

Fatty acid synthesis system in *Glaucophyta*: Further evidence for the lack of chloroplast fatty acid synthesis in glaucophyte algae

Naoki Sato^{1,*} and Koichiro Awai^{2,3,4}

* Corresponding author

¹Graduate School of Arts and Sciences, University of Tokyo, Meguro-ku, Tokyo 153-8902, Japan.

E-mail: NS, naokisat@bio.c.u-tokyo.ac.jp

Supplementary materials

Supplementary Fig. S1. Sequence alignment of eukaryotic acetyl-CoA carboxylases of *Streptophyta* and *Glaucophyta*.

Supplementary Fig. S2. Phylogenetic tree of acyl carrier proteins (ACP).

Supplementary Fig. S3. Phylogenetic tree of chloroplast KAS.

Supplementary Fig. S4. Phylogenetic tree of mitochondrial KAS.

Supplementary Fig. S5. Phylogenetic tree of FabD.

Supplementary Fig. S6. Identification of *Glaucocystis* FAS I fragment in the alignment of *Glaucophyta* FAS I.

Supplementary Fig. S1. Sequence alignment of eukaryotic acetyl-CoA carboxylases of *Streptophyta* and *Glaucophyta*. ATH, *Arabidopsis thaliana*; MeVI, *Mesostigma viride*; CPA, *Cyanophora paradoxa*; GIWI, *Gloeochaete witrockiana*; CyGL, *Cyanoptyche gloeocystis*; GINO, *Glaucocystis nostockinearum*. Partial genomic sequences from *Glaucophyta* are aligned with complete sequences to infer possible complete sequences used in the phylogenetic analysis presented in Fig. 2.

Alignment of Eukaryotic ACCs

Biotin-carboxylation domain

ATH_AT1G36160=ACC1
MeV1_Mv06825-RA.2
CPA_00385_g24767.t1
CPA_00350_g24319.t1
CPA_00350_g24318.t1
GLWI_0184350418
CyGL_2003998
CyGL_2005688
CyGL_2005689
CyGL_2005687
CyGL_2005686
CyGL_2005690
CyGL_2008724
CyGL_2008727
GLNO_2058125
GLNO_2058049
GLWI_0184346346
GLWI_0184370582
GLWI_0184345692

1.....10.....20.....30.....40.....50.....60.....70.....80.....90.....100.....110.....120.....130.....140.....150

ATH_AT1G36160=ACC1
MeV1_Mv06825-RA.2
CPA_00385_g24767.t1
CPA_00350_g24319.t1
CPA_00350_g24318.t1
GLWI_0184350418
CyGL_2003998
CyGL_2005688
CyGL_2005689
CyGL_2005687
CyGL_2005686
CyGL_2005690
CyGL_2008724
CyGL_2008727
GLNO_2058125
GLNO_2058049
GLWI_0184346346
GLWI_0184370582
GLWI_0184345692

.....160.....170.....180.....190.....200.....210.....220.....230.....240.....250.....260.....270.....280.....290.....300

ATH_AT1G36160=ACC1
MeV1_Mv06825-RA.2
CPA_00385_g24767.t1
CPA_00350_g24319.t1
CPA_00350_g24318.t1
GLWI_0184350418
CyGL_2003998
CyGL_2005688
CyGL_2005689
CyGL_2005687
CyGL_2005686
CyGL_2005690
CyGL_2008724
CyGL_2008727
GLNO_2058125
GLNO_2058049
GLWI_0184346346
GLWI_0184370582
GLWI_0184345692

.....310.....320.....330.....340.....350.....360.....370.....380.....390.....400.....410.....420.....430.....440.....450

Biotin-carboxylation domain

ATH_AT1G36160=ACC1
MeV1_Mv06825-RA.2
CPA_00385_g24767.t1
CPA_00350_g24319.t1
CPA_00350_g24318.t1
GLWI_0184350418
CyGL_2003998
CyGL_2005688
CyGL_2005689
CyGL_2005687
CyGL_2005686
CyGL_2005690
CyGL_2008724
CyGL_2008727
GLNO_2058125
GLNO_2058049
GLWI_0184346346
GLWI_0184370582
GLWI_0184345692

.....460.....470.....480.....490.....500.....510.....520.....530.....540.....550.....560.....570.....580.....590.....600

