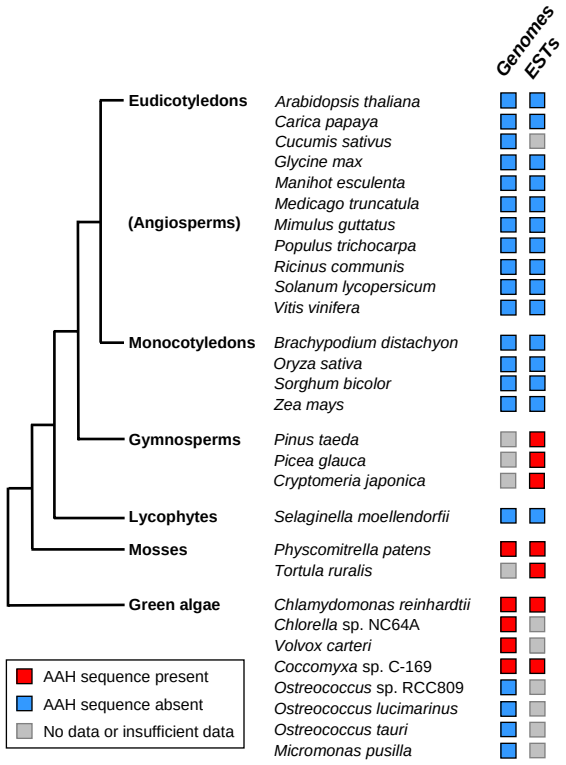


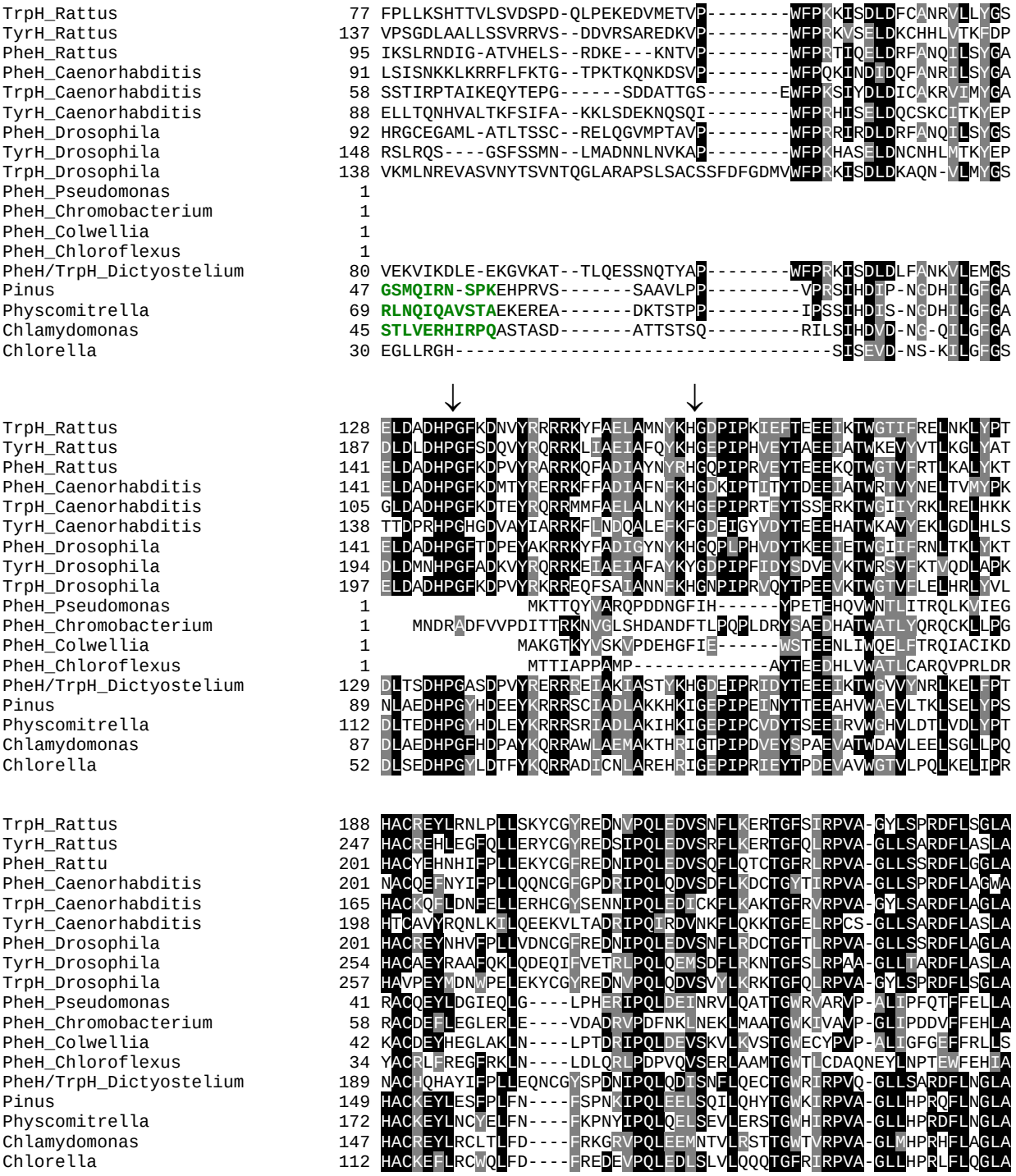
Supplemental Figure 1. The Distribution of AAH-Like Sequences in Plant Genomes and EST Collections.

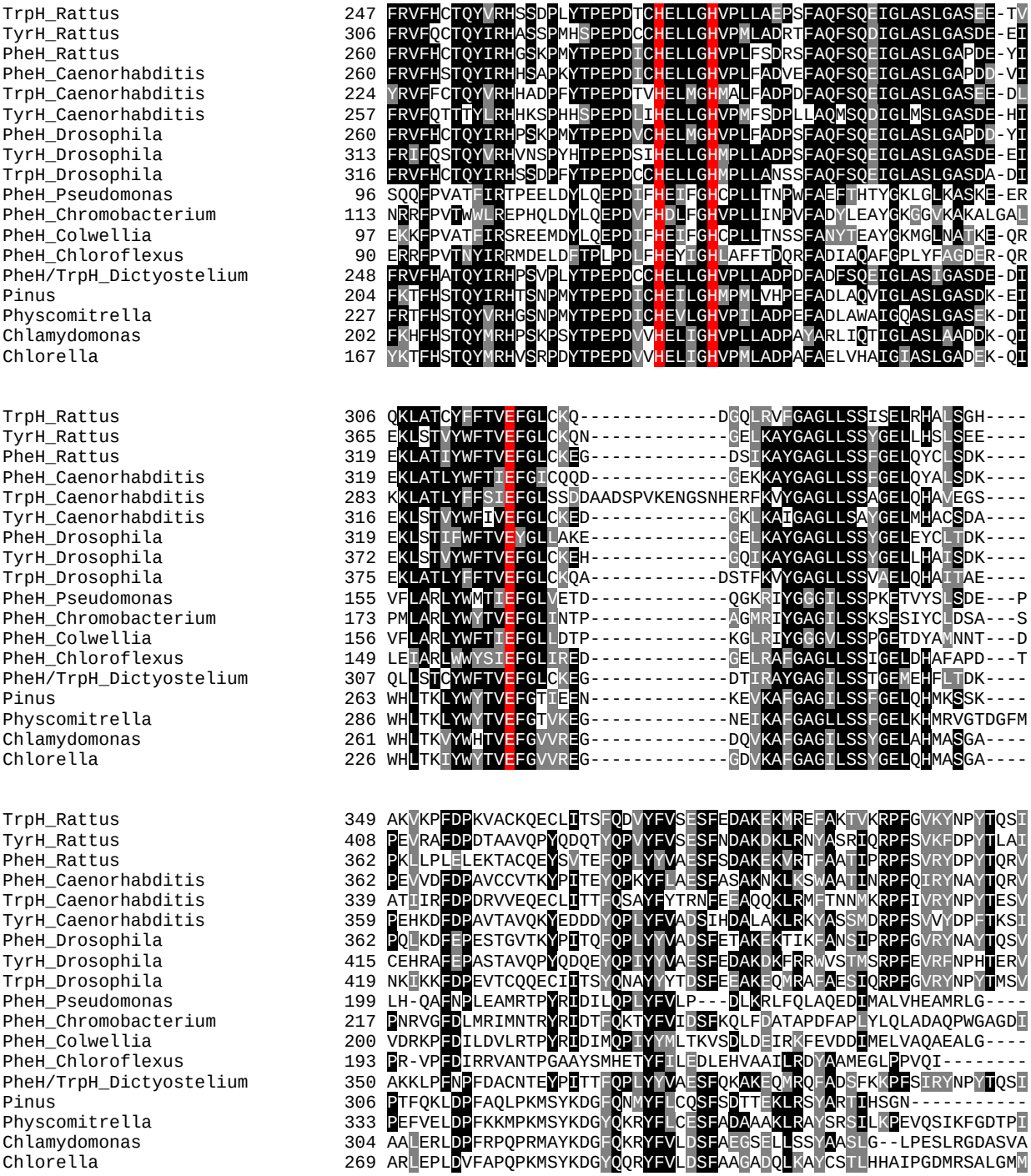
Genomes were searched using tblastn, with eukaryotic or bacterial AAHs as query, at JGI http://genome.jgi-psf.org/, Phytozome http://www.phytozome.net/, and the Sol Genomics Network http://solgenomics.net/index.pl. EST collections were searched at JGI and dbEST http://www.ncbi.nlm.nih.gov/nucest. The 'no data or insufficient data' category refers to species for which no genome sequence is available, or for which $<0.5–1.0 \times 10^5$ EST sequences are available.

