypeptides Encoded by the *A. aestivalis cbfd1* and *cbfd2* cDNAs. The alignment was produced using MAFFT version 6 (Katoh and Toh, 2008; the server at http://mafft.cbrc.jp/alignment/server/ was used). The E-INS-i strategy was employed with the BLOSUM45 scoring matrix, a gap opening penalty of 1.0 and an offset value of 0.1. The output was manually adjusted to optimize the alignment. Residues are in white text on a black background where all sequences have the same amino acid, in white text on a dark grey background where 75% or more residues are identical, and in black text on a light grey background where 50% or more of the sequences have the same amino acid. Three residues of the *A. aestivalis* CBFD1 and CBFD2 polypeptides are highlighted in red where CHYb polypeptides have well-conserved proline residues adjacent to two of the conserved histidine motifs that are believed to comprise part of the active site in enzymes of this family (the membrane-associated fatty acid desaturase family; Shanklin et al., 1994). CBFD1 and CBFD2 do not have prolines in these positions, and are therefore likely to have a somewhat different geometry about the active site. The sequences used to construct this alignment were obtained from the GenBank protein database (all full length amino acid sequences identified using a blastp search on 2/18/2010), or they were generated by assembly and translation of overlapping expressed sequence tags (ESTs) obtained *via* a tblastn search of the GenBank nucleotide database and/or the database of ESTs. **GenBank Accession Numbers:** *A. aestivalis* CBFD2, AY644758; *A. aestivalis* CBFD1 (partial), AY644757; *B. rapa* 1, AAS55552; *B. rapa* 2, ACS45170; *A. thaliana* 1, NP_194300; *A. thaliana* 2, NP_200070; *G. max* 2, assembly of FK002367 and CX702394; *M. truncatula*, assembly of EV258794 and AW774742; *G. max* 1, ACU18541; *P. trichocarpa* 1, XP_002327604; *P. trichocarpa* 2, XP_002305933; *R. communis* 1, XP_002513654; *C. unshiu*, AAG10793; *D. kaki*, ACN86365; *V. vinifera* 2, assembly of CB342437 and EC952475; *I. nil*, BAI47580; *S. lycopersicum* 1, CAB55626; *C. annuum* 1, CAA70427; *C. arabica*, ABA43903; *S. lycopersicum* 2, CAB55625; *C. annuum* 2, assembly of GD111816 and GD133362; *V. vinifera* 1, AAM77007; *G. lutea*, BAE92729; *A. aestivalis* CHYb, ABI93208; *D. carota* 1, ABB52074; *Z. mays* 2, NP_001148085; *S. bicolor* 1, XP_002465958; *O. sativa japonica* 1, ABF93742; *O. sativa japonica* 3, NP_001065155; *E. oleifera*, ABS76147; *Z. mays* 3, NP_001105865; *S. bicolor* 2, XP_002446979; *O. sativa japonica* 2, NP_001053640; *Z. mays* 1, NP_001105907; *T. erecta*, AAG10430; *C. sativus* 1, CAC95130; *C. sativus* 2, AAT84408; *N. pseudonarcissus*, CAC06712; *D. carota* 2, ABB52075; *R. communis* 2, XP_002509848; *Oncidium* Gower Ramsay, ACP27626; *P. sitchensis* 1, ABK24778; *P. sitchensis* 2, ABK25382; *P. taeda* 1, assembly of DT638806 and DR021786; *P. taeda* 2, assembly of DR013495 and DR167448; *M. zofingiensis*, ABS50237; *H. pluvialis*, AAD54243; *C. reinhardtii*, XP_001698698.

Reference for Supplemental Figure 3:

Katoh, K., and Toh, H. (2008) Recent developments in the MAFFT multiple sequence alignment program. Brief. Bioinform. **9:** 286-298.