ACCse BT-domain

ATH_AT1G36160=ACC1
MeV1_Mv06825-RA.2
CPA_00385_g24767.t1
CPA_00350_g24319.t1
CPA_00350_g24318.t1
GLWI_0184350418
CyGL_2003998
CyGL_2005688
CyGL_2005689
CyGL_2005687
CyGL_2005686
CyGL_2005690
CyGL_2008724
CyGL_2008727
GLNO_2058125
GLNO_2058049
GLWI_0184346346
GLWI_0184370582
GLWI_0184345692

.....610.....620.....630.....640.....650.....660.....670.....680.....690.....700.....710.....720.....730.....740.....750

Biotinyl domain

ACCse central domain

ATH_AT1G36160=ACC1
MeV1_Mv06825-RA.2
CPA_00385_g24767.t1
CPA_00350_g24319.t1
CPA_00350_g24318.t1
GLWI_0184350418
CyGL_2003998
CyGL_2005688
CyGL_2005689
CyGL_2005687
CyGL_2005686
CyGL_2005690
CyGL_2008724
CyGL_2008727
GLNO_2058125
GLNO_2058049
GLWI_0184346346
GLWI_0184370582
GLWI_0184345692

.....760.....770.....780.....790.....800.....810.....820.....830.....840.....850.....860.....870.....880.....890.....900

← ACCase central domain

ATH_AT1G36160=ACCI1	EFESIT	-----RNSLT-----	-----DFFPAKLLKSTLEAHLSSE--DEKRGALERLLEPLMSLAKSVGGGRSHARVIVHLPFEYLSVEELFDNDML-ADVIERMRLVKKDLKIVDVLSE	949
MeV1_My06825-RA.2	MYAAATGCGGGGASSEAMVADADNNSTSDRNVAAQGVGKGMETLSPGOQVDFPFAOAILLSALLDYNSCADEETRMARRRLVDPMTLLAPVASGRDEFVQVVDGLLEEYKVEELFAASPSVEVDVIEKLREOYREDLAGVUVVHSE	-----	-----	1024
CPA_00385_g24767.t1	EYAEAL	-----AAGEG-----	-----AGRPPELLRVGPIGLLELAAMALPSKAERDALAAQAKPLAELCSAYAGGLPGREAAALGSLLEFIAELPSSGKPKDDEDAIALREDEFHADLGKAVDLRSH	963
CPA_00350_g24319.t1	-----	-----	-----	582
CPA_00350_g24318.t1	EYSAVAEPRKAG	-----	-----PGAARPELNVAALKTLDEIAGKLPKAEKDAYAAQTKPLADLCAAYAGGSAGREVAVLSDLIREYNOEHAFAVSVGKSDDEDVLEEMREONRSEMAKVDVDRMLSH	478
GLWI_0184350418	GLWI_0184350418	-----	-----PLPRQSSVFNGLSEEDDAGAGOVIKDIEAVIHKIDILDTAARLHVKEKSDSYVANITIGLELCRIYAKGLPGLIDILTSALREYISVEOLFNGKNDDEDVIEELREONRNSLSKVVDITQSH	977
CyGL_2003998	-----	-----	-----	145
CyGL_2005688	KYRLAV	-----	-----BDARKGGGVGVG-----AELDLTAISKILEDYTPKVKAKTERDAFLALTOPLRDLVKVYSGGLGCHLVDVLASIMREYIDVELL	883
CyGL_2005689	KYRLAV	-----	-----BDARKGGGVGVG-----AELDLTAISKILEDYTPKVKAKTERDAFLALTOPLRDLVKVYSGGLGCHLVDVLASIMREYIDVELL	194
CyGL_2005687	KYRLAV	-----	-----EDARKGGAGVGVGVG-----GAPPVELDLTAISKILEDHTPKVKTERDAFVALTOPLRDLVKVYSGGLGCHLVDVLASIMREYIDVELL	928
CyGL_2005686	KYRLAV	-----	-----EDARKGGAGVGVGVG-----GAPPVELDLTAISKILEDHTPKVKTERDAFVALTOPLRDLVKVYSGGLGCHLVDVLASIMREYIDVELL	928
CyGL_2005690	KYRLAV	-----	-----EDARKGGAGVGVGVG-----GAPPVELDLTAISKILEDHTPKVKTERDAFVALTOPLRDLVKVYSGGLGCHLVDVLASIMREYIDVELL	360
CyGL_2008724	-----	-----	-----	-----
CyGL_2008727	-----	-----	-----	-----
GLNO_2058125	EYKSGLP	-----GSGAAG-----	-----VPKPMKIEAAVGKINAAITEATKLSTKAEDAFAPAAQAFIVELVKKNRGLAGHAVSVFSSILRSYITIEAPFNGKSDDEDVTEELREHRAIDLTKLVLDILMSH	929
GLNO_2058049	-----	-----	-----	-----
GLWI_0184346346	-----	-----	-----	-----
GLWI_0184370582	SYRSGLAHKS	-----	-----AMAYALELAFSPPIREVLDDKFOAELPLDER-ASPATVADPIFKLVDRYTEG--TKKVFVTQPIKEYLEVEKIFQ-NRQNRDAILEMRQKTYTDLNKVDFDLSLR	272
GLWI_0184345692	-----	-----	-----	168
.....910.....920.....930.....940.....950.....960.....970.....980.....990.....1000.....1010.....1020.....1030.....1040.....1050	-----	-----	-----	-----