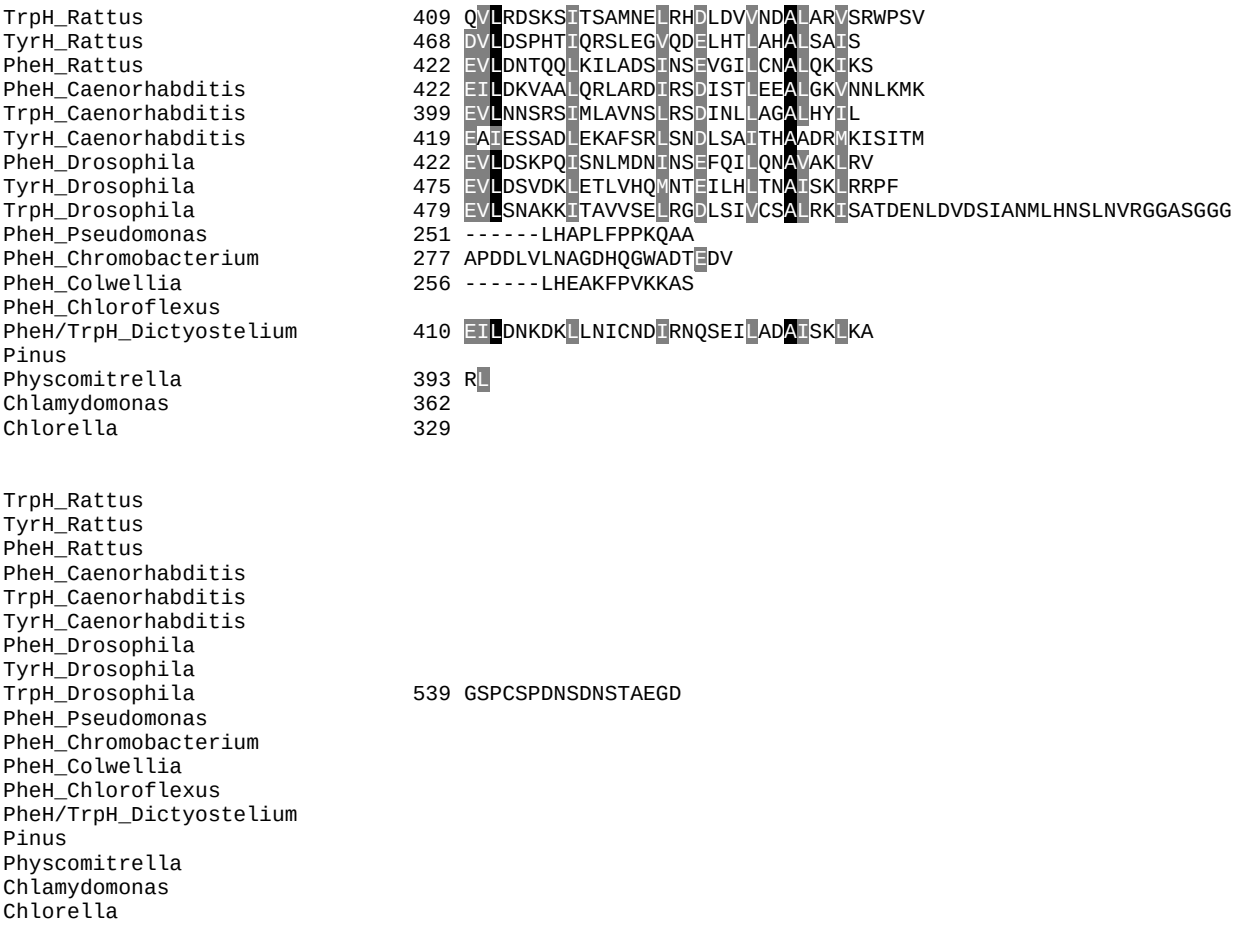
```
TrpH_Rattus                    1
TyrH_Rattus                    1                                MPTPSAPSPQPKGFRRAVSEQDAKQAEAVTSP
PheH_Rattus                    1
PheH_Caenorhabditis            1
TrpH_Caenorhabditis            1
TyrH_Caenorhabditis            1
PheH_Drosophila                1
TyrH_Drosophila                1                                    MMAVAAAQKNREMFAIKKSYSIENGYPS
TrpH_Drosophila                1                        MSASGKSLLGLWLYRSGEQEWAVKQGSPLHQLKKDSTTSG
PheH_Pseudomonas_              1
PheH_Chromobacterium           1
PheH_Colwellia                 1
PheH_Chloroflexus              1
PheH/TrpH_Dictyostelium        1
Pinus                          1
Physcomitrella                 1
Chlamydomonas                  1
Chlorella                      1


TrpH_Rattus                    1                          MIEDNKENKDHSSERGRVTLIFS----------------
TyrH_Rattus                   33 RFIGRRQSLIEDARKEREAAAAAAAAAVASSEPGNPLEAVVFEERDGNAVLNLLFSLRG-
PheH_Rattu                     1   MAAVVLENGVLSRKLSDFGQETSYIEDNSNQNGAIS--LIFS----------------
PheH_Caenorhabditis            1        MPPAGQDDLDFLKYAMESYVADVNADIGKTT--IVFT----------------
TrpH_Caenorhabditis            1                                MIANLPDHTRIKHLET----------------
TyrH_Caenorhabditis            1           MRCQKILQQLNDEGIEVIFTANDVTPIEFSIILTSTDPTLS---------
PheH_Drosophila                1        MYQRQVSFDKPTRVEDSAYIVEGVDIKAARNTCLLFSPRI-------------
TyrH_Drosophila               29 RRRSLVDDARFETLVVKQTKQTVLEEARSKANDYGLTEDEILLANAASESSDAEAAMQSA
TrpH_Drosophila               41 SSSHPSLGRNASAPPEPPRLAIGGGGQDNGRQHSPGERISIIFT----------------
PheH_Pseudomonas               1
PheH_Chromobacterium           1
PheH_Colwellia                 1
PheH_Chloroflexus              1
PheH/TrpH_Dictyostelium        1                        MESNTNSQGQGIIPQSYHSSIFFSISK-------------
Pinus                          1                                  MAFPLQKTFLCS----------------
Physcomitrella                 1                                MAMEVGYLRHSTTIT----------------
Chlamydomonas                  1                                     MLALRQ----------------
Chlorella                      1                                        MGS----------------


TrpH_Rattus                   24 -----LKNEVGGLIKALKIFQENHVNLLHIESRKSKRRNSEFEIFVDCDINREQ--LNDI
TyrH_Rattus                   92 -------TKPSSLSRAVKVFETFEAKIHHLETRPAQRP-------LAGSPHLEY-FVRFE
PheH_Rattus                   41 -----LKEEVGALAKVLRLFEENDINLTHIESRPSRLNKDEY-EFFTYLDKRTKPVLGSI
PheH_Caenorhabditis           36 -----LREKAGALAETLKLFQAHDVNLSHIESRPSRLMKDAMRCSLNLLKLKTIVRLKEL
TrpH_Caenorhabditis           17 -----RDSQDG-SSKTMDVLLE--IELFHYG---------KQEAMDLMRLNGLD--VHEV
TyrH_Caenorhabditis           42 ----------NFVSDILQNMSSAKVQICHVETRGNEAS---HDVLLACKATKNQ-LIHSA
PheH_Drosophila               41 -----RRCPAELLANILAILRKHDINLCILSRILAPWFRVSS-CFWRRMENRS---LGKS
TyrH_Drosophila               89 ALVVRLKEGISSLGRILKAIETFHGTVQHVESRQSRVEGVDHDVLIKLDMTRGN-LLQLI
TrpH_Drosophila               85 -----LRNQVGNLARALQVFQELGINVLHLELSPLEMATNQADVLVDVECDQRR--LDQV
PheH_Pseudomonas               1
PheH_Chromobacterium           1
PheH_Colwellia                 1
PheH_Chloroflexus              1
PheH/TrpH_Dictyostelium       28 -----GSDKIGGLLEYLEIIKKHNINITRIESRPSKTEKKDY-DFFLDLEYPTE--NNKE
Pinus                         13 -------NGQSFPCS----NGRSTSTLLASDLKFQR---------LNKP------FILRV
Physcomitrella                16 -------NGLCCNCDPKPRGARRVQTRLPGTLCLVKDTFTSSKAKLKKPSQREIFLTSRK
Chlamydomonas                  7 -------GALLLSAR----GGQTTHDNLQLCAGPSRR-----------PRARWISSAPRP
Chlorella                      4 -------CRLAADVD----GGQTGHPTARLDAGLSR----------------------P
```