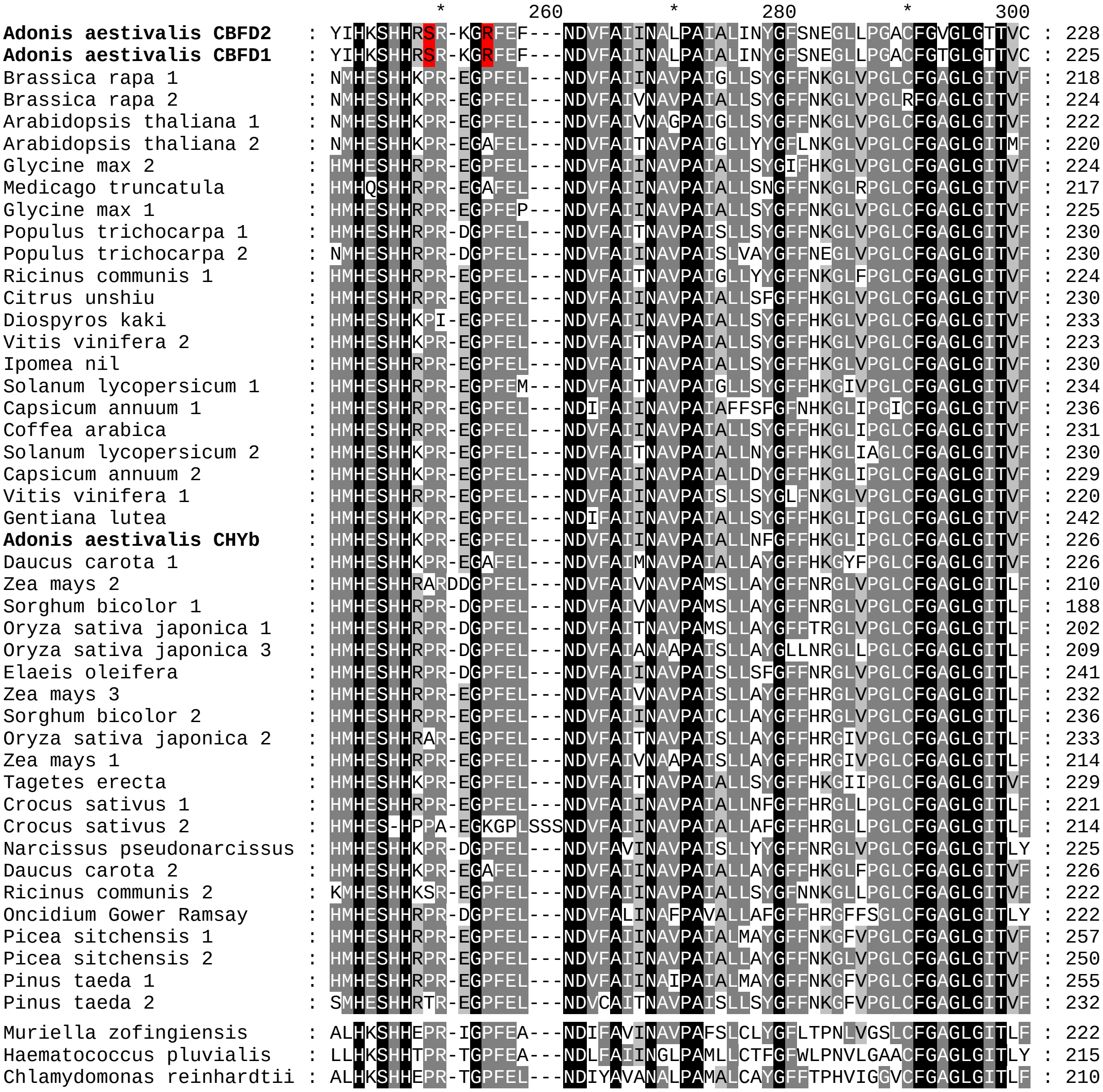
Shanklin, J., Whittle, E., and Fox, B.G. (1994) Eight histidine residues are catalytically essential in a membrane-associated iron enzyme, stearoyl-CoA desaturase, and are conserved in alkane hydroxylase and xylene monooxygenase. Biochemistry **33:** 12787-12794.

```
                                     *        20         *        40         *        60
Adonis aestivalis CBFD2   : M--AAAISV-FSSG----YSFY----KNL-LL--DSK--PNILKPP-----CLLFSPVVI :  39
Adonis aestivalis CBFD1   : -----AISV-FSTS----YSFH----KNL-LL--HSK--QDILNRP-----CLLFSPVVV :  36
Brassica rapa 1           : M--AAGLST-TVTF-NP-LHRSFSSSSS--VR--LHH--PRS---------LTGLPSS-- :  38
Brassica rapa 2           : M--AAGLST-ALTF-KP-LHRAFSSSSS--LR--LHH--PTS---------LTGLPSS-- :  38
Arabidopsis thaliana 1    : M--AAGLST-AVTF-KP-LHRSFSSSSTD-FR--LRL--PKS---------LSGFSPS-- :  39
Arabidopsis thaliana 2    : M--AAGLSTIAVTL-KP-LN------RSS-FS--ANH--PIS---------TAVFPPSLR :  36
Glycine max 2             : M--AAGLST-AAIL-KP-YNLV----QPP-IP--LSK--PTT---------SLFFNPLRC :  37
Medicago truncatula       : M--AAELYT-ATTL-KP-YNLL----QPS-TS--SPS--PSPK--------TLFFTPLRS :  38
Glycine max 1             : M--AAGLSA-AITM-KPLLRFH----QPR-----LPK--PIPT--------TLPFSPLRI :  37
Populus trichocarpa 1     : M--AAGLTA-ATVP-KP-FRYN----SVS-HL--LPK--PVTA--------ASLFFPPIR :  38
Populus trichocarpa 2     : M--ASGITA-ATVS-KP-SGYI----FTS-HL--LQK--PITT--------TSLSLPFIR :  38
Ricinus communis 1        : M--AAGLSA-APVF-KP-FRYI----NAS-YI--LPK--PLP---------TSLVPPSTR :  37
Citrus unshiu             : M--AVGLLA-AIVP-KP-FCLLTTKLQPSSLL--TTK--P-----------APLFAPLGT :  40
Diospyros kaki            : M--PAGISI-AASGPRSFILTR----KYP-LL--GPG--PESTISTP----TCLFSLPAR :  44
Vitis vinifera 2          : M--AVEISV-ATTS-----RLG----RNP-FL--GLK--PTSP-FTP----TSLFIPSIR :  38
Ipomea nil                : M--AVGISI-AASS-RT-VYSC----QFS-LV--RPA--THSA-SPP----SLLFSPLSR :  41
Solanum lycopersicum 1    : M--AAGISA-SASS-RT-IRLR----HNP-FL--SPK--SAST-APP----VLFFSPLTR :  41
Capsicum annuum 1         : M--AAEISI-SASS-RA-ICLQ----RNP-FP--APK--YFAT-APP----LLFFSPLTC :  41
Coffea arabica            : M--AAGIAV-AAGA-QT-VCFR----VNS-FL--TRK--PTSL-VAD----SLTLSPLAQ :  41
Solanum lycopersicum 2    : MAAAARISA-SSTS-RT-FYFR----HSP-FL--GPK--PTST-TSH----VSPISPFSL :  43
Capsicum annuum 2         : M-AAARISF-SSTS-RT-SYYR----HSP-FL--GPK--PTPT-TPS----VYPITPFSP :  42
Vitis vinifera 1          : M--ATGISA-SLNS-MS-CRLG----RNS-FT--ATG--PSSV-ISL----SSFLTPVTH :  41
Gentiana lutea            : METQFLVSG-RNSN----IHCRIDSISSS-SL--TPKSSPVST-STP----TLVVFPPFK :  47
Adonis aestivalis CHYb    : M--AAATSI-TSSS-RA-FRFH----RSL-FL--NTK--PNIR-NPP----CLLFSPLLM :  41
Daucus carota 1           : M--AAGISA-ASSS-TS-FSLG----RNP-FL--GPN--P-----------IWLFAPSVR :  35
Zea mays 2                : MAVARLVSA-----------------PFP------------------------LAPLRV :  18
Sorghum bicolor 1         : MAVARLLAA-----------------PFP------------------------LAPCRV :  18
Oryza sativa japonica 1   : MAVARLVAA-----------------RAP-LL--SPAAVAAAHRSPPALLR-LAFAPLPA :  39
Oryza sativa japonica 3   : MAVARLVVI------------------TPAVL----------------LGRTARVSPSAV :  26
Elaeis oleifera           : M--APGISA-SVTC-----GIG----RNP-FL-------RSPHRCLVAADCVQLFAPLPS :  40
Zea mays 3                : M--AAGLSG-AAMT----GFVA----KNP-LL--AAA--ARRRALPPLAGRALPFSPLTT :  44
Sorghum bicolor 2         : M--AAGLSG-AAMT----GFVA----KNP-LLAAAAA--ARRRA-HPLAGRALPFSPLTT :  45
Oryza sativa japonica 2   : M--ATGLSG-GAMT----SFAV----KKP-LL--AAA--VRRRSWPPPSGRALPFSPL-- :  42
Zea mays 1                : M--AAMTS----------FVA----KNP-LL--AAA--ARRRA-PPLAGRALPFSPLAS :  38
Tagetes erecta            : M--AAIAV-PCSS-RP-FGLG----RMR-LL--GHK--PTTITC------HFPFSFSIK :  40
Crocus sativus 1          : M--SAKISP-SATT-LA-ASFR----RPP------------------SGARIILLSSLPV :  33
Crocus sativus 2          : M--TASISP-AATT-LA-ASSR----RPP------------------AGARVILFSPLSV :  33
Narcissus pseudonarcissus : M--AVWISA-APPA----LAIS----SAP------------------RIRRVILFSPLHS :  31
Daucus carota 2           : M--AARLSV-GL------FRVG----KYE-IQ--VPK--QMTVVKTI----RELSAPFIG :  38
Ricinus communis 2        : M--ATAISA-TSSS-LR-HRFG----QVS-FP--TPK--SNNSTTIH----TLVYG---- :  38
Oncidium Gower Ramsay     : M--AFAMSS-SLTL-FQYQSFG----KKP-FF--SRR--RDFAG-------CSMMNPL-- :  38
Picea sitchensis 1        : MEMMSGVSS-------SCGLSRCDSILSP-LPALIKPAA-APLGRAVYRYDLAFARASSV :  51
Picea sitchensis 2        : M-----RSA-------PCGLSLCDSILSP-LLALTKPVALPPAGRAVSRYYVAFAGASTV :  47
Pinus taeda 1             : M-GLRSVSS-------PYGLSKCGSILSP-PLSSTKPAA-APLGRAVYCYYLALARAFSV :  50
Pinus taeda 2             : M--NSNLPA-SSSK-VPNQTAVKLFMSRP-----SPR--REILG---LGLENIGTAP--- :  43

Muriella zofingiensis     : MQTVFQPSS-VWHL------------RNRHVL--GDR--------------TCVPCR--- :  28
Haematococcus pluvialis   : M--LSKLQSISVKA------------RRVELA---------------------------- :  18
Chlamydomonas reinhardtii : MMLASRPAV-ALGA------------RAQPQV--LRP--------------TLVPRPGMV :  31
```

```
                                     *        80         *       100         *       120
Adonis aestivalis CBFD2   : M--------SPM----RKKKKHGDPCICSVAGR--TRN---LDIPQI---EEEEE---NV :  76
Adonis aestivalis CBFD1   : E--------SPM----RKKKTHRAACICSVAER--TRN---LDIPQI---EEEEE---NE :  73
Brassica rapa 1           : ------------L------RF-RGFSVCYVVEE--QRQ---SS--PV---DNDER---PE :  66
Brassica rapa 2           : -----------LL------RF-RGLSVCYVVSD--RRQ---SS--PI---NNDES---PE :  67
Arabidopsis thaliana 1    : ------------L------RF-KRFSVCYVVEE--RRQ---NS--PI---ENDER---PE :  67
Arabidopsis thaliana 2    : FN--------------GFRRR-KILTVCFVVEE--RKQ---SS--PM---DDDNK---PE :  68
Glycine max 2             : FH------HSTILRV-RPRRRMSGFTVCVLTED--SKE---IK--TV---EQEQE----- :  75
Medicago truncatula       : FP------HSKILETQRTRKS-TCFTVCVLTED--PKH---TS--QL---KTEEE----- :  76
Glycine max 1             : FH------HTASP---IPQNF-STFTVCVLMQD--PKQ---GT--QM---EIQAQ---EP :  74
Populus trichocarpa 1     : HQSF----LHYGT---KVPRK-TSLAVCFVVED--QTK---PSSAHI---ENQQE---EV :  79
Populus trichocarpa 2     : HQNL----LHYGF---KVPRK-TSFAVCFVVED--QTK---PISAHL---ENQQE---EE :  79
Ricinus communis 1        : Y-------NSLFL---GGRRK-TGFAICFVLED--LKQ---SV--ET---ENPVE---EE :  73
Citrus unshiu             : R--------HGFFNG-KNRRKINSFTVCFVLEE--KKQ---ST--QI---ETFTE---EE :  78
Diospyros kaki            : P-------FGAVL---RSGSR-TAKTVCFVVKE--EEL---AA--QL---EAGDQ---DE :  80
Vitis vinifera 2          : R-------HENIF---RCRKK-TRLTVCFVVEE--EKL---ST--EVV--ENRSE----- :  73
Ipomea nil                : RF------RSSVL---RSRRK-PSLTVCFVLED--EKL---ES-GGV---EIRAE---EI :  79
Solanum lycopersicum 1    : N-------FGAIL---LSRRK-PRLAVCFVLEN--EKL---NS--TI---ESESE---VI :  77
Capsicum annuum 1         : N-------LDAIL---RSRRK-PRLAACFVLKD--DKLYTAQS-GKQSDTEAIGD---EI :  84
Coffea arabica            : Q-------FSTTR---RHRRK-PRLTVCFVLED--EEL---KA--QLVTSEEEAR---ER :  80
Solanum lycopersicum 2    : N-------LGPIL---RSRRK-PSFTVCFVLED--EKL---KP--QF---DDEAE---DF :  79
Capsicum annuum 2         : N-------LGSIL---RCRRR-PSFTVCFVLED--DKF---KT--QF---EAGEE---DI :  78
Vitis vinifera 1          : L-------KGNIF---PLQRR-RSLKVCLVLEK--EIE---DG-IEI---EDDSP----- :  76
Gentiana lutea            : L-------VSKSL---RTRSK-PRLTVCFVLEE--KEL---KG--KLVVASDDDDGAGEV :  89
Adonis aestivalis CHYb    : ----------------RNRNGAGALTICFVAER--TRG---REIPQI---EEDEK---NM :  74
Daucus carota 1           : K-------LNPSL---RFKQK-SLTTVCFVVEP--RND---SS-GKP---ENNAD---RD :  72
Zea mays 2                : RA------PRPALP--PGAHAGPRPPVLALAPP--AAS---AA--PRRAVPARAA---PE :  60
Sorghum bicolor 1         : RR------PRQAL---PPAHAGPR----------------------------------PR :  35
Oryza sativa japonica 1   : RR------L--------------------------------AV--PLRVAVGEPE---PE :  56
Oryza sativa japonica 3   : PR------LRPIV---AGRRAV-------------------AA--PTRAVLGDGAGVGGE :  56
Elaeis oleifera           : NRCLRPSSISSALQ--PLRRPSRSAAVCFVVGQ--NRS---SE-MKE---AVEIE---SS :  86
Zea mays 3                : AR--------------APRRRGLGTVTCFVPQD--TEH---PA--AA---PAPVA---PV :  77
Sorghum bicolor 2         : TR--------------AARRRGLGTVTCFVPPD--TEHPAARA-PAA---PVPV----PV :  81
Oryza sativa japonica 2   : TR--------------TPRSRGLGTVTCFVPQG--TESQQAPA--PS---PPPTV---PV :  78
Zea mays 1                : TR--------------APRR----TVTCFVPQD--TAA---PA----------------A :  59
Tagetes erecta            : S-------FTPIV---RGRR----CTVCFVAGG--DSN---SN-SNN---NSDSN---SN :  74
Crocus sativus 1          : RRPVERRIRPPLL----HRRR-RTATVFFVLAE--EKT---TP--FL---DDVEE----- :  73
Crocus sativus 2          : RRVVD--LWPSALGQ-RRRRR-RTGTVFFVLAE--EAE---------------------- :  65
Narcissus pseudonarcissus : RQIG----WPPI----RNRRKRSKSTVFFASDV--DVG---KS-NGG---DGIVD---KI :  71
Daucus carota 2           : L---------------RLGNK-RKLSLCFAVEE--NESPVAAA--EA---EESSR---EV :  72
Ricinus communis 2        : ---------TKGF---SLRKN-QHLSICLVMDK--KAE-------EF---LDVMD---KH :  70
Oncidium Gower Ramsay     : ----------------VARCN-RATEICCVVRK--DGE---AE-SLV---EAEND---QL :  69
Picea sitchensis 1        : NRNGHR--SVRVFSEFRGGR--KVLPLFFALTE--KSQ---QQ-------ETESE---SI :  92
Picea sitchensis 2        : NRNGLR--SVRVFSEFHGRR--KVLPLCFALRE--KAQ---QQ-------ETESE---SI :  88
Pinus taeda 1             : NRNGFR--SARVFSEFRGRR--KVLSLVFASTE--KSQ---QL-------ETKSE---SI :  91
Pinus taeda 2             : ----------------ATYRL-ARLALCKMRMK--ETS------PQT---EAEKE---II :  72

Muriella zofingiensis     : ----------------------TSLSTCRHAVR--LVRANVAE-TQA---TPTTS---QM :  57
Haematococcus pluvialis   : -------------------RDITRPKVCLHAQRCSLVRLRVAA-PQT---EEALG---TV :  52
Chlamydomonas reinhardtii : SN-----------------------------LR--LQPVKVAD-PIV---ASETS---QV :  53
```

```
                                     *       140         *       160         *       180
Adonis aestivalis CBFD2   : EELIE----------QTDSDIVHIKK---------TLGGKQSKRPTGSIVAPVSCLGILS : 117
Adonis aestivalis CBFD1   : EELIE----------QTDSGIIHIKK---------TLGGKQSRRSTGSIVAPVSCLGILS : 114
Brassica rapa 1           : R----------TNVIDPELLALRLAE---------KLERKKSERFTYLIAAVMSSFGITS : 107
Brassica rapa 2           : KTSSL-----DTNAIDAEYLALRLAE---------KLERKKSERFTYLIAAVMSSFGITS : 113
Arabidopsis thaliana 1    : STSS-------TNAIDAEYLALRLAE---------KLERKKSERSTYLIAAMLSSFGITS : 111
Arabidopsis thaliana 2    : STTSS----------SEILMTSRLLK---------KAEKKKSERFTYLIAAVMSSFGITS : 109
Glycine max 2             : -------------QVIPQAVSAGVAE---------KLARKKSQRFTYLVAAVMSSFGITS : 113
Medicago truncatula       : ----------------------IVAQ---------KLARKKSQRFTYLVAAVMSSFGVTS : 105
Glycine max 1             : PPSP-----------PQQVLSQKLAE---------KLARKESESFTYLIAAVMSSFGITS : 114
Populus trichocarpa 1     : PKDVN----------ENQISTPRVAE---------RLERKRKERVTYLIAAVMSSLGITS : 120
Populus trichocarpa 2     : PKDVN----------KNQILTPRVAE---------RLARKQRERDTYLIAAVMSSLGITS : 120
Ricinus communis 1        : NEKV-----------NYQILTPRVAE---------RLARKRSERFTYLVAAVMSSFGITS : 113
Citrus unshiu             : EEESG----------TQISTAARVAE---------KLARKRSERFTYLVAAVMSSFGITS : 119
Diospyros kaki            : TSGV---------ETDKRISAARVEE---------KLARKRSERFTYLVAAVMSSFGITS : 122
Vitis vinifera 2          : -ETV-----------ASQISAARVAE---------KLARKRSERLTYLVAAVMSSFGITS : 112
Ipomea nil                : ERAI-----------EKQISASRLAE---------KLARKRSERFTYLVAAVMSSLGITS : 119
Solanum lycopersicum 1    : EDRIQ-----VEINEEKSLAASWLAE---------KLARKKSERFTYLVAAVMSSLGITS : 123
Capsicum annuum 1         : EVETN----------EEKSLAVRLAE---------KFARKKSERFTYLVAAVMSSLGITS : 125
Coffea arabica            : EKAM-----------AKRISDARTAE---------KLARKRSERFTYLVAAVMSSFGITS : 120
Solanum lycopersicum 2    : EKKI-----------EEQILATRLAE---------KLARKKSERFTYLVAAIMSSFGITS : 119
Capsicum annuum 2         : EMKI-----------EEQISATRLAE---------KLARKKSERFTYLVAAVMSSFGITS : 118
Vitis vinifera 1          : ------------------ESSNRASE---------RLARKKAERYTYLVAAMMSSLGITS : 109
Gentiana lutea            : RKQR---------EKEISASAEKLAQ---------KLARKKSERFTYLVAAVMSSFGITS : 131
Adonis aestivalis CHYb    : DEVFE----------QMNSASVRVAE---------KLARKKSERFTYLIAALMSSMGITS : 115
Daucus carota 1           : EVSR--------EEIEAGSCSVRVEE---------RRARKKSERFTYLVAAVMSSLGITS : 115
Zea mays 2                : DGGRG--------------DAAAAAA---------RAARKQSERRTYLVAAVMSSLGITS :  97
Sorghum bicolor 1         : DRGDT----------EAEAARVAVAE---------RAARKQSERRTYLVAAVMSSLGITS :  76
Oryza sativa japonica 1   : ED-----------------ARRAVAE---------RAARKQSERRTYLVAAVMSSLGITS :  90
Oryza sativa japonica 3   : EDAV-----------VAVVEEDAVAR---------RAARKRSERRTYLVAAVMSSLGFTS :  96
Elaeis oleifera           : DPAAE-------EEEARRILLDRTAE---------KIARKQAERRTYLAAAVLSSLGITS : 130
Zea mays 3                : PETAL-------DEEARAAAARRVAE---------RKARKRSERRTYLVAAVMSSLGVTS : 121
Sorghum bicolor 2         : PETAL-------DEEARAAAARRIAE---------KKARKRSERRTYLVAAMMSSLGFTS : 125
Oryza sativa japonica 2   : PVPSL-------EEEAAAAAARRIAE---------RKARKLSERRTYLVAAVMSSLGFTS : 122
Zea mays 1                : PVPAL-------DEEARAAAARRVAE---------KEARKRSERRTYLVAAVMSSLGVTS : 103
Tagetes erecta            : NPGLD-------LNPAVMNRNRLVEE---------KMERKKSERFTYLVAAIMSTFGITS : 118
Crocus sativus 1          : --------------EKSIAPSNRAAE---------RSARKRSERTTYLITAVMSSFGITS : 110
Crocus sativus 2          : ---------------KRMAPSNRAAE---------RSARKRSERRTYLIAAVMSSLGITF : 101
Narcissus pseudonarcissus : ERLK--------KQEQLMISKSRTTE---------RMERKRSERTTYLIAAMMSSLGITS : 114
Daucus carota 2           : EKQI--------IESFTVAGGSRAEE---------RMARKKTERFTYLVAAVMSSFGITS : 115
Ricinus communis 2        : ENDVN----------KKQPLNSRVER---------KLARKKLERFTYLVAAILSSTGITS : 111
Oncidium Gower Ramsay     : EEEVM---------KPTSIDSFSVVL---------RSERKKAERRTYLVAAMASSLGFTF : 111
Picea sitchensis 1        : EDDDS------TVTEFADSLSSRVDENENKMQWDKRAARRKAERHAYFFAAVASSVGITS : 146
Picea sitchensis 2        : EDEDS-------VTEFADSLSSRVYE--NEKQWDKRAARRKAERHVYFFAAIASSVGITS : 139
Pinus taeda 1             : EDDVS-------ATKFADSLSSRVDEIANKREWDKRAATRKSERHAYFFAAVASSVGITS : 144
Pinus taeda 2             : DEPLQ--SKLASELMKRDENLRKVSG---------QRVRKKAERYAYLFAAIASSVGITS : 121

Muriella zofingiensis     : LEEVHDESHAASASSQIFELAVKTSI---------KRQQRNRQQLTYQGSAIAASLGVGA : 108
Haematococcus pluvialis   : QAAGAGDEH--SADVALQQLDRAIAE---------RRARRKREQLSYQAAAIAASIGVSG : 101
Chlamydomonas reinhardtii : MEA--------PQEKKLSEFELKRLE---------RKQQRAQEAATYKFSAIAATVLVLS :  96
```

```
                                      *       200         *       220         *       240
Adonis aestivalis CBFD2    : MIGPAVYFKFSRLME--GGDIPVAEMGITFATFVAAAVGTEFLSAWVHKELWH--ES-LW : 172
Adonis aestivalis CBFD1    : MIGPAVYFKFSRLME--CGDIPVAEMGITFAAFVAAIGTEFLSGWVHKELWH--DS-LW : 169
Brassica rapa 1            : MAVMAVYYRFSWQME--GGVIPMSEMFGTFALSVGAAVGMEFWARWAHRALWH--AS-LW : 162
Brassica rapa 2            : MAIMAVYYRFSWQME--GGEIPMSEMFGTFALSVGAAVGMEFWARWAHKALWH--AS-LW : 168
Arabidopsis thaliana 1     : MAVMAVYYRFSWQME--GGEISMLEMFGTFALSVGAAVGMEFWARWAHRALWH--AS-LW : 166
Arabidopsis thaliana 2     : MAIMAVYYRFSWQMK--GGEVSVLEMFGTFALSVGAAVGMEFWARWAHRALWH--DS-LW : 164
Glycine max 2              : MAVFAVYYRFSWQME--GGDVPWSEMLGTFSLSVGAAVAMEFWARWAHRALWH--AS-LW : 168
Medicago truncatula        : MAILAVYYRFSWQME-GSGEVPWSEMFGTFALSVGAAVGMEFWARWAHEILWH--AS-LW : 161
Glycine max 1              : MAVFAVYYRFAWQME--GGEVPLSEIFGTFALSVGAAVGMEFWARWAHRALWH--AS-LW : 169
Populus trichocarpa 1      : MAVMAVYYRFYW-LE--GGEVPLSEMFGTFALSVGAAVGMEFWARWAHRALWH--AS-LW : 174
Populus trichocarpa 2      : TAVLAVYYRFYW-LE--GGKVSWPEMFGTFALSVGAAVGMEFWARWAHKELWH--AS-LW : 174
Ricinus communis 1         : MAVMACYYRFYWQME--GGEVPLAEMFGTFSLSVGAAVGMEFWARWAHRALWH--AS-LW : 168
Citrus unshiu              : MAVMAVYYRFWWQME--GGEVPLAEMFGTFALSVGAAVGMEFWARWAHKALWH--AS-LW : 174
Diospyros kaki             : MAVLAVYYRFSWQME--GGEIPYSEMFGTFALSVGAAVGMEFWARWAHRALWH--AS-LW : 177
Vitis vinifera 2           : MAVMAVYYRFSWQME--GGEVPLSEMFGTFALSVGAAVGMEFWARWAHKALWH--AS-LW : 167
Ipomea nil                 : MAVLAVYYRFAWQME--GGEVPYTEMFGTFALSVGAAVGMEFWARWAHRALWH--AS-LW : 174
Solanum lycopersicum 1     : MAILAVYYRFSWQME--GGEVPFSEMLATFTLSFGAAVGMEYWARWAHRALWH--AS-LW : 178
Capsicum annuum 1          : MAVISVYYRFSWQME--GGEMPFSEMFCTFALAFGAAIGMEYWARWAHRALWH--AS-LW : 180
Coffea arabica             : MAVLAVYYRFVWQME--GGEVPYSEMFGTFALSVGAAVGMEFWARWAHKALWH--AS-LW : 175
Solanum lycopersicum 2     : MAVMAVYYRFSWQME--GGEVPVTEMLGTFALSVGAAVGMEFWARWAHKALWH--AS-LW : 174
Capsicum annuum 2          : MAVMAVYYRFYWQME--GGEVPFSEMFGTFALSVGAAVGMEFWARWAHKALWH--AS-LW : 173
Vitis vinifera 1           : MAIVAVYYRLSWQME--GGEIPVLEMLGTFALSVGAAVGMEFWARWAHKALWH--AS-LW : 164
Gentiana lutea             : MAVLSVYYRFSWQME--GGEIPLSEMFGTFALSVGAAVGMEFWARWAHEALWH--AS-LW : 186
Adonis aestivalis CHYb     : MAILSVYYRFSWQME--GGDIPVTEMLGTFALSVGAAVGMEFWARWAHRALWH--AS-LW : 170
Daucus carota 1            : MAVLAVYYRFSWQME--GGEIPYSEMIGTFALSVGAAVGMEFWARWAHEALWH--AS-LW : 170
Zea mays 2                 : MAAAAVYYRFAWQME-GGGAIPVTEMVGTFALSVGAAVGMEFWARWAHRALWH--AS-LW : 153
Sorghum bicolor 1          : MAAAAVYYRFAWQME-GGGEIPVTEMVGTFALSVGAAVGMEFWARWAHRALWH--AS-LW : 132
Oryza sativa japonica 1    : MAAAAVYYRFAWQME-GSGEIPVTEMFGTFALSVGAAVGMEFWARWAHRALWH--AS-LW : 146
Oryza sativa japonica 3    : MAAAAVYYRFAWQMEAGGGDVPATEMVGTFALSVGAAVGMEFWARWAHRALWH--AS-LW : 153
Elaeis oleifera            : MAVAAVYYRFYWQME--GGEVPLTEMLGTFALSVGAAVGMEFWARWAHRALWH--AS-LW : 185
Zea mays 3                 : MAVAAVYYRFSWQME--GGAVPVSEMFGTFALSVGAAVGMEFWARWAHRALWH--AS-LW : 176
Sorghum bicolor 2          : MAVAAVYSRFSWQME--GGEVPVSEMLGTFALSVGAAVGMEFWARWAHRALWH--AS-LW : 180
Oryza sativa japonica 2    : MAVAAVYYRFHWQLE--GGDVPMTEMFGTFALSVGAAVGMEFWAQWAHRSLWH--AS-LW : 177
Zea mays 1                 : MAVAAVYYRFSWQME--GGEVPVIETLGTFALSVGAAVGMEFWARWAHRALWH--AS-LW : 158
Tagetes erecta             : MAVMAVYYRFSWQME--GGEIPYVEMFGTFALSVGAAVGMEYWARWAHEALWH--AS-LW : 173
Crocus sativus 1           : MAAAAVYYRFAWQME--GGDVPVTEMAGTFALSVGAAVGMEFWARWAHRALWH--AS-LW : 165
Crocus sativus 2           : MAAAAVYYRFAWQME--GGDVPMTEMAGTFRLSVGGGPGDEFWPRWPHRASAT--RR-SW : 156
Narcissus pseudonarcissus  : MAIVSVYYRFAWQME--EGEIPVTEMLGTFALSVGAAVGMEFWARWAHRALWH--AS-LW : 169
Daucus carota 2            : MAVFAVYSRFAWQME--GGEVPHSEMLATFSLAFGAAVGMEYWARWAHKALWH--DS-LW : 170
Ricinus communis 2         : IAAMSVYYRFSWQME--GGEFPALEMFGTFALSVGAAVGMEFWARWAHRALWH--AS-LW : 166
Oncidium Gower Ramsay      : MAAAAVYYRFAWQME--GGAVPLTEMMGTFSLAVGSAVGMEYWARWAHRALWH--TS-LW : 166
Picea sitchensis 1         : MAAGAVYYRFVWQMQ--GAEVPYPEIFGTFALAVGATVGMEFWARWAHRALWH--AS-LW : 201
Picea sitchensis 2         : MAAAAVYYRFVWLMQ--GAEVPYTEMFGTFALAVGATVGMEYWARWAHRALWH--AS-LW : 194
Pinus taeda 1              : MTAAAVYYRFVWQMQ--GAQIPYTEIFGTFALAVGATVGMEFWARWAHKALWH--AS-LW : 199
Pinus taeda 2              : MAAGAVYYRFVWQVQ--GGEVPLTEIFGTFSLAVGATVGMEAWARWAHRALWH--SS-LW : 176

Muriella zofingiensis      : LAVAATHYKFSYHMP-SESPFPWLDMAGTLALVIGGVFGMEMWARYAHKALWHDFQP-GW : 166
Haematococcus pluvialis    : IAIFATYLRFAMHMT-VGGAVPWGEVAGTLLLVVGGALGMEMYARYAHKAIWHE-SPLGW : 159
Chlamydomonas reinhardtii  : IAVVATYYRFAWHFA-EDGDLPVDEMAATLLLVFGGMFGMEMYARFAHKVLWHDFEP-GW : 154
```



```
                                     *       260         *       280         *       300
Adonis aestivalis CBFD2   : YIHKSHHRSR-KGRFEF---NDVFAIINALPAIALINYGFSNEGLLPGACFGVGLGTTVC : 228
Adonis aestivalis CBFD1   : YIHKSHHRSR-KGRFEF---NDVFAIINALPAIALINYGFSNEGLLPGACFGTGLGTTVC : 225
Brassica rapa 1           : NMHESHHKPR-EGPFEL---NDVFAIINAVPAIGLLSYGFFNKGLVPGLCFGAGLGITVF : 218
Brassica rapa 2           : NMHESHHKPR-EGPFEL---NDVFAIVNAVPAIALLSYGFFNKGLVPGLRFGAGLGITVF : 224
Arabidopsis thaliana 1    : NMHESHHKPR-EGPFEL---NDVFAIVNAGPAIGLLSYGFFNKGLVPGLCFGAGLGITVF : 222
Arabidopsis thaliana 2    : NMHESHHKPR-EGAFEL---NDVFAITNAVPAIGLLYYGFLNKGLVPGLCFGAGLGITMF : 220
Glycine max 2             : HMHESHHRPR-EGPFEL---NDVFAIINAVPAIALLSYGIFHKGLVPGLCFGAGLGITVF : 224
Medicago truncatula       : HMHQSHHRPR-EGAFEL---NDVFAIINAVPAIALLSNGFFNKGLRPGLCFGAGLGITVF : 217
Glycine max 1             : HMHESHHRPR-EGPFEP---NDVFAIINAVPAIALLSYGFFNKGLVPGLCFGAGLGITVF : 225
Populus trichocarpa 1     : HMHESHHRPR-DGPFEL---NDVFAITNAVPAISLLSYGFFNKGLVPGLCFGAGLGITVF : 230
Populus trichocarpa 2     : NMHESHHRPR-DGPFEL---NDVFAIINAVPAISLVAYGFFNEGLVPGLCFGAGLGITVF : 230
Ricinus communis 1        : HMHESHHRPR-EGPFEL---NDVFAITNAVPAIGLLYYGFFNKGLFPGLCFGAGLGITVF : 224
Citrus unshiu             : HMHESHHRPR-EGPFEL---NDVFAIINAVPAIALLSFGFFHKGLVPGLCFGAGLGITVF : 230
Diospyros kaki            : HMHESHHKPI-EGPFEL---NDVFAIINAVPAIALLSYGFFHKGLVPGLCFGAGLGITVF : 233
Vitis vinifera 2          : HMHESHHKPR-EGPFEL---NDVFAITNAVPAIALLSYGFFHKGLVPGLCFGAGLGITVF : 223
Ipomea nil                : HMHESHHRPR-EGPFEL---NDVFAITNAVPAIALLSYGFFHKGLVPGLCFGAGLGITVF : 230
Solanum lycopersicum 1    : HMHESHHRPR-EGPFEM---NDVFAITNAVPAIGLLSYGFFHKGIVPGLCFGAGLGITVF : 234
Capsicum annuum 1         : HMHESHHRPR-EGPFEL---NDIFAIINAVPAIAFFSFGFNHKGLIPGICFGAGLGITVF : 236
Coffea arabica            : HMHESHHRPR-EGPFEL---NDVFAIINAVPAIALLSYGFFHKGLIPGLCFGAGLGITVF : 231
Solanum lycopersicum 2    : HMHESHHKPR-EGPFEL---NDVFAITNAVPAIALLNYGFFHKGLIAGLCFGAGLGITVF : 230
Capsicum annuum 2         : HMHESHHKPR-EGPFEL---NDVFAIINAVPAIALLDYGFFHKGLIPGLCFGAGLGITVF : 229
Vitis vinifera 1          : HMHESHHRPR-EGPFEL---NDVFAIINAVPAISLLSYGLFNKGLVPGLCFGAGLGITVF : 220
Gentiana lutea            : HMHESHHKPR-EGPFEL---NDIFAIINAVPAIALLSYGFFHKGLIPGLCFGAGLGITVF : 242
Adonis aestivalis CHYb    : HMHESHHKPR-EGPFEL---NDVFAIINAVPAIALLNFGFFHKGLIPGLCFGAGLGITVF : 226
Daucus carota 1           : HMHESHHKPR-EGAFEL---NDVFAIMNAVPAIALLAYGFFHKGYFPGLCFGAGLGITVF : 226
Zea mays 2                : HMHESHHRARDDGPFEL---NDVFAIVNAVPAMSLLAYGFFNRGLVPGLCFGAGLGITLF : 210
Sorghum bicolor 1         : HMHESHHRPR-DGPFEL---NDVFAIVNAVPAMSLLAYGFFNRGLVPGLCFGAGLGITLF : 188
Oryza sativa japonica 1   : HMHESHHRPR-DGPFEL---NDVFAITNAVPAMSLLAYGFFTRGLVPGLCFGAGLGITLF : 202
Oryza sativa japonica 3   : HMHESHHRPR-DGPFEL---NDVFAIANAAPAISLLAYGLLNRGLLPGLCFGAGLGITLF : 209
Elaeis oleifera           : HMHESHHRPR-DGPFEL---NDVFAIINAVPAISLLSFGFFNRGLVPGLCFGAGLGITLF : 241
Zea mays 3                : HMHESHHRPR-EGPFEL---NDVFAIVNAVPAISLLAYGFFHRGLVPGLCFGAGLGITLF : 232
Sorghum bicolor 2         : HMHESHHRPR-EGPFEL---NDVFAIINAVPAICLLAYGFFHRGLVPGLCFGAGLGITLF : 236
Oryza sativa japonica 2   : HMHESHHRAR-EGPFEL---NDVFAITNAVPAISLLAYGFFHRGIVPGLCFGAGLGITLF : 233
Zea mays 1                : HMHESHHRPR-EGPFEL---NDVFAIVNAAPAISLLAYGFFHRGIVPGLCFGAGLGITLF : 214
Tagetes erecta            : HMHESHHKPR-EGPFEL---NDVFAITNAVPAIALLSYGFFHKGIIPGLCFGAGLGITVF : 229
Crocus sativus 1          : HMHESHHRPR-EGPFEL---NDVFAIINAVPAIALLNFGFFHRGLLPGLCFGAGLGITLF : 221
Crocus sativus 2          : HMHES-HPPA-EGKGPLSSSNDVFAIINAVPAIALLAFGFFHRGLLPGLCFGAGLGITLF : 214
Narcissus pseudonarcissus : HMHESHHKPR-DGPFEL---NDVFAVINAVPAISLLYYGFFNRGLVPGLCFGAGLGITLY : 225
Daucus carota 2           : HMHESHHKPR-EGAFEL---NDVFAIINAVPAIALLAYGFFHKGLFPGLCFGAGLGITVF : 226
Ricinus communis 2        : KMHESHHKSR-EGPFEL---NDVFAIINAVPAIALLSYGFNNKGLLPGLCFGAGLGITVF : 222
Oncidium Gower Ramsay     : HMHESHHRPR-DGPFEL---NDVFALINAFPAVALLAFGFFHRGFFSGLCFGAGLGITLY : 222
Picea sitchensis 1        : HMHESHHRPR-EGPFEL---NDVFAIINAVPAIALMAYGFFNKGFVPGLCFGAGLGITVF : 257
Picea sitchensis 2        : HMHESHHRPR-EGPFEL---NDVFAIINAVPAIALLAYGFFNKGLVPGLCFGAGLGITVF : 250
Pinus taeda 1             : HMHESHHRPR-EGPFEL---NDVFAIINAIPAIALMAYGFFNKGFVPGLCFGAGLGITVF : 255
Pinus taeda 2             : SMHESHHRTR-EGPFEL---NDVCAITNAVPAISLLSYGFFNKGFVPGLCFGAGLGITVF : 232

Muriella zofingiensis     : ALHKSHHEPR-IGPFEA---NDIFAVINAVPAFSLCLYGFLTPNLVGSLCFGAGLGITLF : 222
Haematococcus pluvialis   : LLHKSHHTPR-TGPFEA---NDLFAIINGLPAMLLCTFGFWLPNVLGAACFGAGLGITLY : 215
Chlamydomonas reinhardtii : ALHKSHHEPR-TGPFEL---NDIYAVANALPAMALCAYGFFTPHVIGGVCFGAGLGITLF : 210
```

```
                                     *       320         *       340         *       360
Adonis aestivalis CBFD2   : GMAYIFLHNGLSHRRFPVWLIANVPYFHKLAAAHQIHHSGKFQGVPFGLFLGPKELEEVR : 288
Adonis aestivalis CBFD1   : GMAYIFLHNGLSHRRFPVGLIANVPYFHKLAAAHQIHHSGKFQGVPFGLFLGPQELEEVR : 285
Brassica rapa 1           : GIAYMFVHDGLVHKRFPVGPIADVPYLRKVAAAHQLHHTDKFDGVPYGLFLGPKELEEV- : 277
Brassica rapa 2           : GIAYMFVHDGLVHKRFPVGPIADVPYLRKVAAAHQLHHTDKFDGVPYGLFLGPKELEEV- : 283
Arabidopsis thaliana 1    : GIAYMFVHDGLVHKRFPVGPIADVPYLRKVAAAHQLHHTDKFNGVPYGLFLGPKELEEV- : 281
Arabidopsis thaliana 2    : GMAYMFVHDGLVHKRLPVGPIANVPYLRKVAAAHQLHHTDKFKGVPYGLFLGPKEVEEV- : 279
Glycine max 2             : GMAYMFVHDGLVHKRFPVGPIANVPYFRRVAAAHQLHHSDKFNGAPYGLFLGPKEVEEV- : 283
Medicago truncatula       : GIAYMFVHDGLVHKRFPVGPIANVPYFTRVGVAHQLPHWDKFKGVPYGLFLGPKEVEKV- : 276
Glycine max 1             : GMAYMFVHDGLVHKRFPVGPIANVPYLRRVASAHQLHHSEKFDGVPYGLFMGPKEVEEV- : 284
Populus trichocarpa 1     : GIAYMFVHDGLVHKRFPVGPIADVPYFRKVAAAHKIHHSDKFNGVPYGLFLGPREIEEV- : 289
Populus trichocarpa 2     : GMAYMFVHDGLVHKRFPVGPIADVPYFTRVAAAHQIHHSEKFNGVPYGLFLGHKEIEEV- : 289
Ricinus communis 1        : GMAYMFVHDGLVHKRFPVGPIANVPYFRKVAAAHQLHHSEKFNGVPYGLFLGPKEIEEV- : 283
Citrus unshiu             : GMAYMFVHDGLVHKRFPVGPIADVPYFRRVAAAHQLHHSDKFHGVPYGLFLGPKELEEV- : 289
Diospyros kaki            : GMAYMFVHDGLVHKRFPVGPIADVPYFRRVAAAHQLHHSDKFNGVPYGLFLGPKEVEEV- : 292
Vitis vinifera 2          : GMAYMFVHDGLVHRRFPVGPIANVPYFRRVAAAHQLHHSDKFNGVPYGLFLGPKEVEEV- : 282
Ipomea nil                : GMAYMFVHDGLVHKRFPVGPIADVPYFRRVAAAHQLHHTDKFNGVPYGLFLGPKELEEV- : 289
Solanum lycopersicum 1    : GMAYMFVHDGLVHKRFPVGPIANVPYFRRVAAAHQLHHSDKFDGVPYGLFLGPKELEEV- : 293
Capsicum annuum 1         : GMAYMFVHDGLVHKRFPVGPIAKVPYFQRVAAAHQLHHSDKFDGVPYGLFLGPKELEEV- : 295
Coffea arabica            : GMAYMFVHDGLVHKRFPVGPIANVPYFRRVAAAHQLHHSDKFNGVPFGLFLGPKELEKV- : 290
Solanum lycopersicum 2    : GMAYMFVHDGLVHKRFPVGPVANVPYLRKVAAAHSLHHSEKFNGVPYGLFFGPKELEEV- : 289
Capsicum annuum 2         : GMAYMFVHDGLVHKRFPVGPVANVPYLRKVAAAHSLHHSEKFNGVPYGLFLGPKELEEV- : 288
Vitis vinifera 1          : GMAYMFVHDGLVHRRFPVGPIANVPYLRKVASAHQLHHSDKFNGVPYGLFLGPMELEEV- : 279
Gentiana lutea            : GMAYMFVHDGLVHKRFPVGPIADVPYFRRVAAAHTLHHSDKFNGVPYGLFLGPKELEEV- : 301
Adonis aestivalis CHYb    : GMAYMFVHDGLVHRRFPVGPIANVPYFRKVAAAHQIHHTDKFQGVPYGLFLGPKELEEV- : 285
Daucus carota 1           : GIAYMFVHDGLVHKRFAVGPIADVPYFRKVAAAHQLHHMEKFKGVPYGLFLGPKEVEDV- : 285
Zea mays 2                : GMAYMFVHDGLVHRRFPVGPIENVPYFRRVAAAHQIHHMDKFQGVPYGLFLGPKELKEV- : 269
Sorghum bicolor 1         : GMAYMFVHDGLVHRRFPVGPIENVPYFRRVAAAHQIHHMDKFQGVPYGLFLGPKELKEV- : 247
Oryza sativa japonica 1   : GMAYMFVHDGLVHRRFPVGPIANVPYFRRVAAAHQIHHMDKFEGVPYGLFLGPKELEEV- : 261
Oryza sativa japonica 3   : GMAYMFVHDGLVHRRFPVGPIENVPYFRRVAAAHQIHHTDKFEGVPYGLFLGPKELEEV- : 268
Elaeis oleifera           : GMAYMFVHDGLVHRRFPVGPIANVPYFRRVAAAHQIHHMDKFDGVPYGLFLGPKELEEV- : 300
Zea mays 3                : GMAYMFVHDGLVHRRFPVGPIANVPYFRRVAAAHKIHHMDKFEGVPYGLFLGPKELEEV- : 291
Sorghum bicolor 2         : GMAYMFVHDGLVHRRFPVGPIANVPYFRRVAAAHKIHHMDKFEGVPYGLFLGPKELEEV- : 295
Oryza sativa japonica 2   : GMAYMFVHDGLVHRRFPVGPIANVPYFRRVAAAHKIHHTDKFEGVPYGLFLGPKELEEV- : 292
Zea mays 1                : GMAYMFVHDGLVHRRFPVGPIADVPYFRRVAASHKIHHMDKFGGVPYGLFLGPKELEEV- : 273
Tagetes erecta            : GMAYMFVHDGLVHRRFQVGPIANVPYLRRVAAAHQLHHTEKFNGVPYGLFLGPKELEEV- : 288
Crocus sativus 1          : GIAYMFVHDGLVHRRFPVGPIADVPYFQRVAAAHQIHHSEKFEGVPYGLFMGPKELEEI- : 280
Crocus sativus 2          : GIAYMFVHDGLVHRRFPVGPIADLPYFQRVAAAHQIHHSEKFQGVPYGLFMGPKELEEV- : 273
Narcissus pseudonarcissus : GMAYMFVHDGLVHRRFPVGPIADVPYFRRVAAAHRIHHTEKFNGVPYGLFLGPKELEEV- : 284
Daucus carota 2           : GMAYMFVHDGLVHKRFPVGPVANVPYLRKVAAAHQLHHMEKFNGVPYGLFLGPQELEQV- : 285
Ricinus communis 2        : GMAYMFVHDGLVHRRFSVGPIANVPYLRRVAAAHQLHHSDKFGGVPYGLFLGPKELEEKE : 282
Oncidium Gower Ramsay     : GMAYMFVHDGLVHRRFPVGPIATVPYFQWVAAAHQIHHADKFNGVPYGLFLGHKELEEV- : 281
Picea sitchensis 1        : GMAYMFVHDGLVHRRFPVGPIADVPYLLKVAAAHQLHHADKFNGVPYGLFLGPKELEEV- : 316
Picea sitchensis 2        : GMAYMFVHDGLVHRRFPVGPIADVPYLLKVAAAHQLHHADKFNGVPYGLFLGPKELEEV- : 309
Pinus taeda 1             : GMAYMFVHDGLVHRRFPVGPIADVPYLLKVAAAHQLHHAYKFNGVPYGLFLGPKELEEV- : 314
Pinus taeda 2             : GMAYMFVHDGLVHRRFPVGPIANVPYLQRVAAAHQLHHADKYNGVPYGLFLGPKELEEV- : 291

Muriella zofingiensis     : GIMYMFIHDGLVHKRFPVGPIAQMPAMKRVAIAHKLHHSEKYGGVPWGMFFGPQELEAI- : 281
Haematococcus pluvialis   : GMAYMFVHDGLVHRRFPTGPIAGLPYMKRLTVAHQLHHSGKYGGAPWGMFLGPQELQHIP : 275
Chlamydomonas reinhardtii : GIAYMFFHDGLVHRRFPVGPIANLPYMKRIMVAHQIHHTNKFGGVPFGMFLGVQELEAVP : 270
```