ATH_AT1G36160=ACCI1	CGIKNNKKLVLRMLSEOLVY-PNPAAYRDKLIRFSLTNHTNYSLELAKASOLLEOKLSELRSNIARLSLEMFTEDEG	-----	-----NMDPKRKSANERIEDLVASLAVEDALVGLFDHSDHLL	1067
MeV1_My06825-RA.2	CGVKKKHELVLVMSAVVM-PPSRFESQLERFAGLMOTQCAEVAORASILLDEIHLNRLASINRCLLDTPVGERGRHVRGPIHALTTPPLPTLASGRTIPASDFSLDEDVARLSLSEPRIKLSLVSHSLPVEDALEGLDHPDTPQ	-----	-----GPAGARAAGCLH-----IVKANYGIYDILVPMFAHPEDV	1173
CPA_00385_g24767.t1	AGLPRKLLKLAAPSKMAAGPL-EPYAAALHOLVAFAGRPYAEVALRARALL--EAMANPNDNRKORLEELRRAAAS	-----	-----	1074
CPA_00350_g24318.t1	-----	-----	-----	582
CPA_00350_g24319.t1	REVARSKSLVLALILEHTGSTNATEHNVSALODIAGFSGRSYSDVGLKARQLLTHNMPF--YICRRQRLEELRRAARS	-----	-----TEPAARRBSLLPLTSHDCHGFDVLMFPFFOAHSDSI	590
GLWI_0184350418	REFALKSKVVGIEDFIAMHVVVEDFVSVLBELAALGKNYSVEALKARQLL--ISYQPSYSGRRVIMEKVLREADVD	-----	-----TLHRRDLILP-----VVVNSDCHGFDVLMFPFFRHPDAI	1087
CyGL_2003998	-----	-----	-----	145
CyGL_2005688	-----	-----	-----	883
CyGL_2005689	-----	-----	-----	194
CyGL_2005687	ROLQRKKNKVLGILEYISVNGLVHPLERILHDVANFTAROHSEVALRARQIL--ISYQPSFOORTKTLEAVLQAANO	-----	-----TEVAKRRDALR-----PIVTOSEGVDFVLPIFOAEDOSI	1040
CyGL_2005686	ROLQRKKNKVLGILEYISVNGLVHPLERILHDVANFTAROHSEVALRARQIL--ISYQPSFOORTKTLEAVLQAANO	-----	-----TEVAKRRDALR-----PIVTOSEGVDFVLPIFOAEDOSI	1040
CyGL_2005690	ROLQRKKNKVLGILEYISVNGLVHPLERILHDVANFTAROHSEVALRARQIL--ISYQPSFOORTKTLEAVLQAANO	-----	-----TEVAKRRDALR-----PIVTOSEGVDFVLPIFOAEDOSI	472
CyGL_2008724	-----	-----	-----	-----
CyGL_2008727	-----	-----	-----	-----
GLNO_2058125	RELARKKNKVLGLLDILIASCVGVSEHVSVLBELAALGKNYSVEALKARQIL--ISYQPSYSGRRVIMEKVLAKANDR	-----	-----PSNKKDILVPLVNSDNLDFVLLPIFOHADPI	1039
GLNO_2058049	-----	-----	-----	-----
GLWI_0184346346	-----	-----	-----	-----
GLWI_0184370582	NPLSKKHLELITILSIEKKKTTEDFVALLHDLASLTKGKSDIDVSKARQLL--IRFQLPSFKQGVATENLRNAHAAXXXXXXXX	-----	-----RNAHAASGANRDLILTLLIDAPSLDFVLISFLKHPNDI	272
GLWI_0184345692	-----	-----	-----	294
.....1060.....1070.....1080.....1090.....1100.....1110.....1120.....1130.....1140.....1150.....1160.....1170.....1180.....1190.....1200	-----	-----	-----	-----