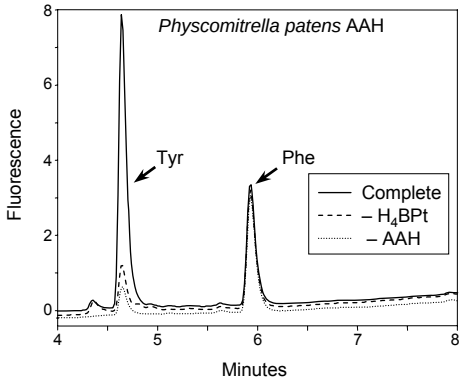

TrpH_Rattus 77 FPLLKSHTTVLSVDSPD-QLPEKEDVMETVP--------WFPKKISDLDFCANRVLLYGS
TyrH_Rattus 137 VPSGDLAALLSSVRRVS--DDVRSAREDKVP--------WFPRKVSELDKCHHLVTKFDP
PheH_Rattus 95 IKSLRNDIG-ATVHELS--RDKE---KNTVP--------WFPRTIQELDRFANQILSYGA
PheH_Caenorhabditis 91 LSISNKKLKRRFLFKTG--TPKTKQNKDSVP--------WFPQKINDIDQFANRILSYGA
TrpH_Caenorhabditis 58 SSTIRPTAIKEQYTEPG------SDDATTGS-------EWFPKSIYDLDICAKRVIMYGA
TyrH_Caenorhabditis 88 ELLTQNHVALTKFSIFA--KKLSDEKNQSQI--------WFPRHISELDQCSKCITKYEP
PheH_Drosophila 92 HRGCEGAML-ATLTSSC--RELQGVMPTAVP--------WFPRRIRDLDRFANQILSYGS
TyrH_Drosophila 148 RSLRQS----GSFSSMN--LMADNNLNVKAP--------WFPKHASELDNCNHLMTKYEP
TrpH_Drosophila 138 VKMLNREVASVNYTSVNTQGLARAPSLSACSSFDFGDMVWFPRKISDLDKAQN-VLMYGS
PheH_Pseudomonas 1
PheH_Chromobacterium 1
PheH_Colwellia 1
PheH_Chloroflexus 1
PheH/TrpH_Dictyostelium 80 VEKVIKDLE-EKGVKAT--TLQESSNQTYAP--------WFPRKISDLDLFANKVLEMGS
Pinus 47 GSMQIRN-SPKEHPRVS-------SAAVLPP---------VPRSIHDIP-NGDHILGFGA
Physcomitrella 69 RLNQIQAVSTAEKEREA-------DKTSTPP---------IPSSIHDIS-NGDHILGFGA
Chlamydomonas 45 STLVERHIRPQASTASD-------ATTSTSQ---------RILSIHDVD-NG-QILGFGA
Chlorella 30 EGLLRGH-----------------------------------SISEVD-NS-KILGFGS
TrpH_Rattus 128 ELDADHPGFKDNVYRRRRKYFAELAMNYKHGDPIPKIEFTEEEIKTWGTIFRELNKLYPT
TyrH_Rattus 187 DLDLDHPGFSDQVYRQRRKLIAEIAFQYKHGEPIPHVEYTAEEIATWKEVYVTLKGLYAT
PheH_Rattus 141 ELDADHPGFKDPVYRARRKQFADIAYNYRHGQPIPRVEYTEEEKQTWGTVFRTLKALYKT
PheH_Caenorhabditis 141 ELDADHPGFKDMTYRERRKFFADIAFNFKHGDKIPTITYTDEEIATWRTVYNELTVMYPK
TrpH_Caenorhabditis 105 GLDADHPGFKDTEYRQRRMMFAELALNYKHGEPIPRTEYTSSERKTWGIIYRKLRELHKK
TyrH_Caenorhabditis 138 TTDPRHPGHGDVAYIARRKFLNDQALEFKFGDEIGYVDYTEEEHATWKAVYEKLGDLHLS
PheH_Drosophila 141 ELDADHPGFTDPEYAKRRKYFADIGYNYKHGQPLPHVDYTKEETETWGIIFRNLTKLYKT
TyrH_Drosophila 194 DLDMNHPGFADKVYRQRRKEIAEIAFAYKYGDPIPFIDYSDVEVKTWRSVFKTVQDLAPK
TrpH_Drosophila 197 ELDADHPGFKDPVYRKRREQFSALANNFKHGNPIPRVQYTPEEVKTWGTVFLELHRLYVL
PheH_Pseudomonas 1 MKTTQYVARQPDDNGFIH------YPETEHQVWNTLITRQLKVIEG
PheH_Chromobacterium 1 MNDRADFVVPDITTRKNVGLSHDANDFTLPQPLDRYSAEDHATWATLYQRQCKLLPG
PheH_Colwellia 1 MAKGTKYVSKVPDEHGFIE------WSTEENLIMQELFTRQIACIKD
PheH_Chloroflexus 1 MTTIAPPAMP------------AYTEEDHLVWATLCARQVPRLDR
PheH/TrpH_Dictyostelium 129 DLTSDHPGASDPVYRERRREIAKIASTYKHGDEIPRIDYTEEEIKTWCVVYNRLKELFPT
Pinus 89 NLAEDHPGYHDEEYKRRRSCIADLAKKHKIGEPIPEINYTTEEAHVWAEVLTKLSELYPS
Physcomitrella 112 DLTEDHPGYHDLEYKRRRSRIADLAKIHKIGEPIPCVDYTSEEIRVWCHVLDTLVDLYPT
Chlamydomonas 87 DLAEDHPGFHDPAYKQRRAWLAEMAKTHRIGTPIPDVEYSPAEVATWDAVLEELSGLLPQ
Chlorella 52 DLSEDHPGYLDTFYKQRRADICNLAREHRIGEPIPRIEYTPDEVAVWGTVLPQLKELIPR
TrpH_Rattus 188 HACREYLRNLPLLSKYCGYREDNVPQLEDVSNFLKERTGFSIRPVA-GYLSPRDFLSGLA
TyrH_Rattus 247 HACREHLEGFQLLERYCGYREDSIPQLEDVSRFLKERTGFQLRPVA-GLLSARDFLASLA
PheH_Rattu 201 HACYEHNHIFPLLEKYCGFREDNIPQLEDVSQFLQTCTGFRLRPVA-GLLSSRDFLGGLA
PheH_Caenorhabditis 201 NACQEFNYIFPLLQQNCGFGPDRIPQLQDVSDFLKDCTGYTIRPVA-GLLSPRDFLACWA
TrpH_Caenorhabditis 165 HACKQFLDNFELLERHCGYSENNIPQLEDICKFLKAKTGFRVRPVA-GYLSARDFLAGLA
TyrH_Caenorhabditis 198 HTCAVYRQNLKILQEEKVLTADRIPQIRDVNKFLQKKTGFELRPCS-GLLSARDFLASLA
PheH_Drosophila 201 HACREYNHVFPLLVDNCGFREDNIPQLEDVSNFLRDCTGFTLRPVA-GLLSSRDFLAGLA
TyrH_Drosophila 254 HACAEYRAAFQKLQDEQIFVETRLPQLQEMSDFLRKNTGFSLRPAA-GLLTARDFLASLA
TrpH_Drosophila 257 HAVPEYMDNWPELEKYCGYREDNVPQLQDVSVYLKRKTGFQLRPVA-GYLSPRDFLSGLA
PheH_Pseudomonas 41 RACQEYLDGIEQLG----LPHERIPQLDEINRVLQATTGWRVARVP-ALIPFQTFFELLA
PheH_Chromobacterium 58 RACDEFLEGLERLE----VDADRVPDFNKLNEKLMAATGWKIVAVP-GLIPDDVFFEHLA
PheH_Colwellia 42 KACDEYHEGLAKLN----LPTDRIPQLDEVSKVLKVSTGWECYPVP-ALIGFGEFFRLLS
PheH_Chloroflexus 34 YACRLFREGFRKLN----LDLQRLPDPVQVSERLAAMTGWTLCDAQNEYLNPTEWFEHIA
PheH/TrpH_Dictyostelium 189 NACHQHAYIFPLLEQNCGYSPDNIPQLQDISNFLQECTGWRIRPVQ-GLLSARDFLNGLA
Pinus 149 HACKEYLESFPLFN----FSPNKIPQLEELSQILQHYTGWKIRPVA-GLLHPRQFLNGLA
Physcomitrella 172 HACKEYLNCYELFN----FKPNYIPQLQELSEVLERSTGWHIRPVA-GLLHPRDFLNGLA
Chlamydomonas 147 HACREYLRCLTLFD----FRKGRVPQLEEMNTVLRSTTGWTVRPVA-GLMHPRHFLAGLA
Chlorella 112 HACKEFLRCWQLFD----FREDEVPQLEDLSLVLQQQTGFRIRPVA-GLLHPRLFLQGLA


TrpH_Rattus 247 FRVFHCTQYVRHSSDPLYTPEPDTCHELLGHVPLLAEPSFAQFSQEIGLASLGASEE-TV
TyrH_Rattus 306 FRVFQCTQYIRHASSPMHSPEPDCCHELLGHVPMLADRTFAQFSQDIGLASLGASDE-EI
PheH_Rattus 260 FRVFHCTQYIRHGSKPMYTPEPDICHELLGHVPLFSDRSFAQFSQEIGLASLGAPDE-YI
PheH_Caenorhabditis 260 FRVFHSTQYIRHHSAPKYTPEPDICHELLGHVPLFADVEFAQFSQEIGLASLGAPDD-VI
TrpH_Caenorhabditis 224 YRVFFCTQYVRHHADPFYTPEPDTVHELMGHMALFADPDFAQFSQEIGLASLGASEE-DL
TyrH_Caenorhabditis 257 FRVFQTTTYLRHHKSPHHSPEPDLIHELLGHVPMFSDPLLAQMSQDIGLMSLGASDE-HI
PheH_Drosophila 260 FRVFHCTQYIRHPSKPMYTPEPDVCHELMGHVPLFADPSFAQFSQEIGLASLGAPDD-YI
TyrH_Drosophila 313 FRIFQSTQYVRHVNSPYHTPEPDSIHELLGHMPLLADPSFAQFSQEIGLASLGASDE-EI
TrpH_Drosophila 316 FRVFHCTQYIRHSSDPFYTPEPDCCHELLGHMPLLANSSFAQFSQEIGLASLGASDA-DI
PheH_Pseudomonas 96 SQQFPVATFIRTPEELDYLQEPDIFHEIFGHCPLLTNPWFAEFTHTYGKLGLKASKE-ER
PheH_Chromobacterium 113 NRRFPVTWWLREPHQLDYLQEPDVFHDLFGHVPLLINPVFADYLEAYGKGGVKAKALGAL
PheH_Colwellia 97 EKKFPVATFIRSREEMDYLQEPDIFHEIFGHCPLLTNSSFANYTEAYGKMGLNATKE-QR
PheH_Chloroflexus 90 ERRFPVTNYIRRMDELDFTPLPDLFHEYIGHLAFFTDQRFADIAQAFGPLYFAGDER-QR
PheH/TrpH_Dictyostelium 248 FRVFHATQYIRHPSVPLYTPEPDCCHELLGHVPLLADPDFADFSQEIGLASIGASDE-DI
Pinus 204 FKTFHSTQYIRHTSNPMYTPEPDICHEILGHMPMLVHPEFADLAQVIGLASLGASDK-EI
Physcomitrella 227 FRTFHSTQYVRHGSNPMYTPEPDICHEVLGHVPILADPEFADLAWAIGQASLGASEK-DI
Chlamydomonas 202 FKHFHSTQYMRHPSKPSYTPEPDVVHELIGHVPLLADPAYARLIQTIGLASLAADDK-QI
Chlorella 167 YKTFHSTQYMRHVSRPDYTPEPDVVHELIGHVPMLADPAFAELVHAIGIASLGADEK-QI
TrpH_Rattus 306 QKLATCYFFTVEFGLCKQ----------DCQLRVFGAGLLSSISELRHALSGH----
TyrH_Rattus 365 EKLSTVYWFTVEFGLCKQN----------GELKAYGAGLLSSYGELLHSLSEE----
PheH_Rattus 319 EKLATIYWFTVEFGLCKEG----------DSIKAYGAGLLSSFGELQYCLSDK----
PheH_Caenorhabditis 319 EKLATLYWFTIEFGICQQD----------GEKKAYGAGLLSSFGELQYALSDK----
TrpH_Caenorhabditis 283 KKLATLYFFSIEFGLSSDDAADSPVKENGSNHERFKVYGAGLLSSAGELQHAVEGS----
TyrH_Caenorhabditis 316 EKLSTVYWFIVEFGLCKED----------GKLKAIGAGLLSAYGELMHACSDA----
PheH_Drosophila 319 EKLSTIFWFTVEYGLLAKE----------GELKAYGAGLLSSYGELEYCLTDK----
TyrH_Drosophila 372 EKLSTVYWFTVEFGLCKEH----------GQIKAYGAGLLSSYGELLHAISDK----
TrpH_Drosophila 375 EKLATLYFFTVEFGLCKQA----------DSTFKVYGAGLLSSVAELQHAITAE----
PheH_Pseudomonas 155 VFLARLYWMTIEFGLVETD----------QGKRIYGGGILSSPKETVYSLSDE---P
PheH_Chromobacterium 173 PMLARLYWYTVEFGLINTP----------AGMRIYGAGILSSKSESIYCLDSA---S
PheH_Colwellia 156 VFLARLYWFTIEFGLLDTP----------KGLRIYGGGVLSSPGETDYAMNNT---D
PheH_Chloroflexus 149 LEIARLWWYSIEFGLIRED----------GELRAFGAGLLSSIGELDHAFAPD---T
PheH/TrpH_Dictyostelium 307 QLLSTCYWFTVEFGLCKEG----------DTIRAYGAGILSSTGEMEHFLTDK----
Pinus 263 WHLTKLYWYTVEFGTIEEN----------KEVKAFGAGILSSFGELQHMKSSK----
Physcomitrella 286 WHLTKLYWYTVEFGTVKEG----------NEIKAFGAGLLSSFGELKHMRVGTDGFM
Chlamydomonas 261 WHLTKVYWHTVEFGVVREG----------DQVKAFGAGILSSYGELAHMASGA----
Chlorella 226 WHLTKIYWYTVEFGVVREG----------GDVKAFGAGILSSYGELQHMASGA----
TrpH_Rattus 349 AKVKPFDPKVACKQECLITSFQDVYFVSESFEDAKEKMREFAKTVKRPFGVKYNPYTQSI
TyrH_Rattus 408 PEVRAFDPDTAAVQPYQDQTYQPVYFVSESFNDAKDKLRNYASRIQRPFSVKFDPYTLAI
PheH_Rattus 362 PKLLPLELEKTACQEYSVTEFQPLYYVAESFSDAKEKVRTFAATIPRPFSVRYDPYTQRV
PheH_Caenorhabditis 362 PEVVDFDPAVCCVTKYPITEYQPKYFLAESFASAKNKLKSWAATINRPFQIRYNAYTQRV
TrpH_Caenorhabditis 339 ATIIRFDPDRVVEQECLITTFQSAYFYTRNFEEAQQKLRMFTNNMKRPFIVRYNPYTESV
TyrH_Caenorhabditis 359 PEHKDFDPAVTAVQKYEDDDYQPLYFVADSIHDALAKLRKYASSMDRPFSVVYDPFTKSI
PheH_Drosophila 362 PQLKDFEPESTGVTKYPITQFQPLYYVADSFETAKEKTIKFANSIPRPFGVRYNAYTQSV
TyrH_Drosophila 415 CEHRAFEPASTAVQPYQDQEYQPIYYVAESFEDAKDKFRRWVSTMSRPFEVRFNPHTERV
TrpH_Drosophila 419 NKIKKFDPEVTCQQECIITSYQNAYYYTDSFEEAKEQMRAFAESIQRPFGVRYNPYTMSV
PheH_Pseudomonas 199 LH-QAFNPLEAMRTPYRIDILQPLYFVLP---DLKRLFQLAQEDIMALVHEAMRLG----
PheH_Chromobacterium 217 PNRVGFDLMRIMNTRYRIDTFQKTYFVIDSFKQLFDATAPDFAPLYLQLADAQPWGAGDI
PheH_Colwellia 200 VDRKPFDILDVLRTPYRIDIMQPIYYMLTKVSDLDEIRKFEVDDIMELVAQAEALG----
PheH_Chloroflexus 193 PR-VPFDIRRVANTPGAAYSMHETYFILEDLEHVAAILRDYAAMEGLPPVQI--------
PheH/TrpH_Dictyostelium 350 AKKLPFNPFDACNTEYPITTFQPLYYVAESFQKAKEQMRQFADSFKKPFSIRYNPYTQSI
Pinus 306 PTFQKLDPFAQLPKMSYKDGFQNMYFLCQSFSDTTEKLRSYARTIHSGN----------
Physcomitrella 333 PEFVELDPFKKMPKMSYKDGYQKRYFLCESFADAAAKLRAYSRSILKPEVQSIKFGDTPI
Chlamydomonas 304 AALERLDPFRPQPRMAYKDGFQKRYFVLDSFAECSELLSSYAASLG--LPESLRGDASVA
Chlorella 269 ARLEPLDVFAPQPKMSYKDGYQQRYFVLDSFAAGADQLKAYCSTLHHAIPGDMRSALGMM