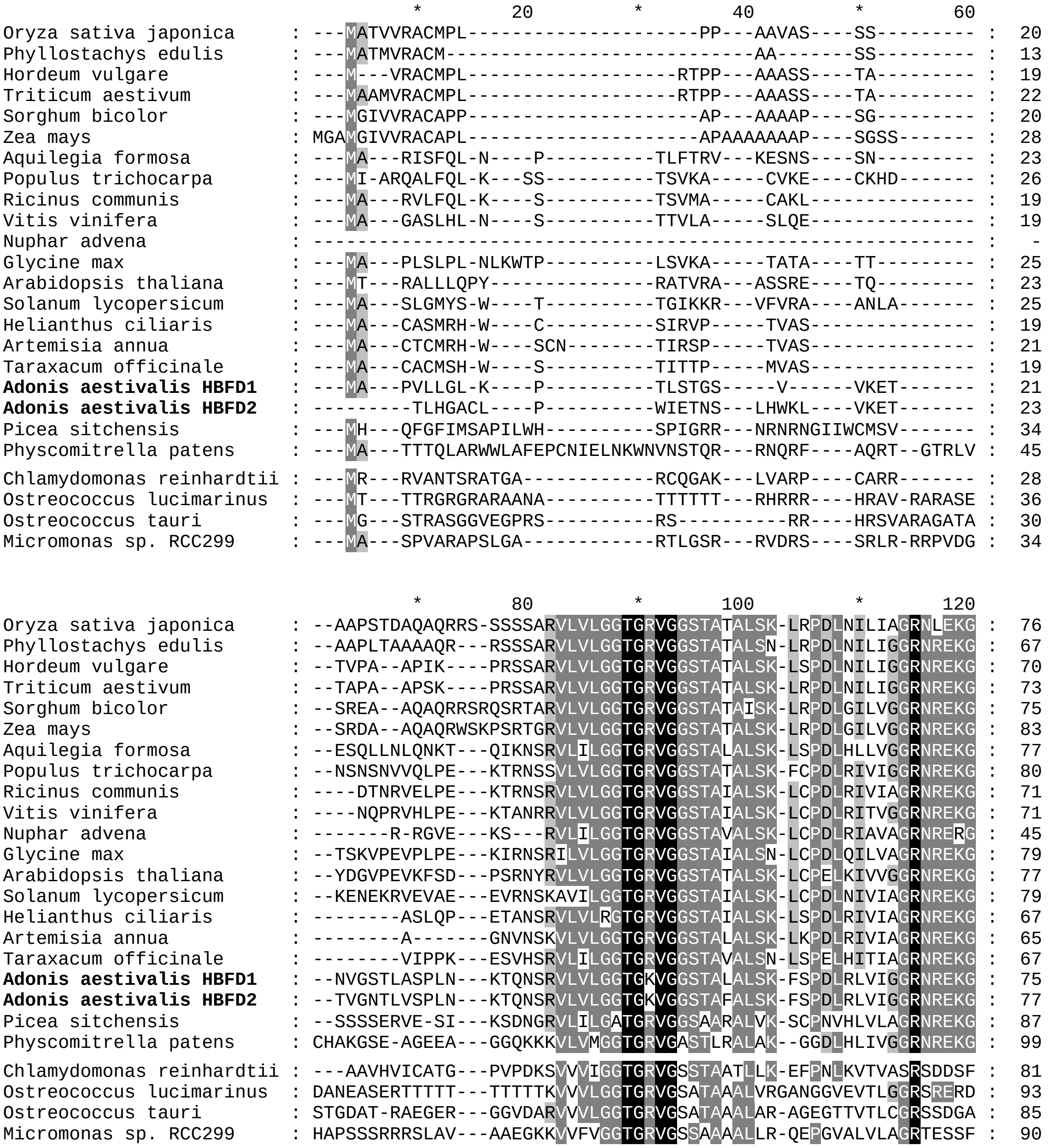
```
                                     *       380         *       400         *
Adonis aestivalis CBFD2   : GGTEEL--------------ERVISR------TTKRT--QPST-------- : 309
Adonis aestivalis CBFD1   : GGTEEL--------------ERVISR------TAKRT--QSST-------- : 306
Brassica rapa 1           : GGDEEL--------------DKEISR------RIKLY--KKSSSS------ : 300
Brassica rapa 2           : GGDEEL--------------EKEISR------RIKLY--KKSSSS------ : 306
Arabidopsis thaliana 1    : GGNEEL--------------DKEISR------RIKSY--KKASGSGSSSSS : 310
Arabidopsis thaliana 2    : GGKEEL--------------EKEISR------RIKLY--NKGSSTS----- : 303
Glycine max 2             : GGLEEL--------------EKEISR------RIRSG--S----------- : 301
Medicago truncatula       : GGIEEL--------------EKKISR------RTRSY--TGS--------- : 296
Glycine max 1             : GGLEEL--------------EKEISR------RARSY--KIAREKS----- : 308
Populus trichocarpa 1     : GGLEEL--------------EREISR------RTKSY--KEL--------- : 309
Populus trichocarpa 2     : GGQEEL--------------EREINR------RTKSS--KGL--------- : 309
Ricinus communis 1        : GGLEEL--------------EKETSR------RKKSY--NGR--------- : 303
Citrus unshiu             : GGLEEL--------------EKEISK------RIKSY--NRVPK------- : 311
Diospyros kaki            : GGDEEL--------------EKEVNR------RTRLS--KGS--------- : 312
Vitis vinifera 2          : GGKDEL--------------EKEINR------RARLS--NGPR-------- : 303
Ipomea nil                : GGLNDL--------------EVEVNR------RIKMS--STGR-------- : 310
Solanum lycopersicum 1    : GGLEEL--------------EKEVNR------RIKIS--KGLL-------- : 314
Capsicum annuum 1         : GVIEEL--------------EKEVNR------RIKSL--KRL--------- : 315
Coffea arabica            : GGLEEL--------------EKEINR------RIKLR--KGS--------- : 310
Solanum lycopersicum 2    : GGTEEL--------------EKEVIR------RTRLS--KGS--------- : 309
Capsicum annuum 2         : GGLEEL--------------EKEVNR------RTRYI--KGS--------- : 308
Vitis vinifera 1          : GGMEEL--------------EKEISR------RIKSS--DSS--------- : 299
Gentiana lutea            : GGLQVL--------------EMEINR------RTKNN--QS---------- : 320
Adonis aestivalis CHYb    : GGNEEL--------------EKEIER------RIKRM--NAL--------- : 305
Daucus carota 1           : GGHEAL--------------ELEINR------RIKSS--ASRASRS----- : 309
Zea mays 2                : GGTEEL--------------EKEIKR------RIRRR--EALDATQ----- : 293
Sorghum bicolor 1         : GGTEEL--------------EKEIKK------RIRRR--EALDAIQ----- : 271
Oryza sativa japonica 1   : GGIEEL--------------EKEIKR------RIKRK--ETLDAIQ----- : 285
Oryza sativa japonica 3   : GGTEEL--------------DKEIKK------RIKRK--EAMDAIR----- : 292
Elaeis oleifera           : GGTEEL--------------QKEINR------RIKLY--NSNTDTSG---- : 325
Zea mays 3                : GGLDEL--------------EKELAR------IGRTI-------------- : 308
Sorghum bicolor 2         : GGLDEL--------------EKELAR------IGRTI-------------- : 312
Oryza sativa japonica 2   : GGLEEL--------------EKELAR------INRSL-------------- : 309
Zea mays 1                : GGLDELVSSPVSEATDTEDAGEEKTRPVVCVVRTSVFMGQSVPNEF----- : 319
Tagetes erecta            : GGTEEL--------------DKEIQR------RIKLY--NNTK-------- : 309
Crocus sativus 1          : GGLKEL--------------EKEVSR------RIKAY--NNSAEIKT---- : 305
Crocus sativus 2          : GGLEEL--------------EKEVSR------LIKAN--H----------- : 291
Narcissus pseudonarcissus : GGEEEL--------------EKLIKR------RIEIN--SRSLDVK----- : 308
Daucus carota 2           : GGNEEL--------------EKEINR------RIKSS--N----------- : 303
Ricinus communis 2        : GGLEAL--------------HTEVQR------RIKAS--C----------- : 300
Oncidium Gower Ramsay     : GGMEAL--------------EREIKR------GVKVF--SSSPNQS----- : 305
Picea sitchensis 1        : GGHDEL--------------EKLFNS------KMKGQGLQKHQ-------- : 339
Picea sitchensis 2        : GGRGEL--------------EKLFNS------KKKGL--QKLQ-------- : 330
Pinus taeda 1             : GGHDEL--------------EKLFNS------KMKGL--QKH--------- : 334
Pinus taeda 2             : GGGEEL--------------DKLMKA------KTNNA--KQGSTQTVKTP- : 319

Muriella zofingiensis     : GAGPEL--------------DRLCAE-------LDS---KSS--------- : 299
Haematococcus pluvialis   : GAAEEV--------------ERLVLE-------LDWS--KR---------- : 293
Chlamydomonas reinhardtii : GGKEEL--------------DKLMAD-------LEA---REAAAAKAAGSS : 297
```

Supplemental Figure 4 (below). Alignment of Polypeptides Encoded by the *A. aestivalis hbfd1* and *hbfd2* cDNAs and Related Polypeptides Encoded by cDNAs or Genes of Other Plants and Green Algae. The polypeptides specified by *hbfd1* and *hbfd2* lack 18-20 C-terminal amino acids that are present in other plant and algal polypeptides. These amino acids are encoded by the C-terminal exon of numerous plant genes (including *A. thaliana*, *Medicago tuncatula*, *Oryza sativa*, *Populus trichocarpa*, *Solanum tuberosum*, *Sorghum bicolor*, *Vitis vinifera*, and *Zea mays*). **Note:** the N terminal sequence of the HBFD2 polypeptide shown here may be incorrect. A comparison of the nucleotide sequence of the cDNA that encodes HBFD2 (GenBank DQ902555) to those of three *A. aestivalis* ESTs (GenBank FL511168, FL510828 and FL510514) indicates that the cDNA should have an A residue inserted at position 57. The N terminal sequence would then read MARVFLGLKPTLSTGSIvkettvgntlvspl, with those amino acid residues shown in uppercase differing from the sequence shown in the alignment below. Alignment of the sequences was achieved using MAFFT version 6 (Katoh and Toh, 2008; the server at http://mafft.cbrc.jp/alignment/server/ was used). The G-INS-i strategy was used with the BLOSUM45 scoring matrix, a gap opening penalty of 1.2 and an offset value of 0.0. The output was manually adjusted to optimize the alignment. Residues are in white text on a black background where all of the sequences have the same amino acid, in white text on a dark grey background where 75% or more are identical, and in black text on a light grey background where 50% or more of the sequences have the same amino acid. The sequences included in this alignment were obtained from the GenBank protein database (all full length amino acid sequences identified using a blastp search on 2/18/2010) or were generated by assembly and translation of overlapping expressed sequence tags (ESTs) identified *via* a tblastn search of the GenBank nucleotide database and/or the database of ESTs. **GenBank Accession Numbers:** *O. sativa japonica*, EEE51627; *P. edulis*, FP093720; *H. vulgare*, BAJ94288; *T. aestivum*, assembly of CD875809, CA736339, CJ710405 and CJ605231; *S. bicolor*, XP_002448957; *Z. mays*, assembly of FL318867, EE185978 and EE185979; *A. formosa*, assembly of DT764577 and DT768502; *P. trichocarpa*, XP_002314088; *R. communis*, XP_002513532; *V. vinifera*, XP_002274681; *N. advena* (partial), EU348740; *A. thaliana* At1g50450, NP_564570; *S. lycopersicum*, AK320320; *H. ciliaris*, assembly of EL425048 and EL422091; *A. annua*, assembly of EY107694 and EY092239; *T. officinale*, assembly of DY819909 and DY827598; *A. aestivalis* HBFD1, ABK41044; *A. aestivalis* HBFD2, ABK41045; *P. sitchensis*, ABR17787; *P.*

patens, assembly of BY985399 and XM_001781386; *C. reinhardtii*, XP_001701437; *O. lucimarinus*, XP_001419501; *O. tauri*, XP_003081093; *Micromonas* sp. RCC299, translation of bases 647766- 649118 of CP001332.

Reference for Supplemental Figure 4:

Katoh, K., and Toh, H. (2008) Recent developments in the MAFFT multiple sequence alignment program. Brief. Bioinform. **9:** 286-298.



```
                                       *        20         *        40         *        60
Oryza sativa japonica       : ---MATVVRACMPL---------------------PP---AAVAS----SS--------- :  20
Phyllostachys edulis        : ---MATMVRACM----------------------------AA-------SS--------- :  13
Hordeum vulgare             : ---M---VRACMPL-------------------RTPP---AAASS----TA--------- :  19
Triticum aestivum           : ---MAAMVRACMPL-------------------RTPP---AAASS----TA--------- :  22
Sorghum bicolor             : ---MGIVVRACAPP---------------------AP---AAAAP----SG--------- :  20
Zea mays                    : MGAMGIVVRACAPL---------------------APAAAAAAAP----SGSS------- :  28
Aquilegia formosa           : ---MA---RISFQL-N----P----------TLFTRV---KESNS----SN--------- :  23
Populus trichocarpa         : ---MI-ARQALFQL-K---SS----------TSVKA-----CVKE----CKHD------- :  26
Ricinus communis            : ---MA---RVLFQL-K----S----------TSVMA-----CAKL--------------- :  19
Vitis vinifera              : ---MA---GASLHL-N----S----------TTVLA-----SLQE--------------- :  19
Nuphar advena               : ------------------------------------------------------------ :   -
Glycine max                 : ---MA---PLSLPL-NLKWTP----------LSVKA-----TATA----TT--------- :  25
Arabidopsis thaliana        : ---MT---RALLLQPY---------------RATVRA---ASSRE----TQ--------- :  23
Solanum lycopersicum        : ---MA---SLGMYS-W----T----------TGIKKR---VFVRA----ANLA------- :  25
Helianthus ciliaris         : ---MA---CASMRH-W----C----------SIRVP-----TVAS--------------- :  19
Artemisia annua             : ---MA---CTCMRH-W----SCN--------TIRSP-----TVAS--------------- :  21
Taraxacum officinale        : ---MA---CACMSH-W----S----------TITTP-----MVAS--------------- :  19
Adonis aestivalis HBFD1     : ---MA---PVLLGL-K----P----------TLSTGS-----V------VKET------- :  21
Adonis aestivalis HBFD2     : ---------TLHGACL----P----------WIETNS---LHWKL----VKET------- :  23
Picea sitchensis            : ---MH---QFGFIMSAPILWH----------SPIGRR---NRNRNGIIWCMSV------- :  34
Physcomitrella patens       : ---MA---TTTQLARWWLAFEPCNIELNKWNVNSTQR---RNQRF----AQRT--GTRLV :  45

Chlamydomonas reinhardtii   : ---MR---RVANTSRATGA------------RCQGAK---LVARP----CARR------- :  28
Ostreococcus lucimarinus    : ---MT---TTRGRGRARAANA----------TTTTTT---RHRRR----HRAV-RARASE :  36
Ostreococcus tauri          : ---MG---STRASGGVEGPRS----------RS----------RR----HRSVARAGATA :  30
Micromonas sp. RCC299       : ---MA---SPVARAPSLGA------------RTLGSR---RVDRS----SRLR-RRPVDG :  34


                                       *        80         *       100         *       120
Oryza sativa japonica       : --AAPSTDAQAQRRS-SSSSARVLVLGGTGRVGGSTATALSK-LRPDLNILIAGRNLEKG :  76
Phyllostachys edulis        : --AAPLTAAAAQR---RSSSARVLVLGGTGRVGGSTATALSN-LRPDLNILIGGRNREKG :  67
Hordeum vulgare             : --TVPA--APIK----PRSSARVLVLGGTGRVGGSTATALSK-LSPDLNILIGGRNREKG :  70
Triticum aestivum           : --TAPA--APSK----PRSSARVLVLGGTGRVGGSTATALSK-LRPDLNILIGGRNREKG :  73
Sorghum bicolor             : --SREA--AQAQRRSRQSRTARVLVLGGTGRVGGSTATAISK-LRPDLGILVGGRNREKG :  75
Zea mays                    : --SRDA--AQAQRWSKPSRTGRVLVLGGTGRVGGSTATALSK-LRPDLGILVGGRNREKG :  83
Aquilegia formosa           : --ESQLLNLQNKT---QIKNSRVLILGGTGRVGGSTALALSK-LSPDLHLLVGGRNREKG :  77
Populus trichocarpa         : --NSNSNVVQLPE---KTRNSSVLVLGGTGRVGGSTATALSK-FCPDLRIVIGGRNREKG :  80
Ricinus communis            : ----DTNRVELPE---KTRNSRVLVLGGTGRVGGSTAIALSK-LCPDLRIVIAGRNREKG :  71
Vitis vinifera              : ----NQPRVHLPE---KTANRRVLVLGGTGRVGGSTAIALSK-LCPDLRITVGGRNREKG :  71
Nuphar advena               : -------R-RGVE---KS---RVLILGGTGRVGGSTAVALSK-LCPDLRIAVAGRNRERG :  45
Glycine max                 : --TSKVPEVPLPE---KIRNSRILVLGGTGRVGGSTAIALSN-LCPDLQILVAGRNREKG :  79
Arabidopsis thaliana        : --YDGVPEVKFSD---PSRNYRVLVLGGTGRVGGSTATALSK-LCPELKIVVGGRNREKG :  77
Solanum lycopersicum        : --KENEKRVEVAE---EVRNSKAVILGGTGRVGGSTAIALSK-LCPDLNIVIAGRNREKG :  79
Helianthus ciliaris         : --------ASLQP---ETANSRVLVLRGTGRVGGSTAIALSK-LSPDLRIVIAGRNREKG :  67
Artemisia annua             : --------A-------GNVNSKVLVLGGTGRVGGSTALALSK-LKPDLRIVIAGRNREKG :  65
Taraxacum officinale        : --------VIPPK---ESVHSRVLILGGTGRVGGSTAVALSN-LSPELHITIAGRNREKG :  67
Adonis aestivalis HBFD1     : --NVGSTLASPLN---KTQNSRVLVLGGTGKVGGSTALALSK-FSPDLRLVIGGRNREKG :  75
Adonis aestivalis HBFD2     : --TVGNTLVSPLN---KTQNSRVLVLGGTGKVGGSTAFALSK-FSPDLRLVIGGRNREKG :  77
Picea sitchensis            : --SSSSERVE-SI---KSDNGRVLILGATGRVGGSAARALVK-SCPNVHLVLAGRNREKG :  87
Physcomitrella patens       : CHAKGSE-AGEEA---GGQKKKVLVMGGTGRVGASTLRALAK--GGDLHLIVGGRNREKG :  99

Chlamydomonas reinhardtii   : ---AAVHVICATG---PVPDKSVVVIGGTGRVGSSTAATLLK-EFPNLKVTVASRSDDSF :  81
Ostreococcus lucimarinus    : DANEASERTTTTT---TTTTTKVVVLGGTGRVGSATAAALVRGANGGVEVTLGGRSRERD :  93
Ostreococcus tauri          : STGDAT-RAEGER---GGVDARVVVLGGTGRVGSATAAALAR-AGEGTTVTLCGRSSDGA :  85
Micromonas sp. RCC299       : HAPSSSRRRSLAV---AAEGKKVVFVGGTGRVGSSAAAALLR-QEPGVALVLAGRTESSF :  90
```

```
                                     *       140         *       160         *       180
Oryza sativa japonica      : ESLASKLGD----ESEFVQVDIRDRNMLEEVLQ--DVDLVVHAAGPFQ-RENECTVLQAA : 129
Phyllostachys edulis       : RSLASKLGE----QSEFLEVDIRNANMLEEALQ--GVDLVVHAAGPFQ-REDKCTVLQAA : 120
Hordeum vulgare            : ESLASKLGE----QSEFVEIDTGNAAMLEKALE--DVDLVVHTAGPFQ-REAECTVLRAA : 123
Triticum aestivum          : ESLASELGE----QSEFVKIDTGNAAMLEKALE--DVDLVVHTAGPFQ-REAECTVLRAA : 126
Sorghum bicolor            : ESVAAKLGS----QSEFVQVDTRDAGMLEEALQ--GVDLVVHTAGPFQ-RAEECTVLQAA : 128
Zea mays                   : ESIAAKLGG----QSEFVQVDTRNAGMLEEALQ--GVDLVVHTAGPFQ-RAEECTVLQAA : 136
Aquilegia formosa          : DALVSKLGE----NSEFVEVNIENVKSLEAALN--DVDLVVHTAGPFQ-QAENCTVLEAA : 130
Populus trichocarpa        : AAMVGQLGR----NSEFTEVNIENVDSLGAALK--DVDLVVHAAGPFQ-QAAKCTVLEAA : 133
Ricinus communis           : AALVDKLGK----NSDFAQVDINNVESLEAALS--DVDVVVHAAGPFQ-QTEKCSVLEAA : 124
Vitis vinifera             : AAMLAKLGE----NSEFAEVNIDNVKSLEAALN--DVDLVIHTAGPFQ-QAEKCTVLEAA : 124
Nuphar advena              : AALVSKLGE----NSEFVEVNINNIDALVEALK--DVDLVVHAAGPFQ-QGGKCNVLEAA :  98
Glycine max                : EVLTAKLGG----NSEFARVDIDDVNSLETALK--NVDLVVHAAGPFQ-QAERCSVLEAA : 132
Arabidopsis thaliana       : EAMVAKLGE----NSEFSQVDINDAKMLETSLR--DVDLVVHAAGPFQ-QAPRCTVLEAA : 130
Solanum lycopersicum       : AAMVSKLGK----NAEFAEVNIDDREALEANLT--DADIVVHAAGPFQ-QSENCKVLEAA : 132
Helianthus ciliaris        : ASMAATLEN----NAEFAQVDINDAKSLESALT--DVDLVVHAAGPFQ-QTENCSLLEAA : 120
Artemisia annua            : ANMVTTLGN----NAEFAEVDINDDKSLDSALT--DVDLVVHAAGPFQ-QTENCRVLEAA : 118
Taraxacum officinale       : ARMVATLGN----NVEFCEFDINDYNLLESALT--DVDLVVHAAGPFQ-QTHNCNVLEAA : 120
Adonis aestivalis HBFD1    : DAVVSKLGE----NSEFVEVNVDSVRSLESALE--DVDLVVHAAGPFQ-QAEKCTVLEAA : 128
Adonis aestivalis HBFD2    : DAVVSKLGE----NSEFVEVNVDSMRSLESAFK--DVDLVVHAAGPFQ-QAEKCTVLEAA : 130
Picea sitchensis           : AKLAAELGA----NTEFYQVSIDNTKALAAALD--GVDLVVHAAGPFQ-REEKCTVLEAA : 140
Physcomitrella patens      : EALARELGG----SVEFSAFNLEDASAVRAAID--GVDLVVHAAGPFQ-RRVECAVLEAA : 152

Chlamydomonas reinhardtii  : KAAVERRPEL--SKAGFQRVDITNADSVQALLKSTGADLVIHTAGPFQ-RSKNYAVLEAA : 138
Ostreococcus lucimarinus   : AEAKARHRGL--ANASFVEVDVCDKASVTRAIQ--GADLVINTAGPFQ-RRKSCAALEAA : 148
Ostreococcus tauri         : SEARSRHPSL--ANASFVEVDVCDKASVRRAIE--GADLVINTAGPFQ-RRTSVAALEAA : 140
Micromonas sp. RCC299      : TAAVERHPTLRSSSCAFAACDCSDPASLEAVIA--GADLVVHSAGPFQGGGDQCAVLDAA : 148


                                     *       200         *       220         *       240
Oryza sativa japonica      : IATKTAYIDVCDDTDYSWRAK-GFHEQAKDCGIPAITTAGIYPGVSNVMAAELVHAARS- : 187
Phyllostachys edulis       : ISTKTAYIDVCDDLDYSWRAK-GFHEQAKACGVPAITTAGIYPGVSNVMAAELVHAARS- : 178
Hordeum vulgare            : ISTKTAYIDVCDDMDYSWRAK-AFHEEAKAQGVPAITTAGIYPGVSNVMAAELVDAARS- : 181
Triticum aestivum          : ISTKTAYIDVCDDMDYSWRAK-AFHEEAKAQGVPAITTAGIYPGVSNVMAAELVNAARS- : 184
Sorghum bicolor            : ISTKTPYIDVCDDTDYSWRAK-GFHEQAKAAGVPAITTAGIYPGVSNVMAAELVHAARS- : 186
Zea mays                   : ISTKTPYIDVCDDTDYSWRAK-GFHEQAKAAGVPAITTAGIYPGVSNVMAAELVHAARS- : 194
Aquilegia formosa          : ISTKTAYMDVCDDTSYSLQAK-SLHEKAVAANIPAITTAGIYPGVSNVMAAELVRSARD- : 188
Populus trichocarpa        : IETKTAYVDVCDDTSYALRAK-SFKDKALAANIPAITTGGIYPGVSNVMAAELVRAAKT- : 191
Ricinus communis           : IATKTAYIDVCDDTSYALRAK-SFKDRALAANIPAITTAGIYPGVSNIMAAELVRAARM- : 182
Vitis vinifera             : IETKTAYVDVCDDTTYAWRAK-SLLEKALSANVPAITTGGIYPGVSNVMAAELVRVARS- : 182
Nuphar advena              : ISTKTAYLDVCDDTTYAYRAK-SVHQKAVDANVSAITTGGIYPGVSNVMAAELVRLARN- : 156
Glycine max                : INTQTAYLDVCDDTSYAWRAK-SFMNRALDANVPAITTGGIYPGISNVMAAELVRAA-N- : 189
Arabidopsis thaliana       : IKTKTAYLDVCDDTSYAFRAK-SLEAEAIAANIPALTTAGIYPGVSNVMAAEMVAAARS- : 188
Solanum lycopersicum       : IGTKTAYLDICDDTSYATRAK-SYMNMALEANIPAITTAGIYPGVSNVMAAELVRTAKL- : 190
Helianthus ciliaris        : IRAKTPYLDVCDDTSYALRAK-SFMNEALAAKVPAITTGGIYPGVSNLMAAELVRVAKS- : 178
Artemisia annua            : IRAKTAYLDVCDDTSYALRAK-SFKDEALAAKVPAITTGGIYPGVSNVMAAELVRVAKE- : 176
Taraxacum officinale       : IRTKTAYLDVCDDTSYALRAK-TYKNEAIAEKVPAITTGGMYPGISNVMAAELVRIAKS- : 178
Adonis aestivalis HBFD1    : ISTRTAYVDVCDNTSYSMQAK-SFHDKAVAANVPAITTAGIFPGVSNVIAAELVRSARD- : 186
Adonis aestivalis HBFD2    : ISTRTAYVDVCDNTSYSMQAK-SFHDKAVAANVPAITTAGIFPGVSNVIAAELVRSARD- : 188
Picea sitchensis           : ISSKTAYIDVCDDRSYSCLAK-TFHEKAVAAQVPAITTAGIYPGVSNLMAAELVRLARTS : 199
Physcomitrella patens      : IDTKTAYIDVCDDQDYSMRAK-AYHDRAVAAGIPAITTGGIYPGISNIMAAELVRL---- : 207

Chlamydomonas reinhardtii  : IASGTGYIDVCDDTPFAEGAKAAYMEKAKAAGVPAIVSGGIYPGTSNVMAAHIISIARA- : 197
Ostreococcus lucimarinus   : LESGVKYLDVCDDASYGAEAK-KLSEKAKAAGVAAITCAGIYPGVSNLMARDIVESMKA- : 206
Ostreococcus tauri         : IECGVKYLDVCDDASYGAEAK-KLSENAKAAGVAAITCAGIYPGVSNLMVRDIVESMKA- : 198
Micromonas sp. RCC299      : IATGVPYLDVCDDAEYAKACR-SKSDAAKSKGVPCVTTGGIYPGVSNLMAAEMIAQNRA- : 206
```

```
                                     *       260         *       280         *       300
Oryza sativa japonica     : ---E-------------NAGEPERLRFFYYTAGTGGAGPTILTTSFLLLAEDVIAYNKGE : 231
Phyllostachys edulis      : ---E-------------NAGEPERLRFFYYTAGSGGAGPTILTTSFLLLGEDVIAYNKGE : 222
Hordeum vulgare           : ---E-------------D-GEPERLRFFYYTAGSGGAGPTILATSFLLLGEDVIAYNKGE : 224
Triticum aestivum         : ---E-------------D-GEPERLRFFYYTAGSGGAGPTILATSFLLLGEDVIAYNKGE : 227
Sorghum bicolor           : ---E-------------N-GEPERLRFFYYTAGTGGAGPTILATSFLLLGEDVIAYNKGT : 229
Zea mays                  : ---E-------------N-GEPERLRFFYYTAGTGGAGPTILTTSFLLLGEDVIAYNKGE : 237
Aquilegia formosa         : ---E-------------NMGEPQRLRFSYYTAGTGGAGPTILATSFLLLGEEVVAYNKGE : 232
Populus trichocarpa       : ---E-------------SKGKPERLRFYYYTAGSGGAGPTILATSFLLLGEEVVAYNKGE : 235
Ricinus communis          : ---E-------------SKGNPERLRFHYYTAGTGGAGPTILATSFLLLGEEVVAYNKGE : 226
Vitis vinifera            : ---E-------------SQGKPERLRFYYYTAGTGGAGPTILATNFLLLGEEVVAYNKGE : 226
Nuphar advena             : ---E-------------GTYEPERLRFYYYTAGSGGAGPTILATSFLLLGEDVVAYNKGQ : 200
Glycine max               : ---E-------------SEDKPERLRFYYYTAGTGGAGPTILATSFLLLGEEVVAYNKGE : 233
Arabidopsis thaliana      : ---E-------------DKGKPEKLRFSYYTAGTGGAGPTILATSFLLLGEEVTAYKQGE : 232
Solanum lycopersicum      : ---E-------------SEGELERLRFYYYTAGTGGAGPTILATSFLLLGEDVIAYNKGE : 234
Helianthus ciliaris       : ---E-------------NKGEPERLRFYYYTAGTGGAGPTILATSFLLLGEEVTTYNKGE : 222
Artemisia annua           : ---E-------------NKGEPERLRFYYYTAGTGGAGPTILATSFLLLGEEVIAYNRGE : 220
Taraxacum officinale      : ---E-------------NKGEPERLRFYYYTAGTGGAGPAILASSILVLGEDAIAYSRGK : 222
Adonis aestivalis HBFD1   : ---E-------------NT-EPQRLRFSYFTAGSGGAGPTSLVTSFLLLGEEVVAYSEGE : 229
Adonis aestivalis HBFD2   : ---E-------------NT-EPQRLRFSYFTAGSGGAGPTSLVTSFLLLGEEVVAYSEGE : 231
Picea sitchensis          : SSKE-------------TPSEPERLRFSYFTAGSGGAGPTILATSFLLLREEVIAYRKGE : 246
Physcomitrella patens     : -----------------N-GKAKRIRYSYYTAGSGGAGPTILATSFLLLGEEAIVYVDGK : 249

Chlamydomonas reinhardtii : ---EYDDNWNYRTPAPGESVEPKWLRYSYYTAGSGGAGPTILETSFLLAGEDVIVYKDNK : 254
Ostreococcus lucimarinus  : ---EFRATEENE----GKEPEVEYVLYNYFTAGSGGVGTTILATSYLLCGEDVVCWEDGQ : 259
Ostreococcus tauri        : ---EFREQEGNE----NTDPEVEYVLYNYFTAGSGGVGTTILATSYLLCGEEVVCWENDK : 251
Micromonas sp. RCC299     : -----SGSEG------AEVAEVEYVLYNYFCAGSGGVGDTILATSYMLCGEDVVCWDKGE : 255


                                     *       320         *       340         *       360
Oryza sativa japonica     : EI----------KLKPYSGALSIDFGKGVRKKDVYLLNLPEVKSAYKVLGVPTVSARFGT : 281
Phyllostachys edulis      : EI----------KLKPYSGALNIDFGKGVRKKDVYLLNLPEVKSAYKILGVPTVSARFGT : 272
Hordeum vulgare           : EI----------KLKPYSGVLNIDFGKGVRKRDVYLLNLPEVKSAHKFLGVPTVSARFGT : 274
Triticum aestivum         : EI----------KLKPYSGVLNIDFGKGVRKRDVYLLNLPEVKSAHKFLGVPTVSARFGT : 277
Sorghum bicolor           : SSLCTSAIAPMEFHTPYSGALNIDFGKGVRKKNVYLLNLPEVKSAFKILGVPTVSARFGT : 289
Zea mays                  : EI----------KLKPYSGALNIDFGKGVRKKNVYLLNLPEVKSAFKILGVPTVSARFGT : 287
Aquilegia formosa         : KI----------KLKPYTGMLNVDFGKGIGKRDIYLLNLPEVRSAHEILGIPTVSARFGT : 282
Populus trichocarpa       : KI----------KLKPYSGMLNIDFGKGIGKRDVYLLNLPEVRSAHDVLGIPTVSARFGT : 285
Ricinus communis          : RI----------KLKPYSGMLNIDFGKGIRKRDVYLLNLPEVQSAHEILGIPTVSARFGT : 276
Vitis vinifera            : KI----------KLKPYSGMLNIDFGKGIGKRDVYLLHLPEVRSAHEILGVPTVSARFGT : 276
Nuphar advena             : KI----------KLRPYSGMLNIDFGKGIGKRDVYLLNLPEVRSTHEVLGVPTVSARFGT : 250
Glycine max               : KI----------RMRPYSGMLNVDFGKGIGKRDVYLLNLPEVSSAHEILGVPSVSARFGT : 283
Arabidopsis thaliana      : KV----------KLRPYSGMITVDFGKGIRKRDVYLLNLPEVRSTHEVLGVPTVVARFGT : 282
Solanum lycopersicum      : KI----------KLTPYSGMLSIDFGMGIGKKDVYLLNLPEVKSVHEVLGVPTVSARFGT : 284
Helianthus ciliaris       : KI----------KLRPYSGMLSIDFGKGIGKKDVYLLNLPEVTSAHEVLGVPTVSARFGT : 272
Artemisia annua           : KI----------KLRPYSSMVNIDFGRGIGKKDVYLLNLPEVTSTHEVLGVPTVSARFGT : 270
Taraxacum officinale      : KI----------KLRPYSGKVNIDFGKGIKKKNVYLLNLPEVTSTHETLGIPSVSARFGT : 272
Adonis aestivalis HBFD1   : KV----------ELKPYTGKLNIDFGKGVGKRDVYLWNLPEVRSGHEILGVPTVSARFGT : 279
Adonis aestivalis HBFD2   : KV----------ELKPYTGKLNIDFGKGVGKRDVYLWNLPEVRSGHEILGVPTVSARFGT : 281
Picea sitchensis          : EF----------RAKPYSGVLDIDFGKGVGKRSVYLLNLPEVGSTHEVLGVPTVSARFGT : 296
Physcomitrella patens     : MQ----------KMKAYSARRDVDFGRGIGKKPVYLLNLPEVRSTHEVLKVPSVSARFGT : 299

Chlamydomonas reinhardtii : EV----------VLPPISNRREVDFGPGVGRKGVYLYNLPEVVSGHKYMRVPDVSARFGT : 304
Ostreococcus lucimarinus  : RI----------VEKPASQRKVVDFGKGVGRREVFLYNLPEVASTREIFGARTVKARFGT : 309
Ostreococcus tauri        : RV----------VEKPASQRKVVDFGQGVGKREVFLYNLPEVASTREYLGASTVKARFGT : 301
Micromonas sp. RCC299     : RV----------VTRPATQRKVVDFGKKCGKREVFLYNLPECASARETWNAETVKARFGT : 305
```

```
                                     *       380         *       400         *       420
Oryza sativa japonica     : APFFWNWVMQAFANFLPVEFLRDKNKVLKLVGFVDPFVRAIDGIAGERVSMRVDLDCSNG : 341
Phyllostachys edulis      : APFFWNWGMQAFAKFLPVEFLRDKNKVLKLVELVDPFVRAIDGIAGERVSMRVDLECSNG : 332
Hordeum vulgare           : APFFWNWGMEAFANFLPVELLRDKDKVRKLVEKIDPLVRAIDGIVGERVSMRVDLECSNG : 334
Triticum aestivum         : APFFWNWGMEAFANFLPVELLRDKDKVGKLVEQIDPLVRAIDGIVGERVSMRVDLECSNG : 337
Sorghum bicolor           : APFFWNWGMQAFANFLPVEFLRDRNKVQKLVQSVDPLVRAVDGIAGERVSMRVDLDCSNG : 349
Zea mays                  : APFFWNWGMQAFANFLPVEFLRDRNKVQKLVQSVDPLVRAVDGIAGEHVSMRVDLDCSNG : 347
Aquilegia formosa         : APFFWNWGMAAMTSLIPSEVLRDRKKVQKLVDLFDPVVRAIDGIAGERVSMRVDLDCSND : 342
Populus trichocarpa       : APFFWNWGMSAMTNLLSPEFLKDRTKVQQLVQLFDPLVRAVDGIAGERVSMRVDLECTDG : 345
Ricinus communis          : SPFFWNWGMEIMTNLFPPEVLRDRSKVQQLVQLFDPLVRALDGIAGERVSMRIDLECSDG : 336
Vitis vinifera            : APFFWNWGMEAMTNLLPVEFLRDRSKVQELVQLFDPIVRAMDGIAGERVSMRVDLECSDG : 336
Nuphar advena             : APFFWNWGMLAMANLVPAEILRDRGKVQQLVTLFDPLVRAFDGIAGERVSMRVDLECSNG : 310
Glycine max               : APFFWNWGMEAMTKLLPSEFLRDRNKVQSLVQLFDPVVRAVDGIAGERVSMRVDLECASG : 343
Arabidopsis thaliana      : APFFWNWGMEIMTKLLPSEVLRDRTKVQQMVELFDPVVRAMDGFAGERVSMRVDLECSDG : 342
Solanum lycopersicum      : DPFFWNWGMIAMRNLLPQEFLRDRSKVQQLVQLFDPVVRAVDGTAGEAVSMRVDLECSDG : 344
Helianthus ciliaris       : APFFWNWAMDAMTRFLPSEVLRDRSKVQEMVRLFDPVVRIIDGYAGERVSLRVDLECSGG : 332
Artemisia annua           : APFFWNWAMDAMTRFVPSEVLRDRSKVQEMVRLFDPVVRVIDGYAGERVSIRVDVECSGG : 330
Taraxacum officinale      : SPFFWNCAMDAMTRVVPSEHLRDRNKVQEMVRLLDPVVRSFDGIVGEHVALRVDLECSGG : 332
Adonis aestivalis HBFD1   : APFFWNWAMVAMTTLLPPGILRDRNKIGMLANFVYPSVQIFDGIAGECLAMRVDLECANG : 339
Adonis aestivalis HBFD2   : APFFWNWAMVAMTSLLPPGILRDRNIIEKLANFVYPSVQVFDGIAGECLAMRVDLECANG : 341
Picea sitchensis          : APFFWNWAMSAVVNLAPVEFLKDRNKVQTLVQLSDPLVRAIDIFSGELMSMRVDLECMDG : 356
Physcomitrella patens     : YPQIWNIAMGLVASLVPKDILQDQQKVQGLVQLSDIAVRAVDGFAGEKVSMRVDYEGENG : 359

Chlamydomonas reinhardtii : DPFIWNWAMWLTARLVPRSLLNDRNFVKGFAKLSDPFVRNVDKIIGEAVAMRVEVDMVGG : 364
Ostreococcus lucimarinus  : SPGIWNGAMVAIANLVPKSLLENQDAMKGLANFSAPIVRSVDAIVGETTSIRVDVKLKDG : 369
Ostreococcus tauri        : SPGLWNGAMVAIANVVPKALLENQDAMKGLAGFTAPIVRAVDAIVGEKTSIRVEVKLKGG : 361
Micromonas sp. RCC299     : SPGLWNGAMALMANLVPKETLLNKDVAKALATFTAPIVRAVDAVVGERTAMRIDVKLKGG : 365


                                     *       440         *       460         *       480
Oryza sativa japonica     : KNTIGLFSHRKLSVSVGYATAAFVLAVLEGSTQPGVWFPEEPEGVAIESRKVLLERASQG : 401
Phyllostachys edulis      : KNTIGLFTHRKLSVSVGHAIAAFVFAVLEGNTQPGVWFPEEPEGVAIEARKLLLERASQG : 392
Hordeum vulgare           : RNTIGLFSHRKLSVSVGHSTAAFVQAVLEGSTQPGVWFPEEPEGIAIESRKLLLERASQG : 394
Triticum aestivum         : RNTIGLFSHRKLSVSVGHSTAAFVQAVLEGSTQPGVWFPEEPEGIAIESRKLLLERASQG : 397
Sorghum bicolor           : RNTIGLFTHKKLSVSVGYATAAFALAVLEGNTQPGVWFPEEPEGIPMEARKLLLERASQG : 409
Zea mays                  : RNTIGLFTHKKLSVSVGFATAAFALAVLEGNTQPGVWFPEEPEGIRIEARKLLFERASQG : 407
Aquilegia formosa         : RSTVGIFSHKRLSVSVGTSTAAFALAVLEGSTQPGVWFPEEPEGIAVDDRELLLQRASQG : 402
Populus trichocarpa       : RNTLGLFSHRKLSVSVGNATAAFALAVLEGSTQPGVWFPEEPEGIAIEARELLLNRATEG : 405
Ricinus communis          : RNTVGIFSHKRLSVSVGNATAAFVLAILEGSTKPGVWFPEEPEGIATEAREVLLQRAAEG : 396
Vitis vinifera            : RNTVGLFSHRRLSVSVGFATAAFALAVLEGSTQPGVWFPEEPEGIAIDARDILLKRAAQG : 396
Nuphar advena             : RSTVGIYSHKRLSESVGTATAAFVMAILEGSTQPGVWFPEEPEGISVEARKALLERASKG : 370
Glycine max               : RNTVGIFSHRRLSVSVGIATAAFALAILEGSTRPGVWFPEEAQGIPIEAREVLLKRASQG : 403
Arabidopsis thaliana      : RTTVGLFSHKKLSVSVGVSTAAFVAAMLEGSTQPGVWFPEEPQGIAVEAREVLLKRASQG : 402
Solanum lycopersicum      : HNRIGIFSHKRLSRSVGISTAAFVLAMLEGSTKPGVWFPEEPEGIAVEAREVLLQRASEG : 404
Helianthus ciliaris       : RQTIGVFSHRSLSISVGYSVAAFALAVLEGNTQPGVWFPEQPEGIAVEARGLLLERATQG : 392
Artemisia annua           : RQTIGVFTHKRLSVSVGHSTAAFALAILEGSTQPGVWFPEEPQGIAIEARELLLERAAQG : 390
Taraxacum officinale      : QHTVSIFSHKKYSVSVGTSIAAFALAVLEGSTQPGVWFPEEPQGIAIEARELLLKRASQG : 392
Adonis aestivalis HBFD1   : RNTFGILSHERLSVLVGTSTAVFAMAILEGSTQPGVWFPEEPGGIAISDRELLLQRASQG : 399
Adonis aestivalis HBFD2   : RNTSAILSHERLSELVGTSTAVFALAILEGSTQAGVWFPEEPEGIAVGDRELLLKRASQG : 401
Picea sitchensis          : RKAVGLYSHKKLSICVGVAISAFVRAVLEGNTQPGVWFPEEPEGIAVEARQQLLERAAEG : 416
Physcomitrella patens     : KKAIGIFSHKMLSVSVGMSVAAFVRALLEGATQPGVWFPEEEGGIAESARPKLLERAAEG : 419

Chlamydomonas reinhardtii : KNSSGIFVHKYLSQSMGYSTAAFAQSVLQGKTQPGVWYPEEKE--ALQDRRQFLQFAATG : 422
Ostreococcus lucimarinus  : KQSVGLYTHPRLSECVGTCTASFATAMLNGECAPGVWYPEEVE--AISDRDALFERAKEG : 427
Ostreococcus tauri        : KQAVGLYTHPRLSECVGTCTAAFALAMLEGGCAPGVWYPEENE--AIADRGALFERAKEG : 419
Micromonas sp. RCC299     : ALAGAIYNHPRLSEAVGDATAAFCAAMLRGETDPGVWYPEEAG--AVKDRARLLEAAAVG : 423
```

```
                                     *       500         *
Oryza sativa japonica     : TTIFVMNKPSWMIETDP--KEVGLGIYV-- : 427
Phyllostachys edulis      : TTNFVMNKPSWLVETDP--KEVGLGIYV-- : 418
Hordeum vulgare           : TTNFVMNKPSWMIETDP--KEVILGIYV-- : 420
Triticum aestivum         : TTNFVMNKPSWMIETDP--KEVILGIYV-- : 423
Sorghum bicolor           : TSTFVMNKPSWMVETDP--KEVGLGIYV-- : 435
Zea mays                  : TSTFVMNKPSWMVETDP--KEVGLGIFV-- : 433
Aquilegia formosa         : TINFVMNKAPWMVETEP--KELVFGIFI-- : 428
Populus trichocarpa       : TINFIMNKPPWMVETDP--KELGLGIYV-- : 431
Ricinus communis          : TINFVMNKPPWMVETEP--KEVGLGIYV-- : 422
Vitis vinifera            : TINFIMNKPPWMVETDP--KELGLGIYV-- : 422
Nuphar advena             : TINFVMNKPAWMVETEP--KELGLGIYV-- : 396
Glycine max               : TFNFIMNRSPWMVETNP--KEFGLGIYL-- : 429
Arabidopsis thaliana      : TFNFILNKPPWMVETEP--KEVVLGIYV-- : 428
Solanum lycopersicum      : TINFILNKAPWMVETNP--KELGFGIYS-- : 430
Helianthus ciliaris       : TINFVMHKAPWMMETNP--KLVGLGIYV-- : 418
Artemisia annua           : TINFAMHKAPWMVETNP--KEVGLGIYV-- : 416
Taraxacum officinale      : TINFVMHKAPSMVETYPNFKEVGFGI---- : 418
Adonis aestivalis HBFD1   : AINFIMKQ---------------------- : 407
Adonis aestivalis HBFD2   : AINFIMKQ---------------------- : 409
Picea sitchensis          : TINFVMNRPPWMVETNP--KEIGFGIYMQ- : 443
Physcomitrella patens     : TLNFVMNKSPWMTDKDP--KEIGFGLYWE- : 446

Chlamydomonas reinhardtii : CSRFELNRSAWALESEI--KQIGGMIYW-- : 448
Ostreococcus lucimarinus  : TSLFALNQAPWMVESKP--VNLGFGLYWT- : 454
Ostreococcus tauri        : TSLFLLNQAPWMVESKA--INIGMGLYWT- : 446
Micromonas sp. RCC299     : CDNYEMNKAAWMLESKP--INLGFGIYIEL : 451
```