ATH_AT1G36160=ACCI1	ORRVVETIIRRLYQPVVVKDSVRMQHRSGLLASNEF	-----	-----LEEMERKNIGLDDPDSEKGLVEKRSKR	1155
MeV1_My06825-RA.2	ORARAEITIRRLYQPVVVKDSVRMQHRSGLLASNEF	-----	-----LEEMERKNIGLDDPDSEKGLVEKRSKR	1321
CPA_00385_g24767.t1	AGAALESFVRRITVOYELEAVSLHNAEGKPSRVDFTF	-----	-----LFAQ-----AAERKKGHGRAPVR	1150
CPA_00350_g24319.t1	-----	-----	-----	582
CPA_00350_g24318.t1	VEAATEGYVRRAYVAYDISILRPVDFSGIARVDFDFTFAGGAGSSSOPSTPTGNPFAAAAAAAMGGRAGRLGTVDSSDNLGLATKAPE--	-----	-----RRGCVVAVASLELERDFFPRIV	703
GLWI_0184350418	AFVAMEYVRRAYQAYDIPDLTITVADGGIP-RAEWNDF	-----	-----ILPGFTLPPAPANPSRGRFGKMPFDEADNLLALQGHVPT	1187
CyGL_2003998	-----	-----	-----	145
CyGL_2005688	-----	-----	-----	883
CyGL_2005689	-----	-----	-----	194
CyGL_2005687	ADTALEAYVMRAYQAYELVLEVKTGPOFS-RAVNDY	-----SFGGAAAGNVG-----	-----IRLGRMVVDFSSDNLTLGLOORAPI	1133
CyGL_2005686	ADTALEAYVMRAYQAYELVLEVKTGPOFS-RAVNDY	-----SFGGAAAGNVG-----	-----IRLGRMVVDFSSDNLTLGLOORAPI	1133
CyGL_2005690	ADTALEAYVMRAYQAYELVLEVKTGPOFS-RAVNDY	-----SFGGAAAGNVG-----	-----IRLGRMVVDFSSDNLTLGLOORAPI	565
CyGL_2008724	-----	-----	-----	-----
CyGL_2008727	-----	-----	-----	-----
GLNO_2058125	KAEAMEYVRRAYQAY	-----	-----	1055
GLNO_2058049	-----	-----	-----	-----
GLWI_0184346346	-----	-----	-----	-----
GLWI_0184370582	ORLALVYVKSRYRAYDIRFES-KFVSSHICAENRFIMP	-----	-----DGPASSFSESTLAINFASNI	272
GLWI_0184345692	-----	-----	-----	353
.....1210.....1220.....1230.....1240.....1250.....1260.....1270.....1280.....1290.....1300.....1310.....1320.....1330.....1340.....1350	-----	-----	-----	-----