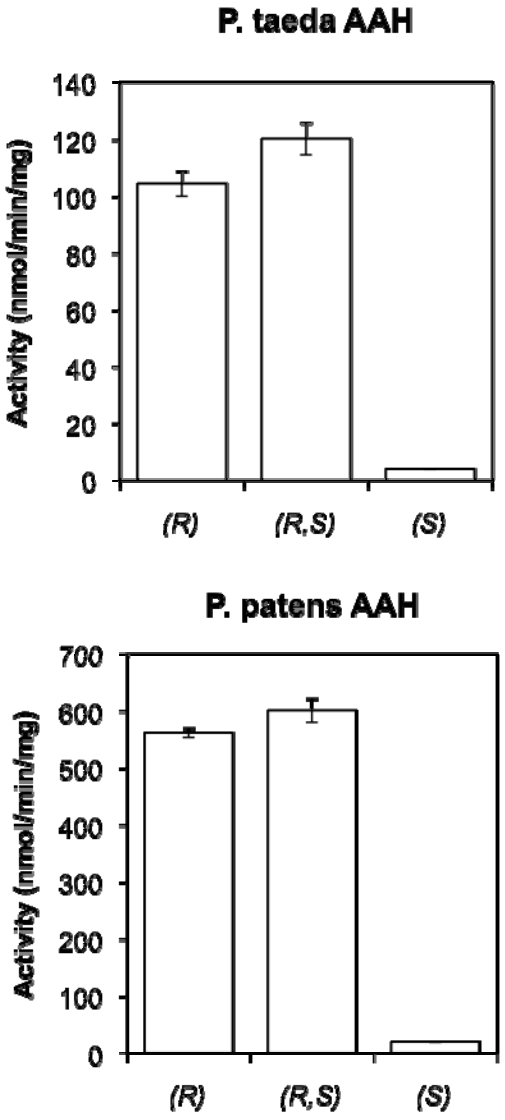
Supplemental Figure 2. Multiple Sequence Alignment of AAH Proteins from Eukaryotes and Prokaryotes.

Alignments were made with ClustalW; residues were shaded with BoxShade (black for identical, gray for similar). Dashes are gaps introduced to maximize alignment. Arrows mark the positions of two introns common to mammalian phenylalanine (PheH), tyrosine (TyrH), and tryptophan (TrpH) hydroxylases and to *Pinus taeda* and *Physcomitrella patens* AAHs. Green font indicates the predicted targeting signals of the plant proteins, which were removed for bacterial complementation and overexpression experiments. Iron-liganding residues are shaded in red. Full organism names: *Caenorhabditis elegans*, *Chlamydomonas reinhardtii*, *Chlorella* sp. NC64A, *Chloroflexus aurantiacus*, *Chromobacterium violaceum*, *Colwellia psychrerythraea*, *Dictyostelium discoideum*, *Drosophila melanogaster*, *Physcomitrella patens*, *Pinus taeda*, *Pseudomonas aeruginosa*, *Rattus rattus*.



Supplemental Figure 3. Phenylalanine Hydroxylase Activity of Recombinant *Physcomitrella patens* AAH.

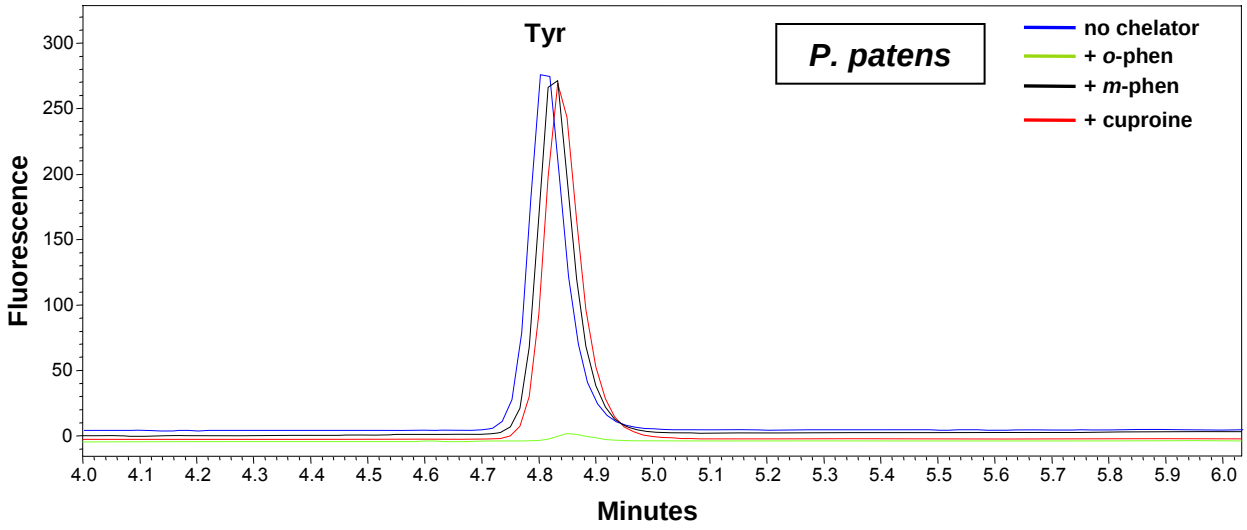
The complete reaction comprised 1 mM phenylalanine, 0.2 mM H_4BPt, 50 µM $Fe(NH_4)_2(SO_4)_2$, 10 mM dithiothreitol, 40 units of catalase, and extract of *E. coli* cells (10 µg protein) harboring pET28b containing *P. patens* AAH cDNA. Controls were run without H_4BPt (–H_4BPt) or with extract from *E. coli* cells harboring pET28b alone (–AAH). Incubation was at 30°C for 15 min. Reactions were analyzed by fluorometric HPLC.