Supplemental Figure 5 (below). Alignment of Selected Members of the Extended Saccharopine Dehydrogenase Family of Enzymes. The polypeptides included in this alignment were chosen so that each clade of the maximum likelihood tree that was produced from this alignment (Figure 10 of the manuscript) included ten examples from a broad diversity of taxa. The alignment was achieved using MAFFT version 6 (Katoh and Toh, 2008; http://mafft.cbrc.jp/alignment/server/). The G-INS-i strategy was used with the BLOSUM30 scoring matrix, a gap opening penalty of 1.9 and an offset value of 0.1. The output was manually adjusted to optimize the alignment. Residues are in white text on a black background where all 40 of the sequences have the same or a similar amino acid (Blosum35 matrix), and in white text on a dark grey background where at least 50% of the sequences have the same or a similar amino acid. **Note:** A number of the polypeptides in this alignment are hypothetical ones, predicted by genome assembly programs. These gene assemblies have not necessarily been confirmed by sequencing of the corresponding cDNAs and some of them may not be entirely correct. **GenBank Accession Numbers: HBFD and related polypeptides (first 10 sequences):** *Adonis aestivalis* HBFD1, ABK41045; *Arabidopsis thaliana* At1g50450, NP_564570; *Picea sitchensis*, ABR17787; *Selaginella moellendorffii*, XP_002960283; *Chlamydomonas reinhardtii*, XP_001701437; *Ectocarpus siliculosus*, CBN76749; *Phaeodactylum tricornutum* CCAP 1055/1, XP_002184209; *Nostoc punctiforme* PCC 73102, YP_001866432; *Synechococcus* sp. PCC 7002, YP_001734542; *Cyanothece* sp. ATCC 51142, YP_001805691. **Polypeptides of unknown function group 1 (sequences 11-20):** *Chlamydomonas reinhardtii*, XP_001690673; *Chlorella variabilis*, EFN50859; *Salpingoeca* sp. ATCC 50818, EGD79814; *Thalassiosira pseudonana* CCMP 1335, XP_002291038; *Phaeodactylum tricornutum* CCAP 1055/1, XP_002177453; *Synechocystis* sp PCC 6803, NP_441234; *Ralstonia pickettii* 12D, YP_002980505; *Phytophthora infestans* T30-4, XP_002900027; *Ectocarpus siliculosus*, CBN75169; *Aspergillus niger* CBS 513.88, XP_001402256. **Polypeptides of unknown function group 2 (sequences 21-30):** *Arabidopsis thaliana* At5g39410, NP_568564; *Picea sitchensis*, ABR17362; *Selaginella moellendorffii*, XP002962105; *Chlorella variabilis*, EFN54908; *Ectocarpus siliculosus*, CBJ33497; *Thalassiosira pseudonana* CCMP 1335, XP_002290994; *Salpingoeca* sp. ATCC 50818, EGD83292; *Homo sapiens*; NP_057086; *Caenorhabditis elegans*, NP_503576; *Drosophila melanogaster*, NP_649517. **Sacharopine dehydrogenases (SDH; sequences 31-40):** *Arabidopsis thaliana* At4g33150, AEE86184; *Selaginella moellendorffii*, XP_002973511 (amino acids 527-end); *Phaeodactylum tricornutum* CCAP 1055/1, XP_002179334 (amino acids 270-end); *Thalassiosira pseudonana* CCMP1335, XP_002286419 (amino acids 566-end); *Phytophthora infestans* T30-4, XP_002906594 (amino acids 576-end); *Drosophila melanogaster*, AAR82744 (amino acids 520-end); *Homo sapiens*, AAF03526 (amino acids 477-end); *Saccharomyces cerevisiae* S288c, NP_014448; *Aspergillus niger* CBS 513.88, XP_001401890; *Salpingoeca* sp. ATCC 50818, EGD80106.

Reference for Supplemental Figure 5:

Katoh, K., and Toh, H. (2008) Recent developments in the MAFFT multiple sequence alignment program. Brief. Bioinform. **9:** 286-298.

```
                                    *        20         *        40         *        60
A. aestivalis HBFD1      : -------------------------------MAPVLLGLKPTLSTGSVVKETNVGSTLAS :  29
A. thaliana At1g50450    : -----------------------------MTRALLLQPYRATVRAASSRETQYDGVPEVK :  31
P. sitchensis            : -------------------MHQFGFIMSAPILWHSPIGRRNRNRNGIIWCMSVSSSSERV :  41
S. moellendorffii        : ----------------------------MILSSSGLPPPWQQPAISCRHCSIVCGASSSL :  32
C. reinhardtii           : -------------------------MRRVANTSRATGARCQGAKLVARPCARRAAVHVIC :  35
E. siliculosis           : -----------------MLRSMTTRRAYQPIGVRCLSPNWFRSSRSSGRDTSTHGHRFRC :  43
P. tricornutum           : -----------------MIAASHYRRTNSDAKVRLVILYLLICRVASFSVKNTEQSSSCI :  43

T. pseudonana unknown 1  : MEVVDKKTPITFGPIVFISSLFLGVPFLLTVLLPLTIVSAPIHFFAALLINGKLFSFSKA :  60
P. tricornutum unknown 1 : -----MPSPVTQDGRTVPSVSAATFLATVASLPVWALTVLPLSLFYQAGKVIMKPVLPKA :  55


                                    *        80         *       100         *       120
A. aestivalis HBFD1      : PLNK----------------TQNSRVLVLGGTGKVGGSTALALSKFS------------- :  60
A. thaliana At1g50450    : FSDP----------------SRNYRVLVLGGTGRVGGSTATALSKLC------------- :  62
P. sitchensis            : ESIK----------------SDNGRVLILGATGRVGGSAARALVKSC------------- :  72
S. moellendorffii        : AGPG----------------PEKKEVVLIGGTGRVGSSTARALARVW------------- :  63
C. reinhardtii           : ATGP----------------VPDKSVVVIGGTGRVGSSTAATLLKEF------------- :  66
E. siliculosis           : RSTPSGLGAKGDDGGGARGYDKHARVIVVGGSGRVGGSTVRALRQLA------------- :  90
P. tricornutum           : SLDN----------------LQNKRVLVVGGSGRVGGSCVTQLVL--------------- :  72
N. punctiforme           : ---------------------MTDRVLILGGRGRIGSSVAQDLANHT------------- :  26
Synechococcus PCC 7002   : ---------------------MVRKILILGGTGCIGQRVAAAIA---------------- :  23
Cyanothece  ATCC 51142   : ---------------------M-QRVLIIGGCGRIGSSIAQDIINHT------------- :  25

C. reinhardtii unknown 1 : --------------MSATAASRPYQVVVWGASGFTGRLVAEHLARDY------------- :  33
C. variabilis unknown 1  : -----------------MAPTRAFDVVIWGATGFTGRLVAEHLARDY------------- :  30
Salpingoeca unknown 1    : ------------------MAGRTYDVVVFGATGFTGSLVVEYLARAY------------- :  29
Synechocystis unknown 1  : -----------------MTKPLPYDLIIVGATGFVGGIVCRYLLSHW------------- :  30
R. pickettii unknown 1   : ------------------MPAPDYDVVVFGATSFVGQILTRYLSEHL------------- :  29
P. infestans unknown 1   : ------------------MS-KRFDVIVYGATGFTGSLVARYLAAES------------E :  29
E. siliculosus unknown 1 : ----------------MSDSSNKYDVVVFGATSFAGQLVCEYYLANY------------- :  31
T. pseudonana unknown 1  : TSNTIVDRSGDQSIIIAPRESRKYDVVLLGSTGFTGKLATAYLAKQY--------GACCF : 112
P. tricornutum unknown 1 : EQPRLDTGYHVSDSDILPRRERKYDMVVLGATGFTGRLAVRHLAKTY------------- : 102
A. niger unknown 1       : -----------------MESNKQYDLIVLGPTGYTGRFCADHIVKNF------------- :  30

A. thaliana At5g39410    : -------------MNPTQKPEPVYDMVILGASGFTGKYVVREALKFL------------- :  34
P. sitchensis unknown 2  : ------------------MEDERLDAVILGASGFTGKYVVRYFLAQL------------- :  29
S. moellendorffii unk. 2 : ------------------MESAAADVIILGASGFTGKRVLGEFLRKL------------- :  29
C. variabilis unknown 2  : -----------------MAASRPHDLVVLGATGFTGNRICKEVLNSG------------- :  30
E. siliculosis unknown 2 : ------------------MASRDMDILVVGASGFTGAHVCKRLARSV------------- :  29
T. pseudonana unknown 2  : -------MTSYPTNRRERFASRSIDVIVHGATGYTGRRVVRHLVTKH------------- :  40
Salpingoeca unknown 2    : -----------MSEAKKTQQPRAVDIVVFGATGFTGQYVVQYLRGTV------------- :  36
H. sapiens unknown 2     : ----------------MATEQRPFHLVVFGASGFTGQFVTEEVAREQ------------- :  31
C. elegans unknown 2     : ------------------ME-ARYDVVIYGATGVTGGKIFETLISCG------------- :  28
D. melanogaster unknown 2: ------------------MSGDRLDVIIFGASGFTGKYTVFEAVTVL------------- :  29

A. thaliana At4g33150 SDH: ------------------MT-KKSGVLILGA-GRVCRPAADFLASVRTISSQQWYKTYFG :  40
S. moellendorffii SDH    : ------------------VVSEYTNILILGA-GRMCEPTLMYLTENA--------FEDYA :  33
P. tricornutum SDH       : ------------------VPNISKRILVLGA-GMVSKSVVDLLGRSA------------- :  28
T. pseudonana SDH        : ------------------LGEREQNILLLGA-GKVASSFAEYLGRSK------------- :  28
P. infestans SDH         : ------------MAEVSVSSPKKRKVVCFGA-GLVASPLVEYLSREQ------------- :  34
D. melanogaster SDH      : ---------------MEGSSESDKKVLVLGA-GMVSAPLVEWLHREK------------- :  31
H. sapiens SDH           : ------------------MG-TRRKVLVLGS-GYISEPVLEYLSRDG------------- :  27
S. cerevisiae SDH        : ------------------MG---KNVLLLGS-GFVAQPVIDTLAAND------------- :  25
A. niger SDH             : ------------------MPVAGSKVLLLGS-GFVTKPTVEVLSKA-------------- :  27
Salpingoeca SDH          : ------------------MP----KILLLGA-GFVAGPCLDYLLRRD------------- :  24
```

```
                                    *       140         *       160         *       180
A. aestivalis HBFD1      : ------PDLRLVIGGRNREKGDAVVSKLG-----------ENSEFVEVNVDS-------- :  95
A. thaliana At1g50450    : ------PELKIVVGGRNREKGEAMVAKLG-----------ENSEFSQVDIND-------- :  97
P. sitchensis            : ------PNVHLVLAGRNREKGAKLAAELG-----------ANTEFYQVSIDN-------- : 107
S. moellendorffii        : ------PDLQLVLAGRNRERGEGIASEIG-----------AGTKFRAVDIED-------- :  98
C. reinhardtii           : ------PNLKVTVASRSDDSFKAAVERRP---------ELSKAGFQRVDITN-------- : 103
E. siliculosis           : -----GPGLELLVGGRSQRNFVKSVERWRTLPGADEEYDYSDVKFVELDLGD-------- : 137
P. tricornutum           : ------RGSQVTVGGTNAERFLSSRKRWQ-SLFPDICEDLEEVDFVALNRER-------- : 117
N. punctiforme           : -------QAQITITGRSAEFGKAVSLSSG-----------GQVQFLVLDLVE-------- :  60
Synechococcus PCC 7002   : ------PLGAVTVTGRSGQTAKAL-----------------QGTFLRLDLED-------- :  52
Cyanothece ATCC 51142    : -------EAEVTITGRNPQLGMAALERFQ-----------DKVQFQVLDLSN-------- :  59

C. reinhardtii unknown 1 : ----QPGKVRWAIAGRDAKKLEQIRSELA-VRVNNPA--VAAVPILTADAND-------- :  78
C. variabilis unknown 1  : -----KTGVKWAIAGRSQERLEKLRLELS-EQYGSE---LQEVPILIGDLKS-------- :  73
Salpingoeca unknown 1    : ----ATKSLKWAIAGRNEAKLKDVRASVA-KKLKDET--VNKVDIIVADAKD-------- :  74
Synechocystis unknown 1  : ---ETAAAKNWAIAGRSQAKLDRLVQSLG--------PQAAHLPTFVVDIKD-------- :  71
R. pickettii unknown 1   : --SGDGEALRWAIAGRSEAKLLDVKRALG--------AAGESLPIIVADAAN-------- :  71
P. infestans unknown 1   : SALSSPSALKWAVAARSEAKLTQMKEQLKDRLPEVAPELIDAIPVVVADSGS-------- :  81
E. siliculosus unknown 1 : --GASPPTFKWAVAARSESKLTALKERLA-SEIDSA---ASTLPTIVADSLD-------- :  77
T. pseudonana unknown 1  : LHSWKRRNGKWAIAGRSKSKLESTLKSIA-KDLGNDE--VLKVDTIIVDTMD-------- : 161
P. tricornutum unknown 1 : ---GVNKSVKWAIAGRSEAKLDKVKRELA-DELDIQE--VESIDTIIVDTTR-------- : 148
A. niger unknown 1       : -----PTNLKWALAGRSLSKLENIAKELK------NVNPDRADPDLLPVQLN-------- :  71

A. thaliana At5g39410    : QTPSSSPLKSLALAGRNPTRLTQSLEWAA-RPNPP----PSSVAILTADTSD-------- :  81
P. sitchensis unknown 2  : -DRDGGQRINIGIAGRSRSKVAEALRWAA-APSLP----PANIPIIEADVTS-------- :  75
S. moellendorffii unk. 2 : -----PEDRKVGIAGRSRQKLEESLSWAL-GHTSEEQRLKSSVPIFEADVHD-------- :  75
C. variabilis unknown 2  : ------FSGKWAVAGRDRVKLERLAASLA-GNDG------NEPSIVVADVAD-------- :  69
E. siliculosis unknown 2 : -ADGSWAGVSWGIAGRSRTKLEDKVLAPL-RAEGLAV--PGEESITVVDNSD-------- :  77
T. pseudonana unknown 2  : ------PSLNVAICGRNKDKLAVVAAEVA-WDDAKK---ASSVFVVSDASKDTSGAESAN :  90
Salpingoeca unknown 2    : ------QDKSIAISGRSADKLAALNRKLG-----------TNYPVIVADVKD-------- :  71
H. sapiens unknown 2     : VDPERSSRLPWAVAGRSREKLQRVLEKAA-LKLGRPTL-SSEVGIIICDIAN-------- :  81
C. elegans unknown 2     : ----KFDNYSIAIAGRSEKKLEEVLEKLE-KSTGTSL--KTRIGLLVCDSTN-------- :  73
D. melanogaster unknown 2: ------RGLRWGIAGRNREKLEAVLKEMG-AKAKKD---LSQVPIFIADVND-------- :  71

A. thaliana At4g33150 SDH: ADSEEKTDVHVIVASLYLKDAKETVEGIS------------DVEAVRLDVSD-------- :  80
S. moellendorffii SDH    : DTSKPPKQVFVHVGSLYLEDASKVVEGVE------------NALAIQIDVMD-------- :  73
P. tricornutum SDH       : -------NQEITVASENHEEARLTAAFSK------------HDRHVGLGVVN-------- :  61
T. pseudonana SDH        : -------TNTITVASQYEADAMKTARYAT------------RGKAVTCDLSQ-------- :  61
P. infestans SDH         : -------GNEVHVVSGIEREVKEMMRKIS----------RRNIKPHVVNVAE-------- :  69
D. melanogaster SDH      : -------DVSITVCSQVKEEADRLAQQYA------------GVDSVYLDVNE-------- :  64
H. sapiens SDH           : -------NIEITVGSDMKNQIEQLGKKYN-------------INPVSMDICK-------- :  59
S. cerevisiae SDH        : -------DINVTVACRTLANAQALAKPSG-------------SKAISLDVTD-------- :  57
A. niger SDH             : -------DVNVTVACRTLESAQKLCEGFK------------NTKAISLDVTD-------- :  60
Salpingoeca SDH          : -------DNTITVACRTLERAQELAGDRQ------------RTAAISLNVKD-------- :  57
```

```
                                    *       200         *       220         *       240
A. aestivalis HBFD1      : --VRS-LESALE--DVDLVVHAAGPFQQA--EKCTVLEAAISTRTAYVDVCDNTSYSMQA : 148
A. thaliana At1g50450    : --AKM-LETSLR--DVDLVVHAAGPFQQA--PRCTVLEAAIKTKTAYLDVCDDTSYAFRA : 150
P. sitchensis            : --TKA-LAAALD--GVDLVVHAAGPFQRE--EKCTVLEAAISSKTAYIDVCDDRSYSCLA : 160
S. moellendorffii        : --LNS-VAQAIE--GAKLVIHSAGPFQRK--NRCTVLEAAIETKVPYIDVCDDLTHAQLA : 151
C. reinhardtii           : --ADS-VQALLKSTGADLVIHTAGPFQRS--KNYAVLEAAIASGTGYIDVCDDTPFAEGA : 158
E. siliculosis           : --AAS-LASALD--GCDLVVHTAGPFQRK--TRPEVLEAAIAAKVPYVDVCDDARLATVA : 190
P. tricornutum           : --SES-VQAVLS--GYDLVVHTAGPFQGKVNTPNGVIEACVENGVPYIDVCDDYCTASAI : 172
N. punctiforme           : --VDK-LQNAIA--NSNLVIHCAGPFHYR---DTNVLETCIAQGVNYVDVSDHRSYTSKA : 112
Synechococcus PCC 7002   : --LAA-LEKAIA--NHDLVIHCAGPFHRR---DGRVLQTCIYQGKNYIDVSDHRCLYQKL : 104
Cyanothece ATCC 51142    : --QHQ-VKTVVS--KADVVVHSAGPFHYR---NADVLRICIEEGVNYVDVSDQRSFTHKA : 111

C. reinhardtii unknown 1 : --APA-VGRVLA--QTQVVLSTAGPFARY---GDNVVAQAVEQGTHYADITGEIPWVKRS : 130
C. variabilis unknown 1  : --QPS-LDSIAS--QTTVMISTAGPFALY---GTPVVDAAVRSGTHYVDITGEAPWVSKI : 125
Salpingoeca unknown 1    : --APA-LNKMAK--ATKVVLSTTGPFWKF---GSDLVAACVTSGTSYCDITGESPWVKEM : 126
Synechocystis unknown 1  : --EAA-VTTLCS--QTKVVVSTVGPYALY---GETLVRVCAITGTDYCDLTGEVQWVQQM : 123
R. pickettii unknown 1   : --EAQ-LRALCA--QTRVVVSTVGPYALY---GEPLIKVCAESGTDYCDLTGETQWIKRM : 123
P. infestans unknown 1   : --EES-LAMMVQ--QTNVVVSLVGPYKLY---GELLVKVCAENGVHYCDLTGEIVWIEEM : 133
E. siliculosus unknown 1 : --DEA-VGGMVS--QAKVIITTVGPYAHY---GSKVVAACSAAGVHCCDLTGESLWVKGL : 129
T. pseudonana unknown 1  : --RST-LKALVD--NTRAVITTAGPFVKY-------------------DITGEVQWNKSM : 197
P. tricornutum unknown 1 : --TSS-MPALVK--QARCVATTAGPFALF---GSPVVEFCAKFGTHYVDITGESDWVKAM : 200
A. niger unknown 1       : --REE-LHPLVQ--KTRVIINCVGPYCLY---STPVIEACASNGTHYVDATGETHWVKEI : 123

A. thaliana At5g39410    : --PDS-LRRLCT--QTKLILNCVGPFRIH---GDPVVSACADSGCDYLDISGEPEFMERM : 133
P. sitchensis unknown 2  : --PPS-LAALCR--RTKLVLNCVGPYRLY---GEPVVSACVEGGVDYLDITGEPGFMEKM : 127
S. moellendorffii unk. 2 : --MAS-LRGLCS--KTKLLVSCVGPYRLY---GEPVVAACVEAGIDYLDITGEPEFMERM : 127
C. variabilis unknown 2  : --PAS-LLEMAK--SCRVLITTVGPFRHW---GEPVVKACVEAGTDYLDVCGEPEFIERV : 121
E. siliculosis unknown 2 : --AAA-LRKAVG--RARLCLNCTGPYRFL---GESVVSACVDSGTDYIDLCGEPEFMQRM : 129
T. pseudonana unknown 2  : DGSQE-LIQVFS--QSKIVIACAGPYRQC---GMPIITAAVASGCDYLDLCGEPQFFDQV : 144
Salpingoeca unknown 2    : --EES-IVAMAQ--QARVCLNCVGPYRFF---GEPVVKACAAVGTHYLDICGEPEFIERM : 123
H. sapiens unknown 2     : --PAS-LDEMAK--QATVVLNCVGPYRFY---GEPVIKACIENGASCIDISGEPQFLELM : 133
C. elegans unknown 2     : --EES-MGKMAR--RAKLIVNAVGPFRLH---GEAVVKAAVENGANQIDVAGEPEWIERM : 125
D. melanogaster unknown 2: --QAS-LLEMAK--KCRIVVNTAGPYRFH---GENVVKCCIESGTHHVDVSGEPQYMETM : 123

A. thaliana At4g33150 SDH: --SES-LLKYVS--QVDVVLSLL-PASCH----AVVAKTCIELKKHLVTAS----YVDDE : 126
S. moellendorffii SDH    : --EQQ-LKSQVQ--KVEVVISLL-PPSFH----ERVAVACIELKKHLVTAS----YVSKD : 119
P. tricornutum SDH       : --DVKRLSDHIE--SADKVVSLL-PPPMH----FQVALDCIKHKTDL------------- :  99
T. pseudonana SDH        : --PGDQLKYLIQ--EADIVVSLL-PAQMH----PTIAEECISMKTDLVTAS----YESEE : 108
P. infestans SDH         : --DAAGVDKLCA--EADCVVSLL-PATMH----TTIAQRCIQHATPLVTAS----YVSPE : 116
D. melanogaster SDH      : --STGHLQELCG--RADVVVSLL-PYSLH----GMVARYCVAEGTHMVTAS----YLNDE : 111
H. sapiens SDH           : --QEEKLGFLVA--KQDLVISLL-PYVLH----PLVAKACITNKVNMVTAS----YITPA : 106
S. cerevisiae SDH        : --DSA-LDKVLA--DNDVVISLI-PYTFH----PNVVKSAIRTKTDVVTSS----YISPA : 103
A. niger SDH             : --DAA-LDKALE--QVDLAISLI-PYTFH----ANVIKSAIRTKKHVVTTS----YVSPA : 106
Salpingoeca SDH          : --SEA-LVAEVQ--KHDIVISLI-PYTYH----PLVIEAAIKVKKHFVSTS----YVSPK : 103
```

```
                                    *       260         *       280         *       300
A. aestivalis HBFD1      : K-SFHDKAVAANVPAITTAGIFPGVSNVIAAELVRSARDENT---------------EPQ : 192
A. thaliana At1g50450    : K-SLEAEAIAANIPALTTAGIYPGVSNVMAAEMVAAARSEDKG--------------KPE : 195
P. sitchensis            : K-TFHEKAVAAQVPAITTAGIYPGVSNLMAAELVRLARTSSSKETPS----------EPE : 209
S. moellendorffii        : K-SLHNKAVEAGVPAIITTGLYPGVSNIMAAELVRLARESSAGS-------------RPR : 197
C. reinhardtii           : KAAYMEKAKAAGVPAIVSGGIYPGTSNVMAAHIISIARAEYDDNWNYRTPAPGESV-EPK : 217
E. siliculosis           : K-ALNEKAQEAGVSATISAGIWPGIDQLMAVEACEMLGGAS----------------EVE : 233
P. tricornutum           : KTKLAEKAKESKTPCIISTGCWPGVSSLMAKQLLQKSMQAYPDVDQ-----------RDI : 221
N. punctiforme           : L-NFSEQAAAAGVTAIINTGIFPGISNSMVRQGVEQFD-------------------KPE : 152
Synechococcus PCC 7002   : K-PLTQAAREAGIIAVCNAGVFPGISNSMVRLGVEQLD-------------------EPH : 144
Cyanothece ATCC 51142    : L-EQHSAAKTAGVTAIINTGVFPGISNSMVRQGVEALD-------------------KAD : 151

C. reinhardtii unknown 1 : VQRHHETAKKKGVKILHCCGYDSIPSDMGTFMMVEYCRDKLGC--------------GVS : 176
C. variabilis unknown 1  : IAAYHEEAAAKRVRIVPCCGFDSTPFDLGALLVIDHMRKQLGK--------------QPA : 171
Salpingoeca unknown 1    : IEKHDATAKANKCRIVSFCGMDCIPSDVGCHVVAKHLREKHKC--------------GVD : 172
Synechocystis unknown 1  : IQKYEAIAKQSGARIVHCCGFDSIPSDLGVYYLQQQSQRRWGE--------------PCI : 169
R. pickettii unknown 1   : IDKYEATAQRSGARIVHCCGFDSVPSDMGVYFLQQHAMRQWGT--------------PVT : 169
P. infestans unknown 1   : TAKYAAIAAKTGAVLVNCCGFESIPSDVTTFLVADRIQKKFNS--------------ATS : 179
E. siliculosus unknown 1 : IDKHHEEAERTGAKIVPSCGFDSIPADLGTLMMVEYMKRTHGL--------------SPD : 175
T. pseudonana unknown 1  : MQQYESTAQRTGAKIVSLCGHDSIPWDLTVRSLSEKLHESCND--------------ELV : 243
P. tricornutum unknown 1 : IMKWGETAQKSGARLVTFCGHDSVPWDLSVMKLQEALQKEFKD--------------DLT : 246
A. niger unknown 1       : ISEYHETAKANGAVIIPCVGIESAPADLLAWATVKRVREDLSC--------------HTR : 169

A. thaliana At5g39410    : EANYHDRAEETGSLIVSACGFDSIPAELGLLFNAKQWVSPSVP--------------NQI : 179
P. sitchensis unknown 2  : EAAYHQRAEETGSLVISACGYDSIPAELGVIFHSKQWKPPSVP--------------HSV : 173
S. moellendorffii unk. 2 : RAKYHEQAVDRESLVVSACGFDSVPAEFGVIFNTKQWQGSSKP--------------NSI : 173
C. variabilis unknown 2  : ELLYNETAKQAGCYLASAVGFDSVPGDLGVAYTMSLFKPPARC--------------TVV : 167
E. siliculosis unknown 2 : TLKFHEAAEAKGVLIMHACAFDSVPADLGCLFAAKQFVSPAVC--------------SSV : 175
T. pseudonana unknown 2  : LISLDRKARSSGVLAISAAAFDCVPAELGAALAERELLKQHPG--------------AKC : 190
Salpingoeca unknown 2    : EYLYDEQARQTGATIISACGFDSIPADLGTMFTVKQFTNGQIP--------------SSV : 169
H. sapiens unknown 2     : QLKYHEKAADKGVYIIGSSGFDSIPADLGVIYTRNKMNGTL----------------TAV : 177
C. elegans unknown 2     : EAKYGQMAKNNNVYIVSACGWDSIPADFGVKLLKRYFDGHL----------------QRI : 169
D. melanogaster unknown 2: QLRYDQLAREKGVYVVSACGFDSIPADMGVVFVEKNFDGVV----------------NSV : 167

A. thaliana At4g33150 SDH: TSMLHEKAKSAGITILGEMGLDPGIDHMMAMKMINDAHIKKGKVKSFTSYCGGLPSPAAA : 186
S. moellendorffii SDH    : MALLDSRAQAAGVTLLCEMGLDPGIDHMMAMKMIDASHERGDKVRVFESYCGGLPSPEAA : 179
P. tricornutum SDH       : -----GRAKEAGVIILNESGLDPGLDHMSAMKKIDDILSRGGRVTVFASVCGGLPSPEAA : 154
T. pseudonana SDH        : MRALCSSAEEAGIAILNEMGLDPGVDHMSAMKIIDDVHERGGEITSFSSVCGGLPSPEVA : 168
P. infestans SDH         : MKELDARAKKAGIPILCEIGLDPGMDHMSAMKVIDEVKALSGKVMTFSSVCGGLPAPEAA : 176
D. melanogaster SDH      : ISGLHEEAKAKGVTIMNEVGLDPGIDHLLALECIHEVQDKGAVVESFVSYCGGLPAPEHS : 171
H. sapiens SDH           : LKELEKSVEDAGITIIGELGLDPGLDHMLAMETIDKAKEVGATIESYISYCGGLPAPEHS : 166
S. cerevisiae SDH        : LRELEPEIVKAGITVMNEIGLDPGIDHLYAVKTIDEVHRAGGKLKSFLSYCGGLPAPEDS : 163
A. niger SDH             : MMELDEECKKAGITVMNEIGLDPGIDHLYAVKTISEVHAEGGKITSFLSYCGGLPAPECS : 166
Salpingoeca SDH          : MASFDQAAKDAGITVMNEIGVDPGIDHLYAKKIIDEAHEEGSKVLEFTSYCGGLPAPEAS : 163
```

```
                                    *       320         *       340         *       360
A. aestivalis HBFD1      : R-LRFSYFTAG-SGGAG-------PTSLVTSFLL-------------------------- : 217
A. thaliana At1g50450    : K-LRFSYYTAG-TGGAG-------PTILATSFLL-------------------------- : 220
P. sitchensis            : R-LRFSYFTAG-SGGAG-------PTILATSFLL-------------------------- : 234
S. moellendorffii        : E-LRFSYFTAG-SGGVG-------PTILATSFLL-------------------------- : 222
C. reinhardtii           : W-LRYSYYTAG-SGGAG-------PTILETSFLL-------------------------- : 242
E. siliculosis           : S-IDFSAYTAG-TGNAG-------TTILSATFLI-------------------------- : 258
P. tricornutum           : S-VKFSFFTAG-SGGAG-------VTLLVATFLI-------------------------- : 246
N. punctiforme           : N-IHLSYLVSG-SGGAG-------ITVMRTTFLG-------------------------- : 177
Synechococcus PCC 7002   : Q-IELYYGVAG-SGGAG-------ETVLTTTFLG-------------------------- : 169
Cyanothece ATCC 51142    : E-IKLSYIVAG-SGGAG-------ITVMRTTFIG-------------------------- : 176

C. reinhardtii unknown 1 : Q-AYSLVGPGR-AGVSG-------GTLESGMNLI-----MNESVSELSELNSNHFTGEGA : 222
C. variabilis unknown 1  : K-VLNAVMGSK-GGVSG-------GTIASGMNAI----SEMNSNPALKASARDVYSLV-- : 216
Salpingoeca unknown 1    : H-INSYILAMK-GGVSG-------GTIASMFNFF-----EQPNFAELRKEMGSPYELIPR : 218
Synechocystis unknown 1  : R-VKMGVKACQ-GGISG-------GTIASGINLI----QEAIADPSTRQALRNPYILCFK : 216
R. pickettii unknown 1   : Q-VKLRVKTLK-GGASG-------GTVASLVNVV----QEAAADPALRRELADPYALCPR : 216
P. infestans unknown 1   : T-VDLYFTEFK-GEASG-------GTLASVFAVM------ETSTSKQLLASRNPFFLTDE : 224
E. siliculosus unknown 1 : D-VRYYFGASK-GGVSG-------GTIASVLDIF---EQVWSGGKAITSKLADPLLLT-- : 221
T. pseudonana unknown 1  : S--VECLNEMK-GDVSG-------GTLATVYESIDRGIWEFEWKSILPSSTLDVYKRLPD : 293
P. tricornutum unknown 1 : S-VTFWDEAV--GGAPG-------GTIATLFSAV------DGKSVKAPHADFDPFLRLMN : 290
A. niger unknown 1       : S-ITGDIHEIKSSGASG-------GTLSTVLTFF------ENVPPSEMRKISTPFALAAA : 215

A. thaliana At5g39410    : E-AYLSLESDK-KIAG------NFGTYESAVLGV--------ANAEKLKELRRSRPRRPR : 223
P. sitchensis unknown 2  : D-SYLALESDK-RIVG------NIGTFQSAVLGV--------ANVHDLQKLRRSRPRRAK : 217
S. moellendorffii unk. 2 : D-AYLTLRSST-RMKG------NFATFESAVLGI--------ASVGELQKLRKSRPIKSP : 217
C. variabilis unknown 2  : E-TALTIRGGP-SGFKG-----HYPTYESAVHGF--------ASAGELRKLRKEAEQ--- : 209
E. siliculosis unknown 2 : S-SFVTLNVGP-SGYSG-----HATTFEAAVHGF--------GSAADLRKVRKEVQAKFP : 220
T. pseudonana unknown 2  : A-GVEVIHTMQ-NVSGA-----NATTFHAAVDGF------YAASSGELAASRKKVKESYP : 237
Salpingoeca unknown 2    : E-SFLQIKSGE-AGTKV-----HFATFESAVHGF--------ASAGELRDLRRKKGKVTV : 214
H. sapiens unknown 2     : E-SFLTIHSGP-EGLSI-----HDGTWKSAIYGF--------GDQSNLRKLRNVSNLKP- : 221
C. elegans unknown 2     : D-AFLQLHFGP-SGYSF-----SATSLQALLLGF--------NGAPDLGSLRHSIMPKKI : 214
D. melanogaster unknown 2: E-TFLETGIKE-GGSGGGTAGLNYGTWESAVYGL--------AHSDELRGIRKQIYPQR- : 216
A. thaliana At4g33150 SDH: N-NPLAYKFSW-NPAGAIRAGQNPAKYKSNGDII------HVDGKNLYDSAARFRVPNLP : 238
S. moellendorffii SDH    : N-NPLAYKFSW-NPTGAIKAGRNAATYKHENKII------RVPGERLFGAAVSFRIPQYP : 231
P. tricornutum SDH       : D-NPLKYKFSW-SPKGVIQTSQSDARYRWEGQVL------QVHGSDLLAAAAPFLH-AWP : 205
T. pseudonana SDH        : NHNPLLYKFSW-SPMGVMKASQNAAVYRKDNQLV------VIDGADLLASAEPFD--AWK : 219
P. infestans SDH         : D-NAIGYKFSW-SPRGVLTAALNAAQYRKDGKVI------NVAGEDLLNRSERVN--FLP : 226
D. melanogaster SDH      : N-NALRYKFSW-SPRGVLLNTLSAAKYLSQGQIV------EISGGGELMSSPRSLD-FLP : 222
H. sapiens SDH           : N-NPLRYKFSW-SPVGVLMNVMQSATYLLDGKVV------NVAGGISFLDAVTSMD-FFP : 217
S. cerevisiae SDH        : D-NPLGYKFSW-SSRGVLLALRNSAKYWKDGKIE------TVSSEDLMATAKPYF--IYP : 213
A. niger SDH             : N-NPLGYKFSW-SSRGVLLALRNAAKFYKDGQEF------SVAGPDLMATAKPYF--IYP : 216
Salpingoeca SDH          : N-NPLGYKFSW-SARGVLLAAGNSCKYFEHGKKV------EIEAPYVLSKGVRNIN-IYP : 214
```

```
                                    *       380         *       400         *       420
A. aestivalis HBFD1      : -----------LGEEVVAYSEGEKVELKPYTGKLNIDFGKGVGKRDVYL----------- : 255
A. thaliana At1g50450    : -----------LGEEVTAYKQGEKVKLRPYSGMITVDFGKGIRKRDVYL----------- : 258
P. sitchensis            : -----------LREEVIAYRKGEEFRAKPYSGVLDIDFGKGVGKRSVYL----------- : 272
S. moellendorffii        : -----------LSEQVLTYSKGKVVKLDPFSGERIVDFGKAVGKRSVFL----------- : 260
C. reinhardtii           : -----------AGEDVIVYKDNKEVVLPPISNRREVDFGPGVGRKGVYL----------- : 280
E. siliculosis           : -----------LCEKVLGFKDGNEIFHEPASGFKKVDFGQSIGEKTRGMCTPAAWLTS-- : 305
P. tricornutum           : -----------LAEKALTVVNGRRKVVTPMETYSTVHFGPVVGDKEVAH----------- : 284
N. punctiforme           : -----------LQYPFETWIDGKWQVIKPYSERELVEFPPPYGRSGVYW----------- : 215
Synechococcus PCC 7002   : -----------LGEPFLVFQGGTWQAKQPYSKPTIIDFPAPIGKTTVYW----------- : 207
Cyanothece ATCC 51142    : -----------IQQPFEAWIDNRWQTIKPYTGRECLDFGKPYGKAHVYW----------- : 214

C. reinhardtii unknown 1 : ARE--------GLPKGPDRPAGLLPRWLPAARTWAGPFIMEGCNAKIVQASHALLRDS-- : 272
C. variabilis unknown 1  : -----------PEEAHGSDGEFWGVEWCPELGKYLAPFVMQVVNNRVVHRSNYLNKYD-- : 263
Salpingoeca unknown 1    : -----------DGRPSIRQPDANAVSYEEEVQRYTTPFVMAGVNSKIVHRTNYLLNHI-- : 265
Synechocystis unknown 1  : -----------PDHRSDHPPSLIPVQNDPIFSGWTAPFVMAGINTPIVLRSNTLQNEV-- : 263
R. pickettii unknown 1   : -----------GHGFTARQRSVRGAAFDADFNVWIAPFVMAAINERVVHRSNALSGNA-- : 263
P. infestans unknown 1   : KTIAEKQ----QAGLVAPNTSGIAIKYDTAMGFWHSLFIGGSVNQAVVHRSNYRLKNK-- : 278
E. siliculosus unknown 1 : -----------ATPGFGKAADPGGLGYDSLAKSWTASSVFASHDSKIVFRSAGLLG---- : 266
T. pseudonana unknown 1  : GCESTFS----VSSDLPVTVSPCRNPATRFVNRWCGPFVMAAINMDVVGRSISLSKHG-- : 347
P. tricornutum unknown 1 : GEKSEYVCRS-DGPTWIHKSKSPWDDTGAWFGRWTTPFIMAQVNGAVVRWSHALREQG-- : 347
A. niger unknown 1       : PPP--------KDIPREPLWTRLLGIRSVRDLGILTTSPSGLADITTVHRSSTLMPEF-- : 265

A. thaliana At5g39410    : -----------PTICGPPA-KGPTLENQKTIGLWALKL--PSADAVVVRRTLTTLTEKPH : 269
P. sitchensis unknown 2  : -----------PQIPGPPARKAKLIEHISSIGLWALKL--PSSDAIVVRRTQTFLAEDPQ : 264
S. moellendorffii unk. 2 : R----------LQIPGVPL-KRPAVHWEDAVNSWAVKI--PSSDAVLVRRTLATVAENPD : 264
C. variabilis unknown 2  : ---------------------------------------ARADASVVRRSMQRLVEA-- : 227
E. siliculosis unknown 2 : PSQIP------RVGARPVERGGPFYEQTPGIEAYCFKF--PGADSAVVRSTQNSLAGRGE : 272
T. pseudonana unknown 2  : EFQETMPPSRPKEWPKIPETPGLMPGHNEGLGLRTLKF--VGADASAIRSSWRYLRSR-- : 293
Salpingoeca unknown 2    : -----------PVVGRKLKRPGMLPEWREDGHGYCIPF--PGSDASIVRRSQQYLQAH-- : 259
H. sapiens unknown 2     : -----------VPLIGPKLKRRWPISYCRELKGYSIPF--MGSDVSVVRRTQRYLYEN-- : 266
C. elegans unknown 2     : D----------HLLAPNKKRHFLWKIEEKGSEGWAMPF--LGADKSIVTRSQYFDYVM-- : 260
D. melanogaster unknown 2: -----------LPRFYPFLKPRPLVFRSTEVDKVCLPF--PGSDRSVVMRSQRFLYDQ-- : 261
A. thaliana At4g33150 SDH: AFALECF----PNRDSLVYGEHYGIESEATTIFRGTLRYEGFSMIMATLSKLGFFDSEAN : 294
S. moellendorffii SDH    : AYALEVL----PNRDSLMYGDLYGISQEAATIFRGTLRYEGFGQIMDTLGKLGYYNSDNH : 287
P. tricornutum SDH       : KLGLECL----PNRDSLRYEHIYNIG-----GAKATLRYRGFSSCSYGL----------- : 245
T. pseudonana SDH        : SLNLECI----PNRDSLVYGEKYGIE-SAATIFRGTLRYQGFSSLLHVLKNMGLLDNKG- : 273
P. infestans SDH         : AFNIEQI----PNRNSLPYGDIYGIP-EAHSLYRGTLRYGGCCQILYQLRKLGLFDMDP- : 280
D. melanogaster SDH      : GFALEGF----PNRDSTKYGNLYGLGRDVHTLLRGTIRYKGFSESIKPMQLLGLIDPEPH : 278
H. sapiens SDH           : GLNLEGY----PNRDSTKYAEIYGIS-SAHTLLRGTLRYKGYMKALNGFVKLGLINREAL : 272
S. cerevisiae SDH        : GYAFVCY----PNRDSTLFKDLYHIP-EAETVIRGTLRYQGFPEFVKALVDMGMLKDDAN : 268
A. niger SDH             : GFAFVAY----PNRDSCPYRERYNIP-EAQTVVRGTLRYQGFPEMIKVLVDIGFLSDEGR : 271
Salpingoeca SDH          : AFAFEGY----PNRDSTPYEERYNIP-ECLTILRGTLRYKGTPLLVQSLALVGFLNDHPQ : 269
```

```
                                     *       440         *       460         *       480
A. aestivalis HBFD1      : ------------------------------WNLPEVRSGHEIL-GVPTVSARFGTAPFFW : 284
A. thaliana At1g50450    : ------------------------------LNLPEVRSTHEVL-GVPTVVARFGTAPFFW : 287
P. sitchensis            : ------------------------------LNLPEVGSTHEVL-GVPTVSARFGTAPFFW : 301
S. moellendorffii        : ------------------------------INLPEVTTAHEIL-NIPNVTAKFGTAPFFW : 289
C. reinhardtii           : ------------------------------YNLPEVVSGHKYM-RVPDVSARFGTDPFIW : 309
E. siliculosis           : -----------------TSGRCRCERQGSAVVLVASFTCNQVL-GIPNISTYFGTSPEPW : 347
P. tricornutum           : ------------------------------LNLLETASVADVL-RIGNVQALFGTAPNFW : 313
N. punctiforme           : ------------------------------FDMPETFTLPKAFPSVKTVITKFGSVPDFY : 245
Synechococcus PCC 7002   : ------------------------------FDVAETFTFAESF-PVETVVTKFGSLPNFY : 236
Cyanothece ATCC 51142    : ------------------------------YDMPETTTLQETF-SVDSVITKFGSVPDFY : 243

C. reinhardtii unknown 1 : -------------AAPLEYGKDFKFTEMIAAS-----GAVGAS-LVSAGTMVAGAVLGLA : 313
C. variabilis unknown 1  : -----------------GGSSTLRYQEAVAAK-----SWLAAK-SVQLGTLAVVGAMGQS : 300
Salpingoeca unknown 1    : ------------------YGPEFRYREKNTVKGDGWGSWLSGW-LTTLGLGAFSVGAYFS : 306
Synechocystis unknown 1  : ------------------YGEGFQYEEGILTGPGII-GWLAAQ-GIKWSLDLMALALAIG : 303
R. pickettii unknown 1   : ------------------YGNPFLYDEAVITGAGLS-GRMKAM-TMVAGLAAFMLGIVIK : 303
P. infestans unknown 1   : ------------------YGDKFVYHERMAIG-----GLFMQL-LATFGTIVVSIMLYFG : 314
E.siliculosus unknown 1  : ------------------YPETFRYKEVMGFKGLLK-GFLPAV-FSTIAFGIGGFLMVIP : 306
T. pseudonana unknown 1  : -------------------NPTVTYREAAVQE-----SFMDAF-SMWFGTVILGTLIINP : 382
P. tricornutum unknown 1 : -------------------SKSLTYTEFAVTK-----DFKTAF-VNHVGLIIVGSMMLNP : 382
A. niger unknown 1       : ------------------YGPRFYFRQFLRAR-----NAFTGI-LWHYAFLFAVTALLLP : 301

A. thaliana At5g39410    : GLPGINESPEQIQKREAFWSSIKPA-HFGVKITSK--SLFGIF-RYVTLGVSLGLLSKFS : 325
P. sitchensis unknown 2  : GLPGVNENSDFAEKRKNFWSEVKPL-HYGVYIGYK--SLLSVA-GTIFTGLNVLLLGSFG : 320
S. moellendorffii unk. 2 : GLPTASKHQEQSLRQ---WTDIKPV-LFGVYLSVK--ELWRVC-FLLLTGFILYVLANFG : 317
C. variabilis unknown 2  : ---------------------GQPAANVSVVFTLP--SRYYLT-LWQGFGSMFQFLAGKP : 263
E. siliculosis unknown 2 : G------------------AGLCP--HYSAYFTAG--QLWGAT-QMTLFGGVFQTLAKSG : 309
T. pseudonana unknown 2  : ------------VPEHARKGKNVPEPRLSVLMGMDSKDTMSAA-KLIVYGATFSTLAQFK : 340
Salpingoeca unknown 2    : ------------------GAEVQPV-QYAAWFSLP--SGFWTG-VMMAAGTVFGALANFS : 297
H. sapiens unknown 2     : -------------------LEESPV-QYAAYVTVG--GITSVI-KLMFAGLFFLFFVRFG : 303
C. elegans unknown 2     : -------------------NNVKPV-RFTPFTRFE--SRWNAL-LLAALMGLLKTFSKYG : 297
D. melanogaster unknown 2: -------------------DKKRPV-QMQAYVGFS--SWLVAG-GVIFFATIFGLLSKFK : 298

A. thaliana At4g33150 SDH: QVL-STGKRITFGALLSNILNKDADNESEPLAGEEEISKRIIKLGHSKETAAKAAKTIVF : 353
S. moellendorffii SDH    : PLL-ASSTETTYAAVLEALILQLSTSYNGLCAEELARIIS----SDNLDVAKRVLSCIRF : 342
P. tricornutum SDH       : ---------------FCLRDQRGGSLNVEDFILACV--------ADNFDEAVAVLDTLIW : 282
T. pseudonana SDH        : ------TGAVSWYDALDDLRKQGGHADLRKFVLACA--------GGDRDLGLRAYNCLSW : 319
P. infestans SDH         : ----SKPIPATWPDLLTQLGGFQGLRE-------------------------DAHGFLQW : 311
D. melanogaster SDH      : ALLHPSGPDVTWRQLVIHLMGMSDSTIFYENL------------KQKLTERIGDVDGIES : 326
H. sapiens SDH           : PAFRPEANPLTWKQLLCDLVGISPSSEHDVLKEAV---------LKKLGGDNTQLEAAEW : 323
S. cerevisiae SDH        : EIF---SKPIAWNEALKQYLGAKSTSKEDLIASIDSKAT-----WKDDEDRERILSGFAW : 320
A. niger SDH             : EYL---NTPIAWKEATKQILGATSSAEKDLEWAIASKTA-----FANNDDRDRIISGLRW : 323
Salpingoeca SDH          : DYLQPTAPDMAWKDVLAKMLGCTND-EASLAKAAAERAG-----IRDHPEERRIISAMKW : 323
```

```
                                    *       500         *       520         *       540
A. aestivalis HBFD1      : NWAMVAMTTLLPP-GILRDRNKIGMLANFVYPSVQIFDGI-AGECLA------------- : 329
A. thaliana At1g50450    : NWGMEIMTKLLPS-EVLRDRTKVQQMVELFDPVVRAMDGF-AGERVS------------- : 332
P. sitchensis            : NWAMSAVVNLAPV-EFLKDRNKVQTLVQLSDPLVRAIDIF-SGELMS------------- : 346
S. moellendorffii        : NWLMSQIAGM----EFLEDYSKVESFARLSDPAVRAIDVL-VGGKVS------------- : 331
C. reinhardtii           : NWAMWLTARLVPR-SLLNDRNFVKGFAKLSDPFVRNVDKI-IGEAVA------------- : 354
E. siliculosis           : NYLLKGMT-LLPD-SIMGNRDLMQALAEFSEPLVRITDKL-VGATNA------------- : 391
P. tricornutum           : NGLLGLMA-KLPT-QLLENEDLMRKLSMFSLPIVRLVDAF-AGATNA------------- : 357
N. punctiforme           : NHLTWIAAHIFPK-WLMQRRYMIEFLSHVSHSMTDVTNNF-SGIGVA------------- : 290
Synechococcus PCC 7002   : NQLTRAMT-LLPE-SLRQHPRIIQGLSKIGYGMTKLTDSF-TGVGVA------------- : 280
Cyanothece ATCC 51142    : NYATWMMAHGLPE-KVLKSPKTVEFLAQVSHVMTDVSDRF-TGTGVA------------- : 288

C. reinhardtii unknown 1 : PLRAVARR-FLPA--------PGQGPSEDVQRNGFWTHDL-VAVTDE------------- : 350
C. variabilis unknown 1  : WMHPLLKK-LLPA--------PGEGPSRETMMTGYFKNRV-VGWTKE------------- : 337
Salpingoeca unknown 1    : PTRNLLRR-MLPK--------PGEGPDEHTRETGKLVMRF-YGVGDS------------- : 343
Synechocystis unknown 1  : PSRWLLLQ-ILPK--------PGEGPTEEAQQQGFYDLRF-WGETAS------------- : 340
R. pickettii unknown 1   : PARIVMQRFLLPK--------PGEGPTPAAQLAGRYDLRL-FGRTDQ------------- : 341
P. infestans unknown 1   : WTRALLKR-LARA--------PGQGPSEESMLQGYFVAEA-AGYTDD------------- : 351
E. siliculosus unknown 1 : VTRKFIAKKFLPA--------PGEGPSKELRDSGYFWINF-LASGRS------------- : 344
T. pseudonana unknown 1  : ITRPLVKR-ALPQ--------PGQGPSETVRKEGYLCVTG-YGVGAE------------- : 419
P. tricornutum unknown 1 : LTSGLLKKYVFPK--------VGEGPSMKNMEDKHYLCIFGEGIGIN------------- : 421
A. niger unknown 1       : PIRTLVRKYIYTP---------GTGPTLEDSVNDFVEYRA-VATADQ------------- : 338

A. thaliana At5g39410    : FGRWLLLK--FPS-VFSLGWFQKKGPSEEEVESATFKMWF-IGRGYS------------- : 368
P. sitchensis unknown 2  : WGRSLLLK--YPE-IFTLGLFKKTGPTEEEVKSATFKMTF-LGHGYS------------- : 363
S. moellendorffii unk. 2 : WGRKLLLS--YPE-VFTGGVFSKTGPSQEEIDNSSFSMVF-VGRGFK------------- : 360
C. variabilis unknown 2  : WGRSLLLQ--YPR-LFTYGMFSHEGPSERQLSEASFQFTN-IAKGYS------------- : 306
E. siliculosis unknown 2 : WGRNLLLN--NVG-AFSRGLFSHEGPTEDQMNETSFEMTF-LAKGYSTPPTAAEPPATPA : 365
T. pseudonana unknown 2  : WGCDVLHN--YPE-AFSGGVFTSKGQSEEELQNAKFTTYV-TAYGSN------------- : 383
Salpingoeca unknown 2    : FGRSLLLK--FPE-LFTLGVFTHDGPTEEQMRATSFEMNF-VAKGYT------------- : 340
H. sapiens unknown 2     : IGRQLLIK--FPW-FFSFGYFSKQGPTQKQIDAASFTLTF-FGQGYS------------- : 346
C. elegans unknown 2     : TMQRFIMR--YPE-LCSGKLFSKSGPSEEQMKEATFTYQF-YGYGYG------------- : 340
D. melanogaster unknown 2: LGRTLLLN--YPG-LFSGGLASRSGPSEESMERTYFKMTF-KASGWL------------- : 341

A. thaliana At4g33150 SDH: LG-FNEEREVPSLCKSVFDATCYLMEEKLAYSGNEQDMVL-LHHEVE------------- : 398
S. moellendorffii SDH    : LG-LDSQEIVPRSCKSAFEVLCSRMEEKLVFRANEQDLVL-LHHELE------------- : 387
P. tricornutum SDH       : LG-VLPGHIPVSG-SNIVEAFFNFLEDKLRYKESESDMVL-MHHIIE------------- : 326
T. pseudonana SDH        : LG-LKEHTPVSEP-SSIAKSFCDVLQQHLQFEEGERDMVL-MHHDIR------------- : 363
P. infestans SDH         : LGAFDKSNPVVKA-PSILDAFCALLQDKLSYQPGERDMAI-MHHEFG------------- : 356
D. melanogaster SDH      : LG-LLDDTPVVKL-NTPLDTLSHYLSKRLAFERDERDLVV-LRHEVG------------- : 370
H. sapiens SDH           : LG-LLGDEQVPQA-ESILDALSKHLVMKLSYGPEEKDMIV-MRDSFG------------- : 367
S. cerevisiae SDH        : LG-LFSDAKITPR-GNALDTLCARLEELMQYEDNERDMVV-LQHKFG------------- : 364
A. niger SDH             : IG-LFSDEQITPR-GNPLDTLCATLEKKMQYEEGERDLVM-LQHKFE------------- : 367
Salpingoeca SDH          : LG-LFSDDVCPKK-VSLLDSLAHTLADKCAYQPGERDMIM-MQHKFR------------- : 367
```

```
                                    *       560         *       580         *       600
A. aestivalis HBFD1      : -------------MRVDLE---CANGRNTFGILS------------HERLSVLVGTSTAV : 361
A. thaliana At1g50450    : -------------MRVDLE---CSDGRTTVGLFS------------HKKLSVSVGVSTAA : 364
P. sitchensis            : -------------MRVDLE---CMDGRKAVGLYS------------HKKLSICVGVAISA : 378
S. moellendorffii        : -------------MRVDLD---CKDGTRSTGLYT------------HPDLSVCAGAATAA : 363
C. reinhardtii           : -------------MRVEVD---MVGGKNSSGIFV------------HKYLSQSMGYSTAA : 386
E. siliculosis           : -------------MRLEAV---AKDGRKAVLNYA------------HEDLEVCVGIATAA : 423
P. tricornutum           : -------------MRCDISNTKNPDLFRCSAIYA------------HKNLEPCVGECVTA : 392
N. punctiforme           : -------------VRSEVTGQKDGKTAVYCSTVV------------HENTALASGCGTGS : 325
Synechococcus PCC 7002   : -------------MRAIVSGIKDATPQQVTVDFV------------HEHTAIAAGLGVAL : 315
Cyanothece ATCC 51142    : -------------MRCDVKGYGSEGAAHYISTFS------------HESASVAAGMGTGS : 323

C. reinhardtii unknown 1 : ----------------------EKPRVVRGKCGD-----------RRDPGYWSTSRMLLE : 377
C. variabilis unknown 1  : -------------------EGGAQPSVVQAEVAD----------PHRDGGYWGTSRMLLE : 368
Salpingoeca unknown 1    : ---------------------AAKPKAVATFVTH------------HDPGYAETATMLAE : 370
Synechocystis unknown 1  : -----------------------GQSTMVKVTGD------------QDPGYGSTAKILAQ : 365
R. pickettii unknown 1   : -----------------------GNTLRVKVTGD------------RDPGYGSTGKMLGQ : 366
P. infestans unknown 1   : ----------------------GKLAVKAKTVGS------------GDPGYRLTSRLISE : 377
E. siliculosus unknown 1 : --------------------GDGKDVVCRGKVGS----------DKGDCGYKETAKMLAE : 374
T. pseudonana unknown 1  : -----------------------GNVAESAMYFP------------HDGGYQSTARMLVE : 444
P. tricornutum unknown 1 : -----------------------GNRAESIMYFD------------KDAGCLETSRMLVE : 446
A. niger unknown 1       : --------------------DTAKPQRVLGKLRY------------QGTMYEFTGLSLAE : 366

A. thaliana At5g39410    : ------------EESLASQGETKPDLEIITRISG------------PEIGYITTPITLVQ : 404
P. sitchensis unknown 2  : ------------DESCLLQQKTKPDMEIVTRISG------------PEVGYVTTPITLVQ : 399
S. moellendorffii unk. 2 : -----------DASKVPPGKKQQPDMEIITRVTG------------PEIGYVTTPIILVQ : 397
C. variabilis unknown 2  : -------------KGAPEAPDQAPDIEIVTRVSG------------PEPGYISCAIFVVQ : 341
E. siliculosis unknown 2 : AGADPGEVEAAAAAPAPAPPTVEPDVTVVTRVKG------------PEPGYVATPIIFLA : 413
T. pseudonana unknown 2  : ------------------YSDGDSSQVARVKVGG------------PEPGYVATPALIVA : 413
Salpingoeca unknown 2    : -------------SAAAAEARGAPDVTVKTRVSG------------PEPGYVATPICIVQ : 375
H. sapiens unknown 2     : --------------QGTGTDKNKPNIKICTQVKG------------PEAGYVATPIAMVQ : 380
C. elegans unknown 2     : ---------------RGDPRDREPKKKLLVTCTG------------PDVGYIATSGCVLS : 373
D. melanogaster unknown 2: --------KSDRLAESSNQYTESPTKTLMVRVSG------------MNPGYGATCVAVLS : 381

A. thaliana At4g33150 SDH: -------------VEFLESKRIEKHTATLLEFGDIKN------GQTTTAMAKTVGIPAAI : 439
S. moellendorffii SDH    : -------------VVYEDSRSAERHSATLVAVGESCNQLKNESRRPHSAMARTVGLTVAI : 434
P. tricornutum SDH       : -------------ASFERA-RSERHLASPQVFGE----------EGISAMAKCVGYTTAS : 362
T. pseudonana SDH        : -------------AVFGDG-SNETLSCSMELYGD----------DRMTAMCKTVGFTAAI : 399
P. infestans SDH         : -------------IEYEDG-KKEKRTSTFVGYGS---------EKGDTIMAKTVGLSAAI : 393
D. melanogaster SDH      : -------------IRWPDG-RREERGINFVVYGQ---------PQGHSAMAMTVGKPAAI : 407
H. sapiens SDH           : -------------IRHPSG-HLEHKTIDLVAYGD---------INGFSAMAKTVGLPTAM : 404
S. cerevisiae SDH        : -------------IEWADG-TTETRTSTLVDYGK---------VGGYSSMAATVGYPVAI : 401
A. niger SDH             : -------------IEHKDG-QKETRTSTLCEYGV---------PGGYSAMAKTVGVPCGV : 404
Salpingoeca SDH          : -------------IQKKTG-DEYVVASTMLEYGI---------PFGATAMARTVGIPCGI : 404
```