ATH_AT1G36160=ACCI1	RETKHND	-----	-----YTAGAP--LGGMMHIAIV-----	1239
MeV1_My06825-RA.2	VARHNSRPAKLAFAFDCDASLSPRGPGSGPGMGVGGGASAGE	-----	-----KIVSLYVALTPIGLPSAPPKGGSTHASPAPAHHHSHPKQIPSHLAQITAESRAKLDDEFLQSPAGHAGCTTAKELQALDVOALTPSOQPSKI	1469
CPA_00385_g24767.t1	AARALILE	-----	-----APOAGNE--PAVNVMDVAIV-----	1226
CPA_00350_g24319.t1	-----	-----	-----	582
CPA_00350_g24318.t1	EMYORSLEG	-----	-----RHVHPQPMVILLDVALLEAG--	778
GLWI_0184350418	ELVRSV	-----	-----AFORDNPSAAPVNVLDVALV-----	1259
CyGL_2003998	-----	-----	-----	145
CyGL_2005688	-----	-----	-----	883
CyGL_2005689	-----	-----	-----	194
CyGL_2005687	ELVAEH	-----	-----VDRGSLKYNANVVDVALI-----	1205
CyGL_2005686	ELVAEH	-----	-----VDRGSLKYNANVVDVALI-----	1205
CyGL_2005690	ELVAEH	-----	-----VDRGSLKYNANVVDVALI-----	637
CyGL_2008724	-----	-----	-----	-----
CyGL_2008727	-----	-----	-----	-----
GLNO_2058125	-----	-----	-----	-----
GLNO_2058049	-----	-----	-----	-----
GLWI_0184346346	-----	-----	-----	-----
GLWI_0184370582	SVAPSSSS	-----	-----	272
GLWI_0184345692	-----	-----	-----	361
.....1360.....1370.....1380.....1390.....1400.....1410.....1420.....1430.....1440.....1450.....1460.....1470.....1480.....1490.....1500	-----	-----	-----	-----

ATH_AT1G36160=ACCI1	RH-----SFHWSEKQYVVEEPLLRHLEPPLSIYLELDKLGYSNIOYPSRDRQWHLTYTVDKVPFI	-----	-----KRMFLRSLVRGATMNDG--FILOOGQDKLISOTLISMAF--SKCVLRSLMD--	1350
MeV1_My06825-RA.2	QSAHFPCLSALEGAAGYRVVPLLRNIOPPLAQLLELGKLEAF-NVYTHSRDPOWHVITCTCPPKATPKMSFSAKSANVRAMPATSGVTFKWPRPMSVFMRSILRIPDAEVSVLSAPDD-IVLPFGMGATGSAIGLRRAGTOSASNC	-----	-----	1617
CPA_00385_g24767.t1	AFYTFRENLGIGEDTILRHVLPAPFSYELEKRLSNF-TYTKCSAERHILHYAEKSKTPAAGAGAGAGKAPAGGEGSGPVDADRRFFVRAI--ASLVFGRNDAAVSIPAEHVFLEGLKVTFRQALAEAGCD	-----	-----	1360
CPA_00350_g24319.t1	-----	-----	-----	582
CPA_00350_g24318.t1	-----	-----	-----	994
GLWI_0184350418	SIFTFRENLNANDPIYRHILPPLAYELEKRLSNF-TIADTPNTRHILHFAEKISPOQRKG	-----	-----OGGAGGAAEADVDRFFVVALVROAAITFGSHEVSVIPDAERVFTEALKLEAVAMCDAR	1383
GLWI_0184350418	-----	-----	-----	145
CyGL_2003998	-----	-----	-----	883
CyGL_2005688	-----	-----	-----	194
CyGL_2005689	-----	-----	-----	1328
CyGL_2005687	GCFTFRERNGYEDLTYRHIVPPMAYELELERLSNF-NIOQYPSNRNVHVHYAEKVPVR	-----	-----ARNTRAIEGGSESESPADADRRFFVRAVLROARIFTC-TEAOVSIPDAERTFVEALKM--LEMAMVD	1328
CyGL_2005686	GCFTFRERNGYEDLTYRHIVPPMAYELELERLSNF-NIOQYPSNRNVHVHYAEKVPVR	-----	-----ARNTRAIEGGSESESPADADRRFFVRAVLROARIFTC-TEAOVSIPDAERTFVEALKM--LEMAMVD	1328
CyGL_2005690	GCFTFRERNGYEDLTYRHIVPPMAYELELERLSNF-NIOQYPSNRNVHVHYAEKVPVR	-----	-----ARNTRAIEGGSESESPADADRRFFVRAVLROARIFTC-TEAOVSIPDAERTFVEALKM--LEMAMVD	760
CyGL_2008724	-----	-----	-----	-----
CyGL_2008727	-----	-----	-----	-----
GLNO_2058125	-----	-----	-----	-----
GLNO_2058049	-----	-----	-----	-----
GLWI_0184346346	-----	-----	-----	-----
GLWI_0184370582	-----	-----	-----	-----
GLWI_0184345692	-----	-----	-----	-----
.....1510.....1520.....1530.....1540.....1550.....1560.....1570.....1580.....1590.....1600.....1610.....1620.....1630.....1640.....1650	-----	-----	-----	-----