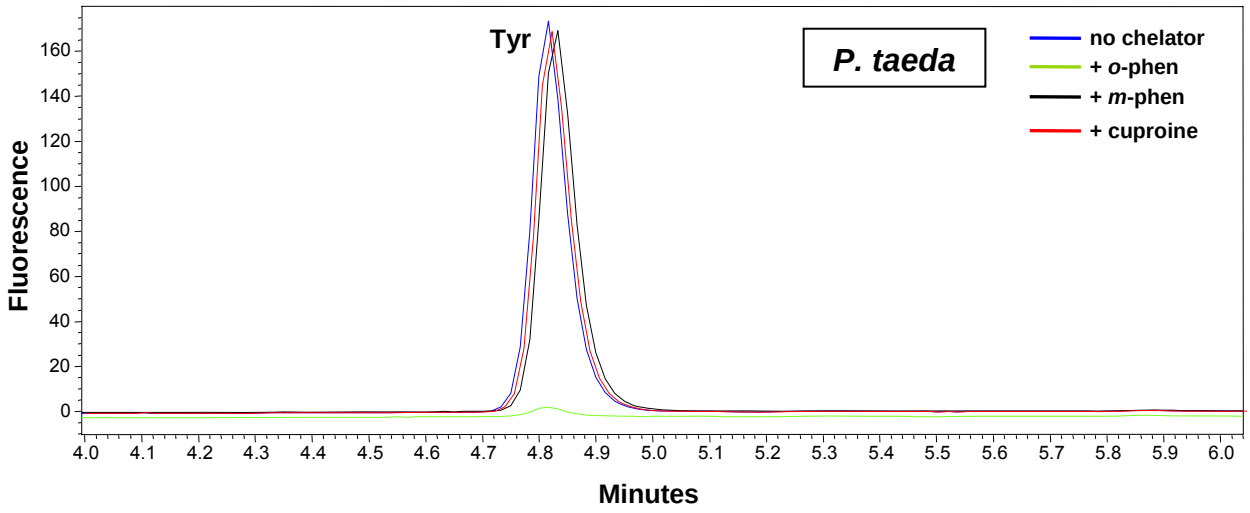
Supplemental Figure 4. Effect of H_4BPt Chirality on Phenylalanine Hydroxylase Activity.

Activities of recombinant AAHs from *P. taeda* and *P. patens* with the natural (*R*) and unnatural (*S*) forms of H4BPt, and a racemic mixture (*R,S*). Cofactor concentrations (μ M) were: *P. taeda*, 200; *P. patens*,100. Activity was assayed as in Figure 5. Data are means and SE of three replicates.

A

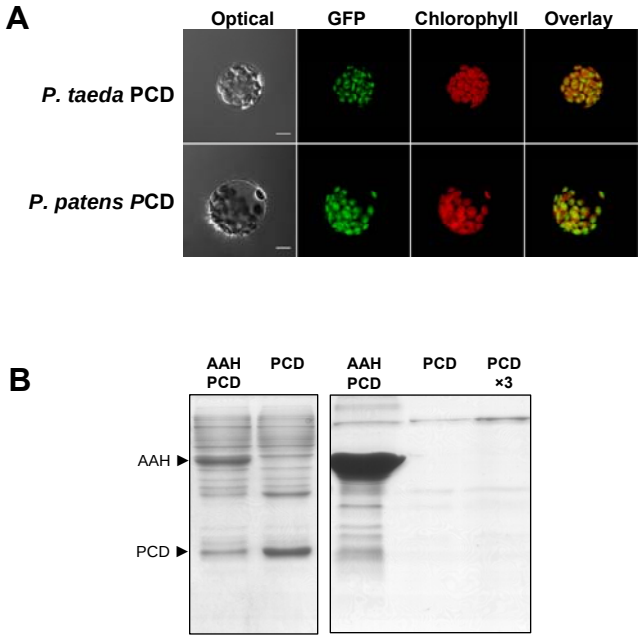


B



Supplemental Figure 5. Evidence That Phenylalanine Hydroxylase Activity Requires Iron.

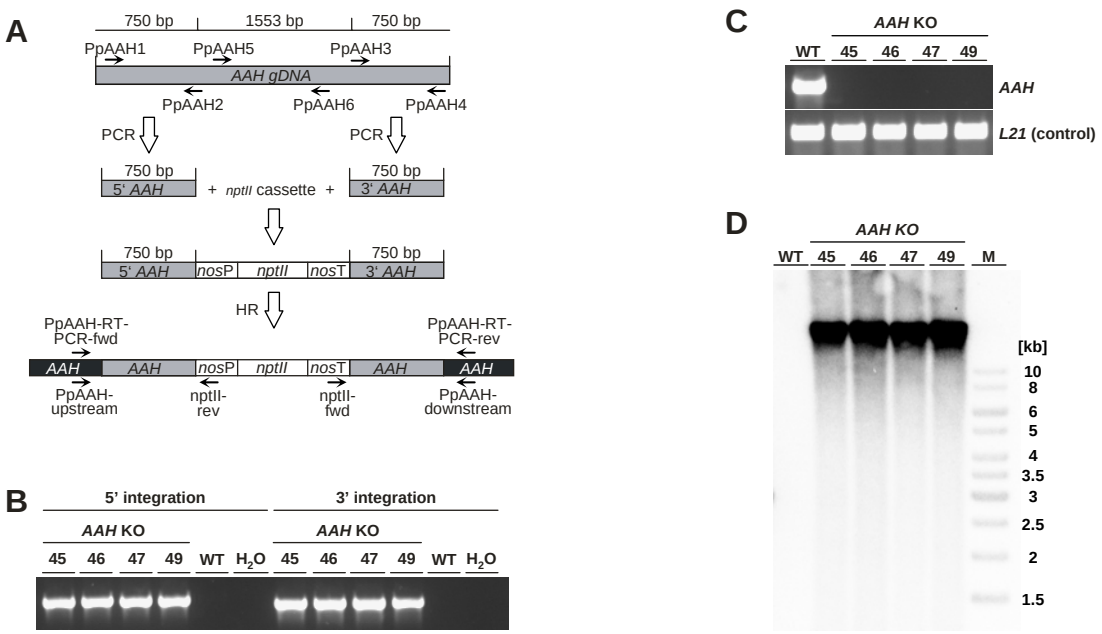
Fluorometric analysis of phenylalanine hydroxylase activity of recombinant *P. patens* **(A)** and *P. taeda* AAH **(B)**. Reactions contained 50 µM chelator (*o*-phen, *o*-phenanthroline; *m*-phen, *m*-phenantholine; cuproine, bathocuproine) or no chelator. Cofactor concentration was 200 µM.



Supplemental Figure 6. Evidence That *P. taeda* and *P. patens* PCD Proteins Are Chloroplast-Targeted and That *P. taeda* PCD and AAH Form a Complex In Vitro.

(A) Transient expression in *Arabidopsis* mesophyll protoplasts of green fluorescent protein (GFP) fused to the C terminus of *P. taeda* PCD (upper panels) or *P. patens* PCD (lower panels). The fusion constructs were as described previousy (Naponelli et al., 2008). GFP (green pseudo-color) and chlorophyll (red pseudo-color) fluorescence were observed by confocal microscopy. Bars = 10 μm.

(B) Complex formation between PCD and AAH. *P. taeda* PCD was expressed in *E. coli* together with *P. taeda* AAH, or alone as a control. The left panel shows SDS-PAGE separations of whole cell protein extracts (10 μg per lane) stained with Coomassie blue. The extracts were applied to Ni-affinity columns, which were washed exhaustively with buffer containing 50 mM imidazole, and then eluted with buffer containing 250 mM imidazole. The right panel shows separations of equivalent amounts of eluate from each sample, plus a three-fold higher loading of the PCD-alone control extract, stained with silver. Note that a PCD band is present only in the sample also containing AAH.



Supplemental Figure 7. Generation and Quality Controls for Four Independent *P. patens* AAH Knockout Lines.

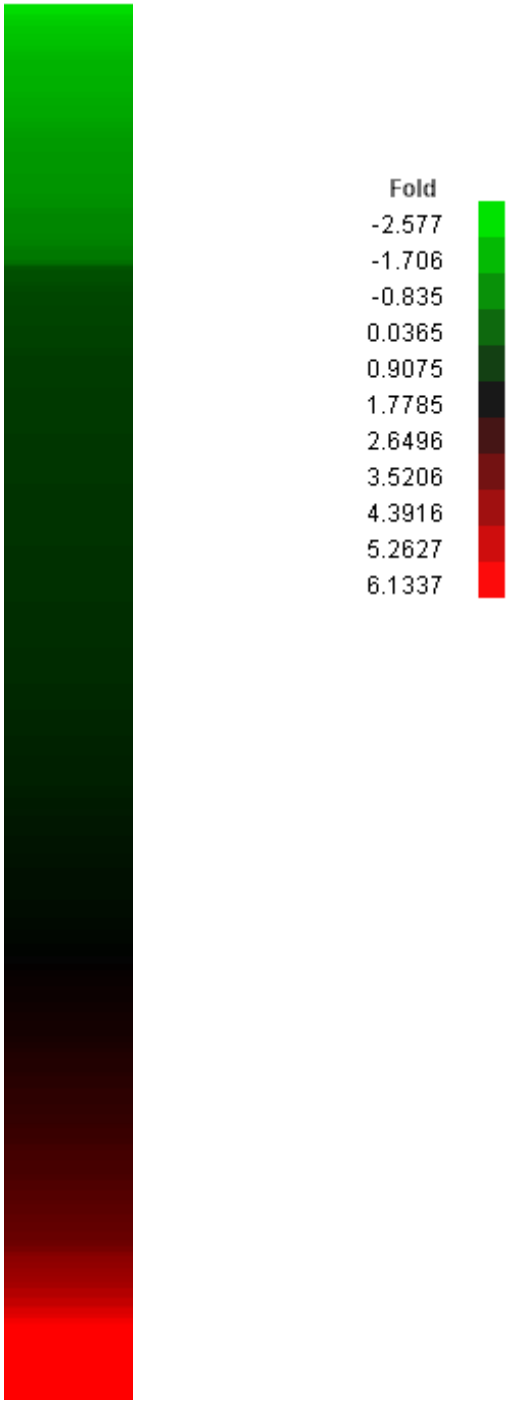
(A) Schematic illustration of *AAH* knockout generation. Upper: 750-bp fragments corresponding to the 5'- and 3'-flanking regions of *AAH* were amplified with the primers PpAAH1 and PpAAH2, and PpAAH3 and PpAAH4. The *AAH* knockout construct was generated by successive cloning of an *nptII* selection marker cassette in between the *AAH* 5'- and 3'-flanking sequences. Bottom: Expected genomic structure of the *AAH* locus after integration of the *AAH* knockout construct by homologous recombination (HR). Arrows indicate primers used for molecular analysis of the transgenic lines. White box: *nptII* cassette; grey boxes: *AAH* gDNA fragments; black boxes: genomic *AAH* locus.

(B) PCR analysis with genomic DNA from four transgenic lines (*AAH* KO 45-49) and wild type (WT) to confirm 5' (PpAAH-upstream/nptII-rev primers) and 3' (PpAAH-down-stream/nptII-fwd primers) integration of the *AAH* knockout construct. H_2O: PCR water control without DNA.