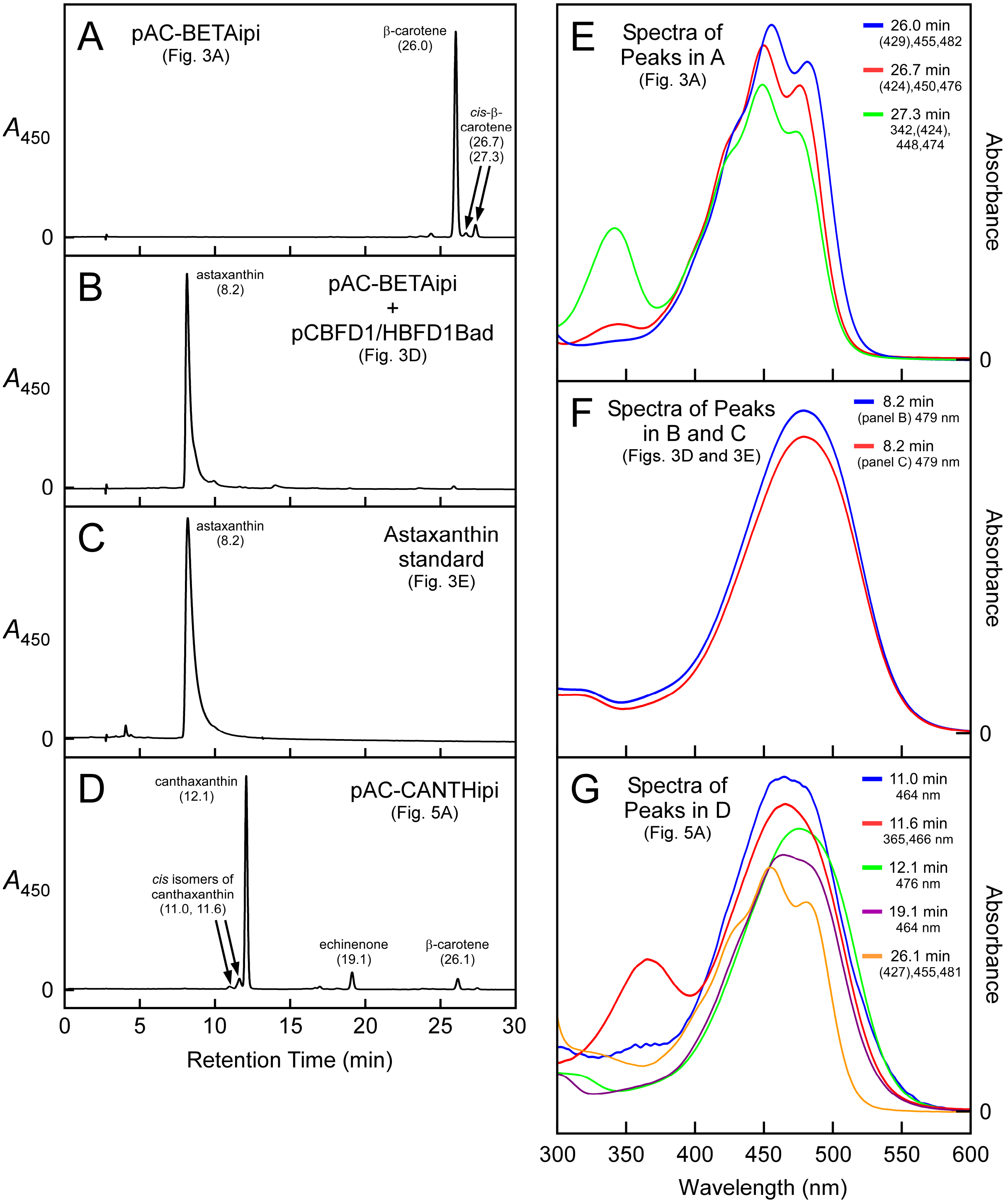
```
                                    *       620         *       640         *       660
A. aestivalis HBFD1      : FAMAILEGST---------QPGVWFPEEPGGIAISDRELLLQRASQGAINFIMKQ----- : 407
A. thaliana At1g50450    : FVAAMLEGST---------QPGVWFPEEPQGIAVEAREVLLKRASQGTFNFILNKPPWMV : 415
P. sitchensis            : FVRAVLEGNT---------QPGVWFPEEPEGIAVEARQQLLERAAEGTINFVMNRPPWMV : 429
S. moellendorffii        : FVVPILQGKT---------LPGVWFPEQREGIPIECREGLLEKASTGTSRYAMIVSYYHL : 414
C. reinhardtii           : FAQSVLQGKT---------QPGVWYPEEKEAL--QDRRQFLQFAATGCSRFELNRSAWAL : 435
E. siliculosis           : FVVATLRGDV---------RPGVWFPEEAFDD-EAKRGLLFDDATRGAFMWERQEQGATA : 473
P. tricornutum           : FAAAVLSGAV---------RDGIWFPEEAIQG-GVDAAAVLAAASVGAHTVLVSSEGMEL : 442
N. punctiforme           : IAQLLLEGKL--------KKPGVFAVEEALPT--DLFEEVMQSRGIKINHSWL------- : 368
Synechococcus PCC 7002   : VAELVLSEQI--------NQPGLYPVEQIIPS--DLF--LAWARQHQLQLSWNIQPSEKN : 363
Cyanothece ATCC 51142    : IVELLLKGEL--------VQPGVHPVEQALST--ELFQSILKSRQLTIEKTINFLQENEQ : 373

C. reinhardtii unknown 1 : TGLALSLDAERLAADPRLAPGGVLTAAAGCGH--VLLERLR--G-AGFTFEVAGVEGEGV : 432
C. variabilis unknown 1  : AALCLALQQQELDASPEVLKGGVLTPASAMGI--FLVERLR--N-AGQTFKILEGGASTA : 423
Salpingoeca unknown 1    : SAICLAQDTAKLPE-----RYGFLTPIAAMGD--TLVARLQ--A-AGIDITVEDDKNTSS : 420
Synechocystis unknown 1  : AALCLAKDRPKSSL-----QGGFWTPAASFGQ--DLIHRLI--DYAGLTFTEENGSTL-- : 414
R. pickettii unknown 1   : AAISLALDHTKDGLKTG-RPGGFWTPATMFDD--RFIERLT--RHAGLRFERI------- : 414
P. infestans unknown 1   : CAFCLAKDEFGDAKS---LKGGFYTPASAFGH--KLADRLQ--TKKFITFELKDMA---- : 426
E. siliculosus unknown 1 : CGLCLALDDLEY------KKGGVLTTASAMGM--PLVDRLN--K-AGMTFKILDE----- : 418
T. pseudonana unknown 1  : SGLCLALNEDRHEM----SQGGFYSPAGLMCA--ALLDRLV--N-SGTKFAIRSS----- : 490
P. tricornutum unknown 1 : SGLCLALEEDKVPA----KHGGFWPPAAGLGN--VLLKRLL--E-TGTSFEARAISEASV : 497
A. niger unknown 1       : AAMTIIENEEKVKKV---SRCGIVTTAT-LGQ--EFIDRWD--K-VGCHIETQVVDN--- : 414

A. thaliana At5g39410    : CGLIVLGQRESLVK------GGVYTPGIVFGS-TDIQQRLE--D-NGISFELISKIKTQG : 454
P. sitchensis unknown 2  : CALILIDQRHSLPK------GGVFPPGIVFGP-TDLQERLE--K-NGISFEVLSKRTIS- : 448
S. moellendorffii unk. 2 : AALLVLENRDKLPK------GGVWTPGVAFGT-TDYEQRLQ--N-NGLSFDVISKH---- : 443
C. variabilis unknown 2  : AAITLLEERQSLGL------PGVHTPASLLRD-TTYIDRLR--S-RGIKFEQVQDATVQ- : 390
E. siliculosis unknown 2 : VARCLLEERSSLPV-----SGGVHTPGSVLVS-SSLVDRLG--K-DGVTFEVVG------ : 458
T. pseudonana unknown 2  : LALTILEAGK--YEVGLAFDSGCTLPGALFGDCEKVYDNMR--A-EGVSFDVVQEFGGES : 468
Salpingoeca unknown 2    : CALALIGNEGKVPR-------GVLTSATAFRH-TDILQRLN--G-AGVKFEVL------- : 417
H. sapiens unknown 2     : AAMTLLSDASHLPK-----AGGVFTPGAAFSK-TKLIDRLN--K-HGIEFSVISSSEV-- : 429
C. elegans unknown 2     : SVLVFLKEKEKLPK-----SGGVYTTAYAFEN-TRLIDFLL--D-FGIKFDIEMPSKL-- : 422
D. melanogaster unknown 2: AALTILRESDKMPS-----TGGVLPPAAAFSK-TGLISELEKHD-HGIKFEILANK---- : 430

A. thaliana At4g33150 SDH: GALLLIEDKI--------KTRGVLRPLEAEVY-LPALDILQ--AY-GIKLMEKAE----- : 482
S. moellendorffii SDH    : GAELLFTGRL--------KSRGVIRPLQPEVY-VPGLEILG--KL-GLGVVES------- : 475
P. tricornutum SDH       : SAGLILSFALD-------DLLGLLLPTSPRVY-EPVLAAKK--ESPSTKRAHQIRQTR-- : 410
T. pseudonana SDH        : GTKLILEGGI--------TNKGLLLPTSKDVY-TPSLELLR--EE-GI------------ : 435
P. infestans SDH         : GVQLILQDAV--------QGRGVLTPTTPDIY-GPALARLE--VE-GVRFIEKTFPQH-- : 439
D. melanogaster SDH      : AAKMILDGEI--------QERGVLLPFTPDIY-RPMLQRLR--SE-GLTATETSRWLN-- : 453
H. sapiens SDH           : AAKMLLDGEI--------GAKGLMGPFSKEIY-GPILERIK--AE-GIIYTTQSTIKP-- : 450
S. cerevisiae SDH        : ATKFVLDGTI--------KGPGLLAPYSPEIN-DPIMKELK--DKYGIYLKEKTVA---- : 446
A. niger SDH             : AVKLVLDGTI--------NQKGVVAPMTMDIC-APLIKTLK--EDYGIEMIEKTL----- : 448
Salpingoeca SDH          : AVQLILDGKI--------TRKGVFAPMTKDIY-QPLIAELE--KE-GIGCKEEYLDEWTH : 452


                                    *
A. thaliana At1g50450    : ETEPKEVVLGIYV------ : 428
P. sitchensis            : ETNPKEIGFGIYMQ----- : 443
S. moellendorffii        : H------------------ : 415
C. reinhardtii           : ESEIKQIGGMIYW------ : 448
E. siliculosis           : VEGAPL------------- : 479
P. tricornutum           : SLEQVWGTRRKETITMTST : 461
C. reinhardtii unknown 1 : AAAK--------------- : 436
Salpingoeca unknown 1    : TPPSSL------------- : 426
P. tricornutum unknown 1 : RHSLE-------------- : 502
T. pseudonana unknown 2  : QSPV--------------- : 472
```

Supplemental Figure 6 (below). Absorption Spectra for Peaks in the HPLC Elution Profiles Displayed in Figures 3A, 3D, 3E, and 5A of the Manuscript. Panels (A), (B), (C) and (D) reproduce HPLC elution profiles of the manuscript figures. Shown are elution profiles for extracts of *E. coli* cultures containing the following plasmids: **(A)**, pAC-BETAipi; **(B)**, pAC-BETAipi and pCBFD1/HBFD1Bad; **(D)**, pAC-CANTHipi. Panel **(C)** displays an HPLC elution profile for a synthetic astaxanthin standard. Panels **(E)**, **(F)**, and **(G)** display absorption spectra for peaks in the elution profiles of panels (A), (B), (C) and (D). Numbers in parentheses below the carotenoid names in panels (A), (B), (C) and (D) are HPLC retention times in minutes. Labels for the absorption spectra in (E), (F), and (G) refer to the HPLC retention times for the corresponding peaks in (A), (B), (C) and (D). Absorption maxima are listed below these HPLC retention times, with peak "shoulders" enclosed in parentheses. Spectra were recorded in HPLC mobile phase just as the individual pigments were eluted from the HPLC column. **Note:** it can be seen that the peak for the astaxanthin standard in (C) is noticeably broader than that for astaxanthin in (B), despite a lower loading of pigment (*ca.* 1/3 the amount of that for the sample analyzed in panel B). This is probably because the standard compound was a racemic mixture, whereas astaxanthin produced through the action of the *A. aestivalis* enzymes in *E. coli* is most likely a single isomer, (3*S*, 3'*S*)-astaxanthin, as was reported for astaxanthin extracted from the flower petals of *A. aestivalis* (Kamata and Simpson, 1987) and *A. annua* (Renstrøm et al., 1981).