ACCCase central domain →

ATH_AT1G36160=ACCI1	ASASASYDDILSLISEASEDEPSPAPNFKSLRCSNLVVGAVNLOS	-----	-----AMEELELNANNAKMPDHAMFLCILLREQDIDLVPPPRVSVNADEDEETVLEALBEAAREHRSVGVGMRLHGVCEVRLWL-VSSGLA	1441
MeV1_My06825-RA.2	-----	-----	-----	1757
CPA_00385_g24767.t1	-----	-----	-----	1421
CPA_00350_g24319.t1	-----	-----	-----	582
CPA_00350_g24318.t1	-----	-----	-----	954
GLWI_0184350418	-----	-----	-----	1445
CyGL_2003998	-----	-----	-----	145
CyGL_2005688	-----	-----	-----	883
CyGL_2005689	-----	-----	-----	194
CyGL_2005687	-----	-----	-----	1390
CyGL_2005686	-----	-----	-----	1390
CyGL_2005690	-----	-----	-----	822
CyGL_2008724	-----	-----	-----	-----
CyGL_2008727	-----	-----	-----	-----
GLNO_2058125	-----	-----	-----	-----
GLNO_2058049	-----	-----	-----	-----
GLWI_0184346346	-----	-----	-----	-----
GLWI_0184370582	-----	-----	-----	-----
GLWI_0184345692	-----	-----	-----	-----
.....1660.....1670.....1680.....1690.....1700.....1710.....1720.....1730.....1740.....1750.....1760.....1770.....1780.....1790.....1800	-----	-----	-----	-----

Carboxyltransferase N-term. domain

ATH_AT1G36160=ACC1	-----SDSETSPEN-----	GFQYLYLSPEDHRI	-----GSSVIAEVLKSSGETRWVDTIVGKEDGIGVENLGGSGAIGAYSKAYNEFTLLFVFGSRVVGIGAYLARLGMRC	1755
MeV1_Mv06825=RA.2	-----TDDRVPFSK-----	GFQYLYLTDYDQKV	-----GSSVRAEQVAENGOTRWVVDIIIGREELGVGSLGSGTIAGIFYPSKANVETFLFIISGRSVGIGAYLARLGQR	2115
CPA_00385_g24767.t1	ARINLAAEARPLPLGPEIDPSRAEKERAIAGRRHFRPLPFPPELGLSLFSFFLLG	GHRYLYLTADHRL	-----KGKSVVCEPVTVVCVGGGGSETRYRLIDIIGAEDGLGVENLGGSLIAGETSRAYEETFLFIPTVTSRTVIGIGAYLVRLGQR	1841
CPA_00350_g24319.t1	-----	-----	-----	582
CPA_00350_g24318.t1	-----	OKGFYLYLSQADYERLR	-----EKSVVCEPVOVHLGGRTETRYRIVDVGAEHGLGVENLGGSGMIAGETSRAYDEIFLLFVTARTVIGIGAYLVRLGQR	1277
GLWI_0184350418	-----NDPLNFSK-----	GFQYLYLDEATRYTL	-----SERKSVVCEPVDVTLDDGOVRRYKIVDIIIGAGLGVENLGGSGMIAGETSRHAYDDIFLFTVARTVIGIGAYLVRLGQR	1771
CyGL_2003998	-----	-----	-----	145
CyGL_2005688	-----	-----	-----	883
CyGL_2005689	-----	-----	-----	194
CyGL_2005687	-----IDPQVPDKG-----	-----	-----	1619
CyGL_2005686	-----IDPQVPDKG-----	-----	-----	1619
CyGL_2005690	-----IDPQVPDKG-----	-----	-----	1051
CyGL_2008724	-----	-----	-----KITDIIGAVHGLGVENLGGSGMIAGETSRHAYEDVFLTLVLTSTRVIGIGAYLVRLGQR	58
CyGL_2008727	-----	-----	-----	1055
GLNO_20581125	-----	LKGFXYLYVDDDFRL	-----LAKKNVSNCELRLEVCNGDETQVRYRLVDIIIGAGLGVENLGGSGMIAGETSRAYDDIFLFTVARTVIGIGAYLVRLGQR	264
GLNO_2058049	-----	-----	-----	272
GLWI_0184346346	-----	-----	-----	361
GLWI_0184370582	-----	-----	-----	326
GLWI_0184345692	-----NDISKGFXYLYLDEADYORLAPHNSVLTSLVODEHTKKYRITDIIGAEDGLSVENLRGSGMIAGETSIAYHDFITLFTVTRVIGIGAYLVRLGQR	-----	-----	2250
.....2110.....2120.....2130.....2140.....2150.....2160.....2170.....2180.....2190.....2200.....2210.....2220.....2230.....2240.....2250	-----	-----	-----	-----