(C) RT-PCR from four *AAH* KO and wild type (WT) with PpAAH-RT-PCR-fwd/PpAAH-RT-PCR-rev primers spanning the integrated *AAH* knockout construct and primers for the *L21* control to monitor efficient cDNA synthesis. *AAH* and *L21* reactions were performed in parallel.

(D) Southern blot with genomic DNA from four *AAH* KO mutants and wild type digested with *Bsm*I, which does not cut within the knockout construct. An *nptII* fragment was used as hybridization probe. The blot does not show the predicted 4.5-kb band, but a shifted band due to the concatameric integration of knockout constructs, which is common in *P. patens*.; M: Molecular weight markers.

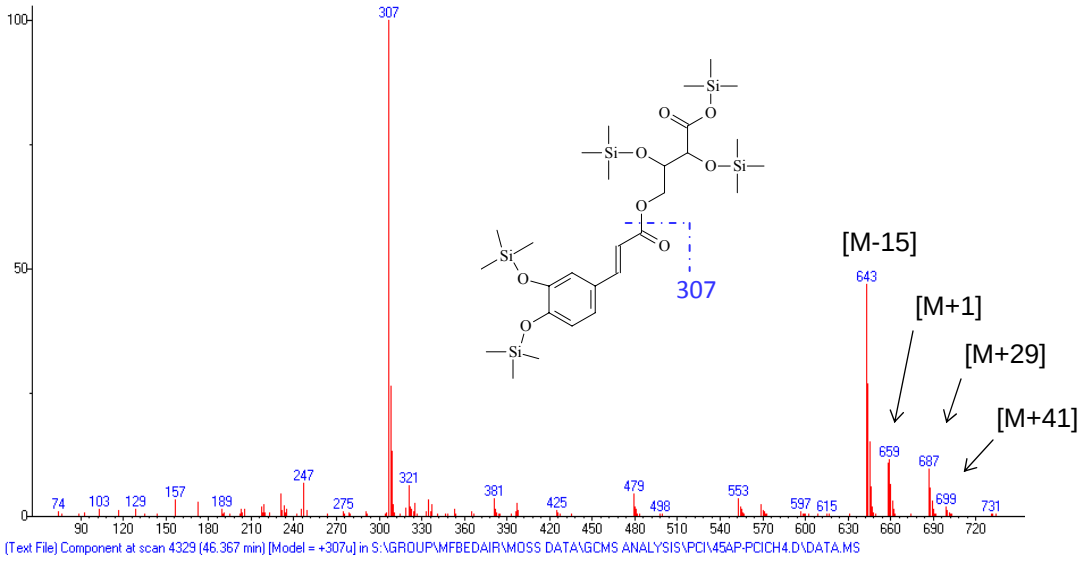
Sequences of primers used for molecular analyses are listed in Supplemental Table 2 online.



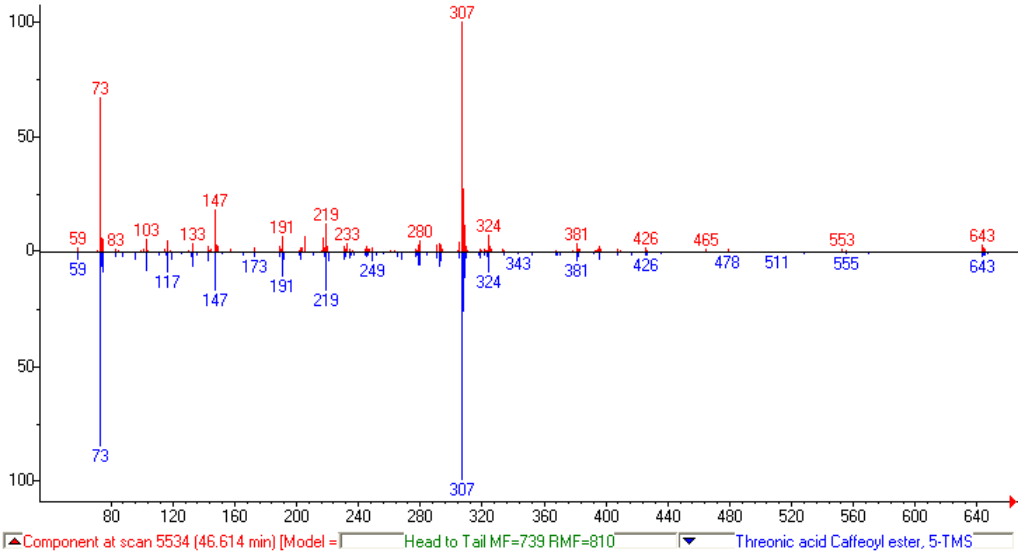
Supplemental Figure 8. Heat Map Depicting the Changes in Metabolite Levels that Characterize *P. patens AAH* Knockout Lines.

Signal intensities for each metabolite were scaled to an internal control and then $\log_2$ transformed; values were subjected to a mixed model ANOVA with lines as a fixed effect. Contrasts (i.e., fold difference) between knockout and wild-type lines are presented, with green indicating compounds downregulated in the knockout lines and red indicating compounds upregulated in the knockout lines. There are a total of 425 polar and nonpolar metabolites that are shown after being declared significant, with q-values (reflecting the false discovery rate) ranging from 0.0495 to 0.0021.

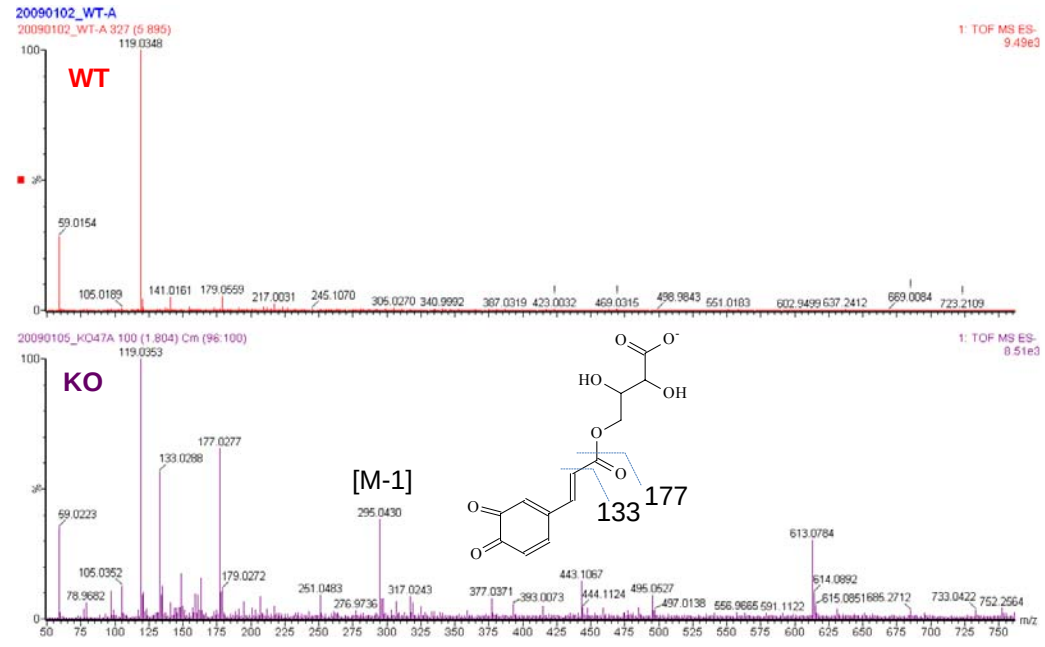
A


307
[M-15]
643
[M+1]
[M+29]
[M+41]
659
687
699
731
307
(Text File) Component at scan 4329 (46.367 min) [Model = +307u] in S:\GROUP\MFBEDAIR\MOSS DATA\GCMS ANALYSIS\PCI\45AP-PCICH4.D\DATA.MS

B


307
73
147
191
219
233
280
324
381
426
465
553
643
Component at scan 5534 (46.614 min) [Model =
Head to Tail MF=739 RMF=810
Threonic acid Caffeoyl ester, 5-TMS

C


20090102_WT-A
WT
KO
[M-1]
295.0430
177
133

Supplemental Figure 9. Mass Spectral Evidence for the Identity of Caffeic Acid Esters in *P. patens*.

(A) Chemical ionization spectra of caffeoyl ester at 46.62 min using methane as reagent gas. Molecular ions peaks $[M+1]^+$, $[M+29]^+$ and $[M+41]^+$ are the $[M+H]^+$, $[M+C_2H_5]^+$, and $[M+C_3H_5]^+$ respectively. The ion $[M-15]^+$ is the $[M-CH_3]^+$.

(B) Electron ionization spectra of caffeoyl ester at 46.62 min in AAH knockout (KO) (red) and of synthesized caffeoylthreonic acid standard (blue).

(C) Liquid chromatography quadrupole time of flight (LC-Q-TOF) MS evidence for the presence of caffeoylthreonic acid in the *AAH* KO. Ion $[M-1]^-$ is the $[M-H]^-$.

Three unknown peaks at 46.04, 46.62 and 47.76 min were found to be four- to five-fold higher in the *AAH* KO compared to the wild-type (WT). The first peak at 46.04 min was weakly identified as 1-trans-caffeoylquinic acid, but this identification was dismissed due to the lower retention index value of the unknown peak. However, the presence of the characteristic caffeoyl ester ions at m/z 307 and 324 and the highest Pearson correlation coefficient with caffeic acid (0.9835 and 0.9788, $p<0.0001$) suggested that these two peaks are esters of caffeic acid with a smaller ester moiety than the C7 quinic acid. The unknown caffeoyl esters were analyzed with chemical ionization using methane as reagent gas; the molecular ion peaks were m/z 659, 687 and 699, i.e., M+H, $M+C_2H_5$ and $M+C_3H_5$, respectively (Supplemental Figure 9A). Thus the nominal mass of the trimethylsilyl derivative of the two caffeoyl esters was found to be 658. The isotopic ratio of the molecular ion peak suggested the presence of five trimethylsilyl groups, thus indicating a nominal mass of 298 for the underivatized molecules. Searching the plant database KNApSAcK (http://kanaya.naist.jp/KNApSAcK/) suggested three possibilities: caffeoyltrihydroxybutyric acid, caffeoyldexoyribitol or caffeoylmethylerythritol. Analyzing *AAH* KO and WT samples using Q-TOF UPLC-MS showed the presence of two peaks that eluted at 1.4 and 1.7 min and were three-fold higher in the KO (Supplemental Figure 9C). The accurate masses of these peaks were within 5 ppm of the exact mass of caffeoyltrihydroxybutyric acid but within 120 ppm of the other two possibilities, the deoxyribitol or methylerythritol esters of caffeic acid. The identification of the unknown caffeoyl ester peaks as caffeoyltrihydroxybutyric acid was further confirmed by synthesis of caffeoylthreonic acid. The GCMS analysis of the synthesized caffeoylthreonic acid matched the peak at 46.62 min as can be seen from Supplemental Figure 9B and was within a retention index difference of 15.