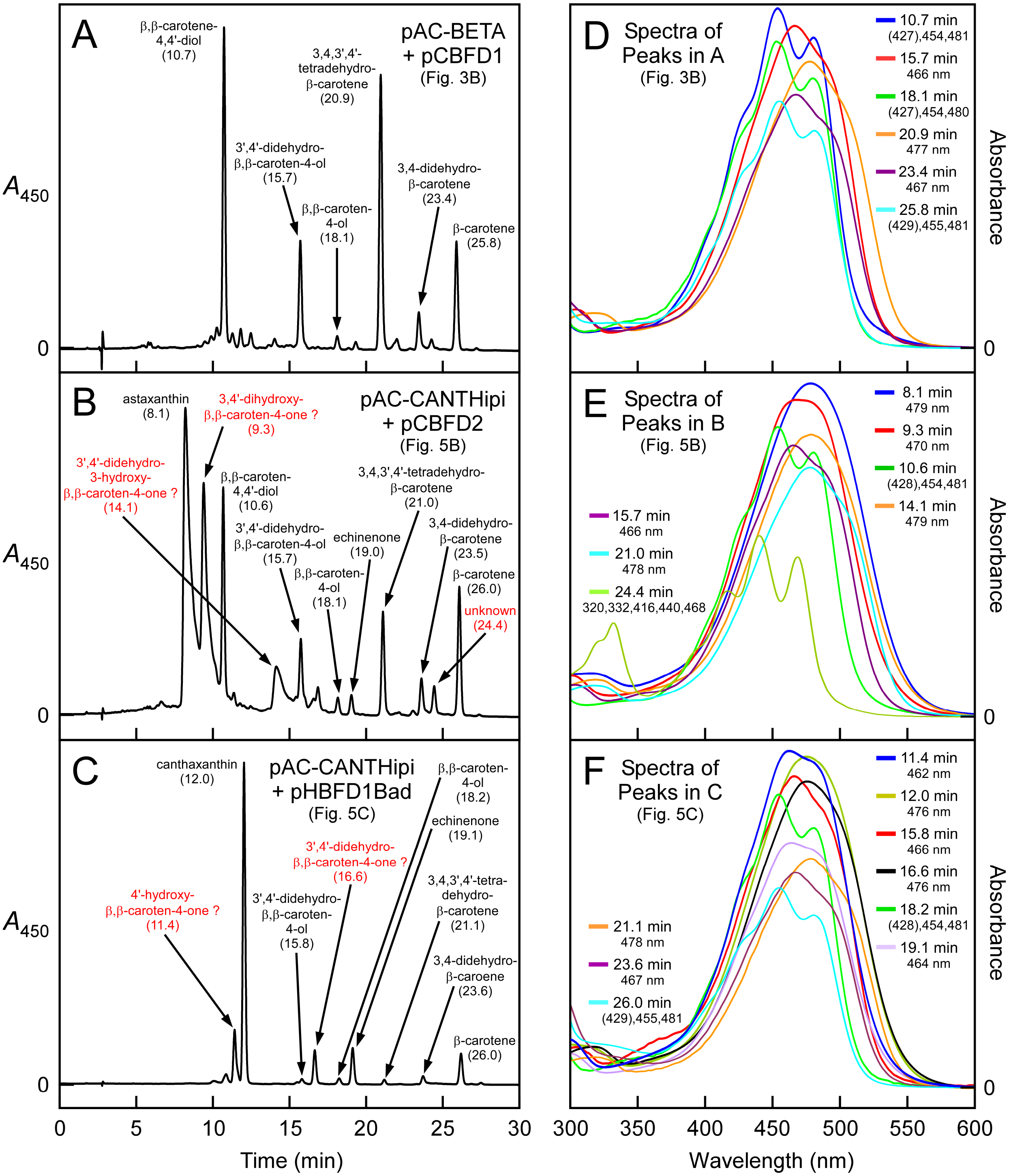
References for Supplemental Figure 6:

Kamata, T., and Simpson, K.L. (1987) Study of astaxanthin diester extracted from *Adonis aestivalis*. Comp. Biochem. Physiol. B **86:** 587-591.

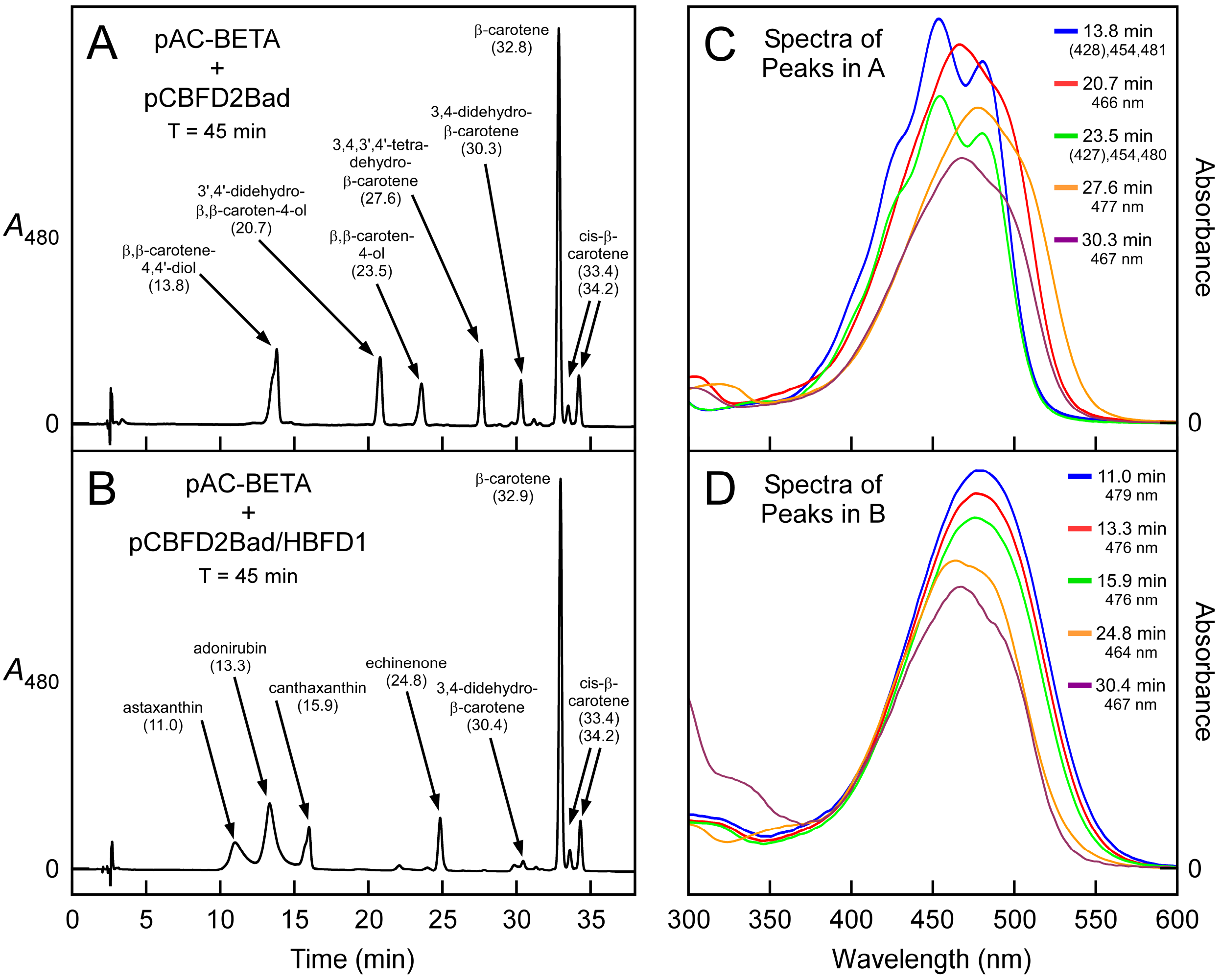
Renstrøm, B., Berger, H., and Liaaen-Jensen, S. (1981) Esterified, optically pure (3*S*, 3'*S*)-astaxanthin from flowers of *Adonis annua*. Biochem. Syst. Ecol. **9:** 249-250.


A
pAC-BETAipi
(Fig. 3A)
β-carotene
(26.0)
cis-β-carotene
(26.7)
(27.3)
A450
0
B
astaxanthin
(8.2)
pAC-BETAipi
+
pCBFD1/HBFD1Bad
(Fig. 3D)
A450
0
C
astaxanthin
(8.2)
Astaxanthin standard
(Fig. 3E)
A450
0
D
canthaxanthin
(12.1)
pAC-CANTHipi
(Fig. 5A)
cis isomers of canthaxanthin
(11.0, 11.6)
echinenone
(19.1)
β-carotene
(26.1)
A450
0
0
5
10
15
20
25
30
Retention Time (min)
E
Spectra of Peaks in A
(Fig. 3A)
26.0 min
(429),455,482
26.7 min
(424),450,476
27.3 min
342,(424),
448,474
Absorbance
0
F
Spectra of Peaks in B and C
(Figs. 3D and 3E)
8.2 min
(panel B) 479 nm
8.2 min
(panel C) 479 nm
Absorbance
0
G
Spectra of Peaks in D
(Fig. 5A)
11.0 min
464 nm
11.6 min
365,466 nm
12.1 min
476 nm
19.1 min
464 nm
26.1 min
(427),455,481
Absorbance
0
300
350
400
450
500
550
600
Wavelength (nm)

Supplemental Figure 7 (below). Absorption Spectra for Peaks in the HPLC Elution Profiles Displayed in Figures 3B, 5B, and 5C of the Manuscript. Reproduced here, from figures in the manuscript, are HPLC elution profiles for extracts of *E. coli* cultures wherein cells contained the following plasmids: **(A)**, pAC-BETA and pCBFD1; **(B)**, pAC-CANTHipi and pCBFD2; **(C)**, pAC-CANTHipi and pHBFD1Bad. Absorption spectra for peaks in the elution profiles of (A), (B) and (C) are shown in panels **(D)**, **(E)** and **(F)**, respectively. Numbers in parentheses below the carotenoid names in panels (A), (B) and (C) are HPLC retention times in minutes. Labels for the absorption spectra in (D), (E) and (F) refer to the HPLC retention times for the corresponding peaks in (A), (B) and (C). Absorption maxima are listed below these retention times, with peak “shoulders” enclosed in parentheses. Spectra were recorded in the HPLC mobile phase just as the individual pigments eluted from the HPLC column. Those peak identifications indicated with red text and followed by question marks in panels (B) and (C) are speculative: they are consistent with the absorption spectra, the HPLC retention times, and the known catalytic capabilities of the carotenoid pathway enzymes that are present in the *E. coli* cells, but the requisite standards were not available for comparison.


A
β,β-carotene-4,4'-diol (10.7)
pAC-BETA + pCBFD1 (Fig. 3B)
3,4,3',4'-tetradehydro-β-carotene (20.9)
3',4'-didehydro-β,β-caroten-4-ol (15.7)
β,β-caroten-4-ol (18.1)
3,4-didehydro-β-carotene (23.4)
β-carotene (25.8)
A_{450}
0
B
astaxanthin (8.1)
3,4'-dihydroxy-β,β-caroten-4-one ? (9.3)
pAC-CANTHipi + pCBFD2 (Fig. 5B)
3',4'-didehydro-3-hydroxy-β,β-caroten-4-one ? (14.1)
β,β-caroten-4,4'-diol (10.6)
3,4,3',4'-tetradehydro-β-carotene (21.0)
3',4'-didehydro-β,β-caroten-4-ol (15.7)
echinenone (19.0)
3,4-didehydro-β-carotene (23.5)
β,β-caroten-4-ol (18.1)
β-carotene (26.0)
unknown (24.4)
A_{450}
0
C
canthaxanthin (12.0)
pAC-CANTHipi + pHBFD1Bad (Fig. 5C)
β,β-caroten-4-ol (18.2)
echinenone (19.1)
3',4'-didehydro-β,β-caroten-4-one ? (16.6)
4'-hydroxy-β,β-caroten-4-one ? (11.4)
3',4'-didehydro-β,β-caroten-4-ol (15.8)
3,4,3',4'-tetra-dehydro-β-carotene (21.1)
3,4-didehydro-β-caroene (23.6)
β-carotene (26.0)
A_{450}
0
0
5
10
15
20
25
30
Time (min)
D
Spectra of Peaks in A (Fig. 3B)
10.7 min (427),454,481
15.7 min 466 nm
18.1 min (427),454,480
20.9 min 477 nm
23.4 min 467 nm
25.8 min (429),455,481
Absorbance
0
E
Spectra of Peaks in B (Fig. 5B)
8.1 min 479 nm
9.3 min 470 nm
10.6 min (428),454,481
14.1 min 479 nm
15.7 min 466 nm
21.0 min 478 nm
24.4 min 320,332,416,440,468
Absorbance
0
F
Spectra of Peaks in C (Fig. 5C)
11.4 min 462 nm
12.0 min 476 nm
15.8 min 466 nm
16.6 min 476 nm
18.2 min (428),454,481
19.1 min 464 nm
21.1 min 478 nm
23.6 min 467 nm
26.0 min (429),455,481
Absorbance
0
300
350
400
450
500
550
600
Wavelength (nm)

Supplemental Figure 8 (below). HPLC Elution Profiles, with Absorption Spectra of Selected Peaks, for Two Pigment Extracts that Provided Data for the Time Course Studies Shown in Figure 4 of the Manuscript. Panels **(A)** and **(B)** show HPLC elution profiles. Panels **(C)** and **(D)** display absorption spectra for selected peaks in (A) and (B). Numbers in parentheses below the carotenoid names in panels (A) and (B) are HPLC retention times in minutes. Labels for the absorption spectra in (C) and (D) refer to the HPLC retention times for the corresponding peaks in (A) and (B). Absorption maxima are listed below these retention times, with peak "shoulders" enclosed in parentheses. Spectra were recorded in the HPLC mobile phase just as the individual pigments eluted from the HPLC column. Samples were taken 45 min after the addition of arabinose to induce production of the enzyme encoded by *cbfd2*. **Note:** a somewhat different mobile phase gradient was used here (4 to 50% mobile phase B in mobile phase A over the course of 35 min; other analyses employed a gradient of 10 to 50% A in B over 30 min) in order to achieve a near baseline separation for all pigments of interest. The various intermediates and end products can be seen to be well separated in the HPLC system employed. Notice that peaks for carotenoids with 3-hydroxy-4-keto-β-rings [*i.e.* astaxanthin and adonirubin (3-hydroxy-β,β-carotene-4,4'-dione)] tailed much more than did other peaks.


A
pAC-BETA
+
pCBFD2Bad
T = 45 min
β-carotene (32.8)
3,4-didehydro-β-carotene (30.3)
3,4,3',4'-tetra-dehydro-β-carotene (27.6)
3',4'-didehydro-β,β-caroten-4-ol (20.7)
β,β-caroten-4-ol (23.5)
β,β-carotene-4,4'-diol (13.8)
cis-β-carotene (33.4) (34.2)
A480
0
B
pAC-BETA
+
pCBFD2Bad/HBFD1
T = 45 min
β-carotene (32.9)
adonirubin (13.3)
echinenone (24.8)
canthaxanthin (15.9)
3,4-didehydro-β-carotene (30.4)
astaxanthin (11.0)
cis-β-carotene (33.4) (34.2)
A480
0
0 5 10 15 20 25 30 35
Time (min)
C
Spectra of Peaks in A
13.8 min (428),454,481
20.7 min 466 nm
23.5 min (427),454,480
27.6 min 477 nm
30.3 min 467 nm
Absorbance
0
D
Spectra of Peaks in B
11.0 min 479 nm
13.3 min 476 nm
15.9 min 476 nm
24.8 min 464 nm
30.4 min 467 nm
Absorbance
0
300 350 400 450 500 550 600
Wavelength (nm)

Supplemental Table 1 (below) lists some of the many cDNAs that were selected in a color complementation screen (see Figure 2B of the manuscript) of an *A. aestivalis* flower library in *E. coli* that contained the plasmid pAC-BETA-CBFD1/2. With the exception of those containing *AdKC17* (*hbfd1*) and *AdKC28* (*hbfd2*), plasmids containing the selected cDNAs had little effect, if any, on the composition of the carotenoids that accumulated in *E. coli* containing plasmid pAC-BETA-CBFD1/2, but rather increased the total amount of pigment, or perhaps only affected the distribution of the pigment within the cell, so that the colonies that contained them appeared a darker orange in color than the preponderance of colonies on agar plates.

Certain of the cDNAs in library plasmids recovered from the selected colonies encoded enzymes of the isoprenoid pathway (isopentenyl diphosphate isomerase and 1-deoxy-*D*-xylulose-5-phosphate synthase) that produces the substrates required for carotenoid biosynthesis. Such cDNAs were expected, and are routinely selected in screens of plant cDNA libraries for cDNAs that enhance carotenoid accumulation in *E. coli* (Cunningham and Gantt, 2007). Very much unexpected were the many and various cDNAs that encoded polypeptides with no apparent relevance to carotenoid biosynthesis. We conjecture that the products of these cDNAs either affect the distribution of carotenoid pigments within the cell (thereby giving the appearance of increased pigment content), or enhance the accumulation of carotenoids by somehow providing or increasing the size of a “sink”. The enhancement (or apparent enhancement) of carotenoid accumulation engendered by these cDNAs appears, in many cases, to be dependent on the structure of the carotenoid molecule. Many of these same cDNAs were selected when screening various other plant cDNA libraries in *E. coli* engineered to accumulate zeaxanthin, but they were not selected when screens were carried out using *E. coli* engineered to accumulate the much less polar carotenoids lycopene and β-carotene (F.X. Cunningham, Jr. and E. Gantt, unpublished).

Supplemental Table 1. *Adonis aestivalis* cDNAs Selected in a Color Complementation Screen[a] Carried out Using *E. coli* Containing the Plasmid pAC-BETA-CBFD1/2 as Host

cDNA	GenBank Accession	Top blastx hit (% identity)[b]	Relevance to carotenoids	Carotenoid composition[c]
AdKC22	AF188060	*A. aestivalis* isopentenyl diphosphate isomerase 1 (100%)	precursor supply	unchanged
AdKC10	AF188061	*A. aestivalis* isopentenyl diphosphate isomerase 2 (100%)	"	"
AdBDY19	EF489295	*Apis mellifera* (honey bee) isopentenyl diphosphate isomerase (57%)	"	"
AdKC13	EF489296	*Ustilago maydis* (corn smut fungus) isopentenyl diphosphate isomerase (76%)	"	"
AdDY7	EF043284	*Chrysanthemum x morifolium* deoxyxylulose 5-phosphate isomerase (85%)	"	"
AdKC18	EU348732	*Vitis vinifera* serine/threonine kinase (94%)	sink ?	unchanged
AdKC24	EU348733	*V. vinifera* ribosomal protein S25 (89%)	"	"
AdKC25	EU348734	*Ricinis communis* iron transport protein 2 (54**%)**	"	"
AdKC26	EU348735	*Pistacia vera* dehydrin (53%)	"	"
AdKC30	EU348736	*Arabidopsis thaliana* unknown protein, mucin-related (65%)	"	"
AdKC31	EU348737	*V. vinifera* ADP-ribosylation factor-like protein (91%)	"	"
AdKC32	EU348738	*V. vinifera* nuclear RNA binding protein (67%)	"	"
AdKC33	EU348739	*Phaseolus vulgaris* polyubiquitin (100%)	"	"
AdKC14	AY644757	*A. aestivalis* carotenoid β-ring 4-dehydrogenase CBFD1 (AdKeto1; 100%)	pathway enzyme	unchanged
AdKC17	DQ902555	*A. thaliana* unknown protein, saccharopine dehydrogenase family (65%)	pathway enzyme	altered
AdKC28	DQ902556	*A. thaliana* unknown protein, saccharopine dehydrogenase family (66%)	pathway enzyme	altered

[a] Colonies of a darker orange color than that exhibited by the preponderance of colonies on agar plates were selected in this screen.

[b] From a search of the GenBank protein database on March 15, 2010.

[c] Cultures inoculated with the selected colonies were harvested in early stationary phase, the pigments were extracted with absolute acetone, and the acetone extracts were analyzed by reverse-phase TLC.

Supplemental Table 2. HPLC Retention Times and Absorption Peaks for Known Carotenoids

Ret. Time (min)[a]	Common Name	Semi-Systematic Name	Absorption Peaks[b] (nm)	Absorption Spectrum
8.0-8.2	astaxanthin	3,3'-dihydroxy-β,β-carotene-4,4’-dione	478	Sup. Fig. 6F
10.6-10.7	isozeaxanthin	β,β-carotene-4,4’-diol	(427), 454, 481	Sup. Fig. 7D
12.0-12.1	canthaxanthin	β,β-carotene-4,4’-dione	476	Sup. Fig. 6G
13.2-13.3	zeaxanthin	β,β-carotene-3,3’-diol	(430), 455, 482	Sup. Fig. 1E
15.7	3’,4’-didehydro-isocryptoxanthin	3’,4’-didehydro-β,β-caroten-4-ol	466, asymmetric peak	Sup. Fig. 7D
18.1-18.2	isocryptoxanthin	β,β-caroten-4-ol	(427), 454, 480	Sup. Fig. 7D
19.1	echinenone	β,β-caroten-4-one	464	Sup. Fig. 6G
20.9-21.1		3,4,3’,4’-tetradehydro-β,β-carotene	477, asymmetric peak	Sup. Fig. 7D
23.4–23.6		3,4-didehydro-β,β-carotene	467, asymmetric peak	Sup. Fig. 7D
25.9–26.1	β-carotene	β,β-carotene	(429), 455, 482	Sup. Fig. 6E

[a] For HPLC with a mobile phase gradient of 10-50% B in A over the course of 30 min.

[b] In the HPLC mobile phase.