Supplemental Table 1. Separation and Detection of Potential Products of AAH Action Using Fluorometric HPLC

Substrates[a]		Potential products		Fluorescence wavelengths[b] (nm)	Minimum detectable product formation		
Name	RT[c] (min)	Name	RT (min)		pmol per assay	% of *P. taeda* Phe[d]	% of *P. patens* Phe[e]
L-Phenylalanine	6.03	L-Tyrosine	4.71	265/315	0.5	0.06	0.02
D-Phenylalanine	6.05	D-Tyrosine	4.73	265/315	0.5	0.06	0.02
L-Tyrosine	4.98	L-3,4-Dihydroxyphenylalanine	4.05	265/315	0.6	0.08	0.03
L-Tryptophan	8.15	L-5-Hydroxytryptophan	6.27	265/315	0.2	0.03	0.01
Cinnamate	18.00	*p*-Coumarate	13.40	310/375	13	0.56	0.41
p-Coumarate	13.40	Caffeate	11.40	310/375	26	1.00	0.28
Benzoate	14.60	4-Hydroxybenzoate	10.50	250/330	4	0.51	0.13
Anthranilate	12.94	3-Hydroxyanthranilate	8.60	315/400	2	0.25	0.06
Mandelate	7.52	4-Hydroxymandelate	4.40	250/330	17	0.74	0.19
Indoleacetate	15.16	5-Hydroxyindoleacetate	9.94	280/350	0.2	0.003	0.0003
Caffeate	11.40	5-Hydroxycaffeate[f]	<11.40	310/375	26	0.40	0.04
Ferulate	13.98	5-Hydroxyferulate[f]	<13.98	310/375	9	0.14	0.02

[a]Substrate concentration was 15 mM for indoleacetate, caffeate, and ferulate, 1 mM in other cases. L-Phenylalanine was tested at both 15 and 1 mM.

[b]Excitation/emission.

[c]RT, retention time. Gradient 1 was used with phenylalanine, tyrosine, and tryptophan substrates; gradient 2 was used with all other substrates.

[d]Minimum detectable product formation expressed as a percentage of that formed by *P. taeda* AAH with the same concentration of L-phenylalanine as substrate (1 or 15 mM) and the same amount of protein per assay (2.9–8.6 µg). Activities with 1 and 15 mM L-phenylalanine were 27 and 75.5 nmol min^{-1} mg^{-1} protein, respectively.

[e]Minimum detectable product formation expressed as a percentage of that formed by *P. patens* AAH with the same concentration of L-phenylalanine as substrate (1 or 15 mM) and the same amount of protein per assay (2.8–7.7 µg). Activities with 1 and 15 mM L-phenylalanine were 107 and 690 nmol min^{-1} mg^{-1} protein, respectively.

[f]Authentic standards not available; fluorescence excitation and emission wavelengths and detection limit based on the substrate.

Supplemental Table 2. Primers Used in this Study

Primer function and primer name	Primer direction[a]	Primer sequence (5′→3′)[b]
Pinus taeda cDNA cloning		
PtAAH-GSP1	R	CAGTATGCCAGCTCCAAATGCCTTAAC
PtAAH-GSP2	R	GAGCAAGATCAGCAAACTCAGGGTGTAC
PtAAH-GSP3	F	GTACACCCTGAGTTTGCTGATCTTGCTC
PtAAH-GSP4	F	GTTAAGGCATTTGGAGCTGGCATACTG
PtAAH7	R	TGTGTTTCTTGGCCAGGTCAGCA
PtAAH8	R	TTCTCTTGTATTCTTCATCATGGTATCC
pSL1	F	GCTTAAGGTGCACGGCCCA
pSL2	R	AAAACGACGGCCAGTGCCAAG
Subcellular localization		
PpAAH-GFP-fwd	F	ATGATCGTCGACATGGCTATGGAAGTGGGTTATC
PpAAH-GFP-rev	R	CACCATGGTGAGCCGGATTGGCGTATC
PtAAH-4z-fwd	F	CAGTCAGAATTCACCATGGCGTTTCCACTCCAG
PtAAH-4z-rev	R	CAGTCACTGCAGTTAATTACCAGAGTGAATAGTTC
PpAAH-4z-fwd	F	CAGTCAGAATTCACCATGGCTATGGAAGTGGGTTATC
PpAAH-4z-rev	R	CAGTCACTGCAGTTAGAGCCGGATTGGCGTATC
Recombinant protein expression		
PtAAH-expfwd	F	CAGTCACCATGGAACACCCAAGAGTGAGCAG
PtAAH-exprev	R	GATGCTCTCGAGATTACCAGAGTGAATAGTTCTTGCATA
PpAAH-expfwd	F	CAGTCACCATGGAAAAGGAGAGAGAGGCAGA
PpAAH-exprev	R	CAGTCACTCGAGGAGCCGGATTGGCGTATCT
Complex formation		
PtPCD-pet43fwd	F	GGGAATTCCATATGAAATGTTCTCAAGCTAATG
PtPCD-pet43rev	R	AGCTGAATTCTTAGTTCTTTGCAGCTTTC
PtPCD-rbsfwd	F	CAGTCATCTAGAAATAATTTTGTTTAACTTTAAG
PtPCD-rbsrev	R	CAGTCATCTAGATTAGTTCTTTGCAGCTTTCTTTC
Functionnal complementation		
PaAAH-pSU18fwd	F	CATGGAATTCCTGAAGGAAACAGCTATGAAAACGACGCAGTACGTG
PaAAH-pSU18rev	R	ACGTACGCATGCTCAGGCCGCCTGCTTGGG
PtAAH-pSU18fwd	F	CATGGAATTCCTGAAGGAAACAGCTATGGAACACCCAAGAGTGA
PtAAH-pSU18rev	R	ACGTACGCATGCTTAATTACCAGAGTGAATAGTTCTT
PpAAH-pSU18fwd	F	CATGGAATTCCTGAAGGAAACAGCTATGAAGGAGAGAGAGGCAGACAA
PpAAH-pSU18rev	R	ACGTACGCATGCTTAGAGCCGGATTGGCGTAT
CrAAH-pSU18fwd	F	CATGGAATTCCTGAAGGAAACAGCTATGTCGACGGCCTCTGACGC
CrAAH-pSU18rev	R	ACGTACGCATGCTCACGCCACGCTGGCGTC
Moss Knockout		
PpAAH1	F	ACGTGGTACCATCCTTAAATGGGGATTGTCG
PpAAH2	R	TCAAGCTTGCGGCCGCATACTCACCACAATACAAACTTTC
NPT1	F	GTGAGTATGCGGCCGCAAGCTTGATATCGGATCCTGT
NPT2	R	ATTATTCACTCGAGGGATCCCCCGGGCTG
PpAAH3	F	GGATCCCTCGAGTGAATAATACAGGAGCAGAAATC
PpAAH4	R	TAATGAGCTCCTTTAATCTCGTTACCTTCCTTTA
PpAAH5	F	CCTATTCCCAGTAGCATCCACG
PpAAH6	R	GATCATGGTAGCCCTGCAAAAG
PpAAH-upstream	F	ATGCCATGTGCATTGTATCC
nptII-rev	R	CCAAACGTAAAACGGCTTGT
PpAAH-downstream	R	GCTTCAGCTCCCCAAAACTA
nptII-fwd	F	AGCGCGCAAACTAGGATAAA
PpAAH-RT-PCR-fwd	F	CCCTGTGTCTCGTCAAGGAT
PpAAH-RT-PCR-rev	R	TGTAGCTCATCTTGGGCATCT
PpL21-fwd	F	GGTTGGTCATGGGTTGCG
PpL21-rev	R	GAGGTCAACTGTCTCGCC

[a] F, forward; R, reverse.
[b] Restriction sites are underlined.

SUPPLEMENTAL METHODS

AAH Ablation in Moss

The knockout construct was designed to remove part of the AAH genomic sequence (*AAH*) and to replace it with the *nptII* selection marker. To avoid polar effects, the deleted region stretched from the second exon of *AAH* to the middle of its last exon, representing about half the coding sequence. The construct was assembled in pBluescript II SK (Stratagene) by inserting on each side of the *nptII* cassette a 750-bp fragment homologous to the 5′ or 3′ flanking region of the planned deletion (see Supplemental Figure 7A online). These fragments were amplified by PCR from *P. patens* genomic DNA with primers PpAAH1/AAH2, and PpAAH3/AAH4, respectively (see Supplemental Table 2 online) and subcloned into pGEM-T (Promega). When internal pGEM-T restriction sites were needed for further cloning purposes, fragments were given the desired orientation by excising the insert with *Not*I, religating the fragments, and selecting the correct orientation by restriction. The construct was assembled stepwise as follows. First, the *nptII* selection marker cassette (*nos*-promoter: neomycin phosphotransferase: *nos*-terminator) amplified from the vector pBSNNNEV (Egener et al., 2002) with primers NPT1/NPT2 was cloned into the *Xho*I and *Not*I sites of pBluescript II SK. The *nptII* gene having *Bam*HI sites at its extremities was given the same orientation as the *AAH* gene on the chromosome by excising the gene with *Bam*HI, religating the fragments, and selecting the correct orientation by restriction. Secondly, the 3′ flanking region of *AAH* was subcloned into pBluescript II SK using an internal *Not*I site from pGEM-T and the *Sac*I site from the primer PpAAH4. Thirdly, the 5′ flanking region having a *Kpn*I site from the primer PpAAH1 and an internal *Apa*I site from pGEM-T was inserted into pBluescript II SK (*Apa*I brought in the vector by the *nptII* fragment). The whole construct was verified by sequencing. The vector backbone and the insert having the same size, the construct was codigested by *Apa*LI (cutting twice the vector), *Kpn*I, and *Sac*I to generate a single 3 kb fragment used for the recombination. This fragment (knockout construct) was purified and transfected into *P. patens* protoplasts. Protoplast isolation, polyethylene glycol-mediated transfection, and regeneration of stably transformed plants were performed according to standard procedures (Frank et al., 2005). Plants were selected on standard growth medium containing 12.5 mg/L G418. G418-resistant transgenic lines were screened by PCR to identify lines with a disrupted *AAH* locus with the primers PpAAH5 and PpAAH6 derived from the 1.5 kb *AAH* region that was replaced by the *nptII* cassette. A PCR amplicon was obtained from wild type

genomic DNA; transformants that did not give rise to an amplicon were considered to have undergone targeted integration of the knockout construct at the *AAH* locus.

Four independent transgenic lines were selected to confirm precise 5′ and 3′ integration of the construct by PCR with primers located up- or downstream of the expected integration site and primers derived from regions of the *nptII* selection marker cassette (PpAAH-upstream/nptII-fwd and PpAAH-downstream/nptII-fwd, respectively; see Supplemental Figure 7B online). All four lines showed precise 5′ and 3′ integration of the knockout construct at the *AAH* locus. These lines are null mutants since they failed to generate an AAH-derived PCR product in a subsequent RT-PCR using the primers PpAAH-RT-PCR-fwd and PpAAH-RT-PCR-rev spanning the integrated *AAH* knockout construct (see Supplemental Figure 7C online). cDNA was synthesized from 2 µg of total RNA with an oligo(dT) primer using 200 units of SuperScriptIII Reverse Transcriptase (Invitrogen) according to the manufacturer's instructions in a total volume of 20 µl. RT-PCR with 35 cycles was performed using 1.3 µl of the synthesized cDNA with primers PpAAH-RT-PCR-fwd and PpAAH-RT-PCR-rev spanning the integrated PpAAH knockout construct. Sequences of all primers and RT-PCR primers for the *P. patens L21* control gene encoding a small ribosomal protein are reported in Supplemental Table 2 online. Genomic Southern blot analyses were performed to exclude additional integrations of the knockout construct into the nuclear DNA. Two µg of genomic DNA of knockout and wild type plants were digested with *Bsm*I (which does not cut within the knockout construct), separated on a 0.8% agarose gel, blotted onto nylon membrane, and hybridized with a radiolabeled *nptII* cassette DNA fragment. The resulting hybridization pattern demonstrates an insertion of concatemeric knockout cassettes (Kamisugi et al*.*, 2006) at the *AAH* locus in all four analyzed knockout mutants, but no illegitimate integrations at other loci (see Supplemental Figure 7D online). These lines were used for further analyses.

P. patens was grown axenically either in liquid or solid Knop medium containing 250 mg/L KH_2PO_4, 250 mg/L $MgSO_4.7\ H_2O$, 250 mg/L KCl, 1 g/L $Ca(NO_3)_2.4\ H_2O$, 12.5 mg/L $FeSO_4.7\ H_2O$, pH 5.8; solid medium was supplemented with 1% (w/v) agar. The plants were cultured under standard conditions in a growth chamber at 25 ± 1°C under a 16 h/8 h light/dark regime with a light intensity of 55 µmol $m^{-2}\ s^{-1}$. Plants grown in liquid culture were subcultured weekly; plants grown on solid medium were transferred to fresh medium monthly. For metabolic analyses, 1.5 g batches of plant material were grown in 450 mL liquid medium under standard

conditions, harvested by filtration, and freeze-dried in glass vials for 48 h. The samples were then shipped on dry ice to the Samuel Roberts Noble Foundation for metabolomic analysis.

Metabolomic Analysis

For GC-MS analysis, lyophilized samples were ground and 6.00 ± 0.05 mg was weighed into a 4 mL glass vial. Samples were extracted with methanol and chloroform (1:1.5, v/v) containing docosanol (non-polar internal standard, 10 µg/mL) for 1 h at 50ºC, with agitation (200 rpm). Water (1.5 mL) containing 25 µg/mL ribitol (polar internal standard) was added and samples were incubated for a second 1 h period. The biphasic system was then centrifuged at 2900 *g* for 30 min. Two mL of the aqueous polar phase and 1 mL of the non-polar phase were transferred to individual 2 mL glass vials. Polar extracts were dried in a vacuum centrifuge, and the non-polar extracts were dried under a stream of nitrogen gas. Dried polar extracts were methoximated in pyridine with 50 µL of freshly prepared methoxyamine-HCl (15 mg/mL), briefly sonicated, and incubated at 50ºC until the residue was resuspended. Metabolites were then derivatized by adding 50 µL of a commercial derivatization solution (Pierce Biotechnology) containing *N*-methyl-*N*-trimethylsilyltrifluoroacetamide (MSTFA) + 1% trimethylchlorosilane (TMCS) and incubating for 1 h at 50ºC. The sample was cooled to room temperature, transferred to a 300 µL glass insert, and analyzed using an Agilent 6890 GC coupled to a 5973*i* MSD scanning from *m/z* 50-650 with the acquisition rate of 2 spectra/s. Samples (1 µL) were injected at a 15:1 split ratio and injector was held at 280ºC. Separation was achieved on DB-5MS column (J&W Scientific, 60 m, 0.25 mm i.d., and 0.25 µm film) at 1 mL/min helium. The temperature program was 2 min at 80ºC followed by a 47 min ramp to 315ºC and was held at 315ºC for 12 min. The transfer line to the mass spectrometer was set to 280ºC and the MS source was set to 200ºC. Chemical ionization of the polar fraction was done using methane as the reactant gas and positive ion detection; 1 µL was injected in a splitless mode and MSD scanning was from *m/z* 50-800.

The dried non-polar extract was resuspended in 0.5 mL chloroform and transesterified by adding 0.5 mL 1.25 M HCl in methanol and incubating for 6 h at 50ºC. Following transesterification, HCl and solvent were evaporated under nitrogen until dryness, then resuspended in 70 µL pyridine and derivatized with 30 µL of MSTFA + 1% TMCS for 1 h at 50ºC. The sample was cooled to room temperature, transferred to a 300 µL glass insert and analyzed using the parameters described above for the EI-MS polar extracts (Broeckling et al., 2005). AMDIS software (http://chemdata.-nist.gov/mass-spc/amdis/) was used for mass spectra deconvolution and meta-

bolite identification by searching several EI-MS libraries including an in-house build library, the published GOLM library (http://csbdb.mpimp-golm.mpg.de/csbdb/gmd/msri/gmd_msri.html) and the commercial NIST08 library. MET-IDEA software (http://bioinfo.noble.org/download) was used for peak picking, alignment and quantification. The relative peak areas of metabolites were normalized against the peak areas of the internal standard. Normalization allows quantitative comparisons of accumulated metabolites or tentative peaks. After normalization, peak areas were $\log_2$-transformed. Metabolite abundance was contrasted between genotypes by running a mixed-model ANOVA (PROC MIXED) with lines as fixed effects. In order to generate a conservative estimate of significantly different metabolite quantities, contrasts were declared significant after adjustment for multiple testing using a false discovery rate (FDR) of 5% (Benhamini and Hochberg, 1995). The analysis was carried out using PROC MIXED in SAS 9.1 software (SAS Institute). FDR was calculated using R (R Foundation for Statistical Computing). Multi-Experiment Viewer software (Saeed et al., 2003) was used for hierarchical clustering analysis.

Quantification of selected metabolites was done using three knockout lines and three wild type samples. Standard calibration curves were constructed for phenylalanine, caffeic acid, and chlorogenic acid (used to quantify unknown caffeic esters). Ions m/z 218, 396, and 307 were used for the quantification of phenylalanine, caffeic acid, and caffeic esters, respectively. Caffeoylthreonic acid was synthesized using Steglich esterification (Neises and Steglich, 1978).

For LC-MS analysis, lyophilized samples were ground and 10.00 ± 0.05 mg was weighed into a 4 mL glass vial. Samples were extracted in 1 mL of 80% (v/v) aqueous methanol containing 18 µg/mL umbelliferone as an internal standard for 2 h at room temperature. Samples were centrifuged and the supernatant analyzed by Waters ACQUITY^{TM} UPLC/ QTOF Premier™ MS. Separations were achieved using a Waters ACQUITY^{TM} UPLC 2.1 x 100 mm, BEH C18 column maintained at 60ºC. The mobile phase consisted of (A) 0.1% (v/v) aqueous acetic acid and (B) acetonitrile and a linear gradient of 95%:5% to 30%:70% eluents A:B over 30 min with a flow rate of 0.56 mL/min. Mass spectrometry was operated in negative ion electrospray. The nebulization gas was 850 L/h at 350ºC, cone gas was 50 L/h and the source temperature was 120ºC. TOF data were acquired from *m/z* 100-2000 using lock-mass ion at an interval of 10 s. Raffinose was used as the reference compound and delivered at a concentration of 50 fmol/mL and flow rate of 0.2 mL/h. Peak picking, alignment and quantification were performed using MarkerLynx software (Waters).

REFERENCES

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate – a practical and powerful approach to multiple testing. J. R. Stat. Soc. Ser B. Methodol. **57:** 289-300.

Broeckling, C.D., Huhman, D.V., Farag, M.A., Smith, J.T., May, G.D, Mendes, P., Dixon, R.A., and Sumner, L.W. (2005). Metabolic profiling of *Medicago truncatula* cell cultures reveals the effects of biotic and abiotic elicitors on metabolism. J. Exp. Bot. **56:** 323-336.

Egener, T., Granado, J., Guitton, M.C., Hohe, A., Holtorf, H., Lucht, J.M., Rensing, S.A., Schlink, K., Schulte, J., Schween, G., Zimmermann, S., Duwenig, E., Rak, B., and Reski, R. (2002). High frequency of phenotypic deviations in *Physcomitrella patens* plants transformed with a gene-disruption library. BMC Plant Biol. **2:** 6.

Frank, W., Decker, E.L., and Reski, R. (2005). Molecular tools to study *Physcomitrella patens*. Plant Biol. **7:** 220-227.

Kamisugi, Y., Schlink, K., Rensing, S.A., Schween, G., von Stackelberg, M., Cuming, A.C., Reski, R., and Cove, D.J. (2006). The mechanism of gene targeting in *Physcomitrella patens*: homologous recombination, concatenation and multiple integration. Nucleic Acids Res. **34:** 6205-6214.

Neises B. and Steglich W. (1978). Simple method for the esterification of carboxylic acids. Angew. Chem. Int. Ed. Engl. **17:** 522-524.

Saeed, A.I., Sharov, V., White, J., Li, J., Liang, W., Bhagabati, N., Braisted, J., Klapa, M., Currier, T., Thiagarajan, M., Sturn, A., Snuffin, M., Rezantsev, A., Popov, D., Ryltsov, A., Kostukovich, E., Borisovsky, I., Liu, Z., Vinsavich, A., Trush, V., and Quackenbush, J. (2003). TM4: a free, open-source system for microarray data management and analysis. Biotechniques **34:** 374-378.