► Carboxyltransferase C-term. domain

At1G36160=ACCL1	IPVGIVVAVETRTMELVLPADPANPMSQETVVVQAGGVWVFPDPSAAKTAAGALDMFDR	EGLPFLIFANWRGFGGGORDLFGQILQAGSTIVEILRLTRYKQPVFVVI	PPGGLRGGAAWVVVD	QINSDKVENIYADETARGNVL	EPGGIVEIKF	2042		
MeV1_Mv06825-RA.2	IPVAVIAVETRTMELVLPADPANPMSQETVVVQAGGVWVFPDPSAAKTAAGALDMFDR	EGLPFLIFANWRGFGGGORDLFGQILQAGSTIVEILRLTRYKQPVFVVI	PPGGLRGGAAWVVVD	QINSDKVENIYADETARGNVL	EPGGIVEIKF	2407		
CPA_00385_g24767.t1	ALQGLFLAA	LGPAASEDPRRWGFLA	DAVRAATGAPGFSV	GRYP	RHAATSTSPFGSCSS	ILCD	2041	
CPA_00350_g24319.t1						RDGEAGGYL	DPYN	582
CPA_00350_g24318.t1	IPGLGVIVETRTMELVLPADPANPMSQETVVVQAGGVWVFPDPSAAKTAAGALDMFDR	EGLPFLIFANWRGFGGGORDLFGQILQAGSTIVEILRLTRYKQPVFVVI	PPGGLRGGAAWVVVD	QINSDKVENIYADETARGNVL	EPGGIVEIKF	1567		
GLWI_1814350418	IPMGVISVETRTMELVLPADPANPMSQETVVVQAGGVWVFPDPSAYKTAAGALDMFDR	EGLPFLIFANWRGFGGGORDLFGQILQAGSTIVEILRLTRYKQPVFVVI	PPGGLRGGAAWVVVD	QINSDKVENIYADETARGNVL	EPGGIVEIKF	2060		
CyGL_2003998								145
CyGL_2005588								883
CyGL_2005589								194
CyGL_2005587								1619
CyGL_2005586								1619
CyGL_2005590								1051
CyGL_2008724	IPVGFIIVAVETRSVEQILPADPANPMSTEAVVQAGGVWVFPDPSAYKTSALADMSH	EGLPFLIFANWRGFGGMRDMFDEILKFGSYIVDNLRLTRYKQPVFVVI	PPGGLRGGAAWVVVD	QINSDKVENIYADETARGNVL	EPGGIVEIKF	347		
CyGL_2008727								24
GLNO_2051125								1054
GLNO_2050849	IPVGFIIVAVETRSVEQILPADPANPMSQETVVVQAGGVWVFPDPSAYKTAAGALDMFDR	EGLPFLIFANWRGFGGMRDMFDEILKFGSYIVDNLRLTRYKQPVFVVI	PPGGLRGGAAWVVVD	QINSDKVENIYADETARGNVL	EPGGIVEIKF	504		
GLWI_1814346346								272
GLWI_1814370582								361
GLWI_1814345692	LPPIGVICVETRTVENMILPADPANPEETVEIVQAGGVWVFPDPSAYKTAAGALDMFDR	EGLPFLIFANWRGFGGMRDMFDEILKFGSYIVDNLRLTRYKQPVFVVI	PPGGLRGGAAWVVVD	QINSDKVENIYADETARGNVL	EPGGIVEIKF	614		
	2410.....2420.....2430.....2440.....2450.....2460.....2470.....2480.....2490.....2500.....2510.....2520.....2530.....2540.....2550							

➔ to 2250

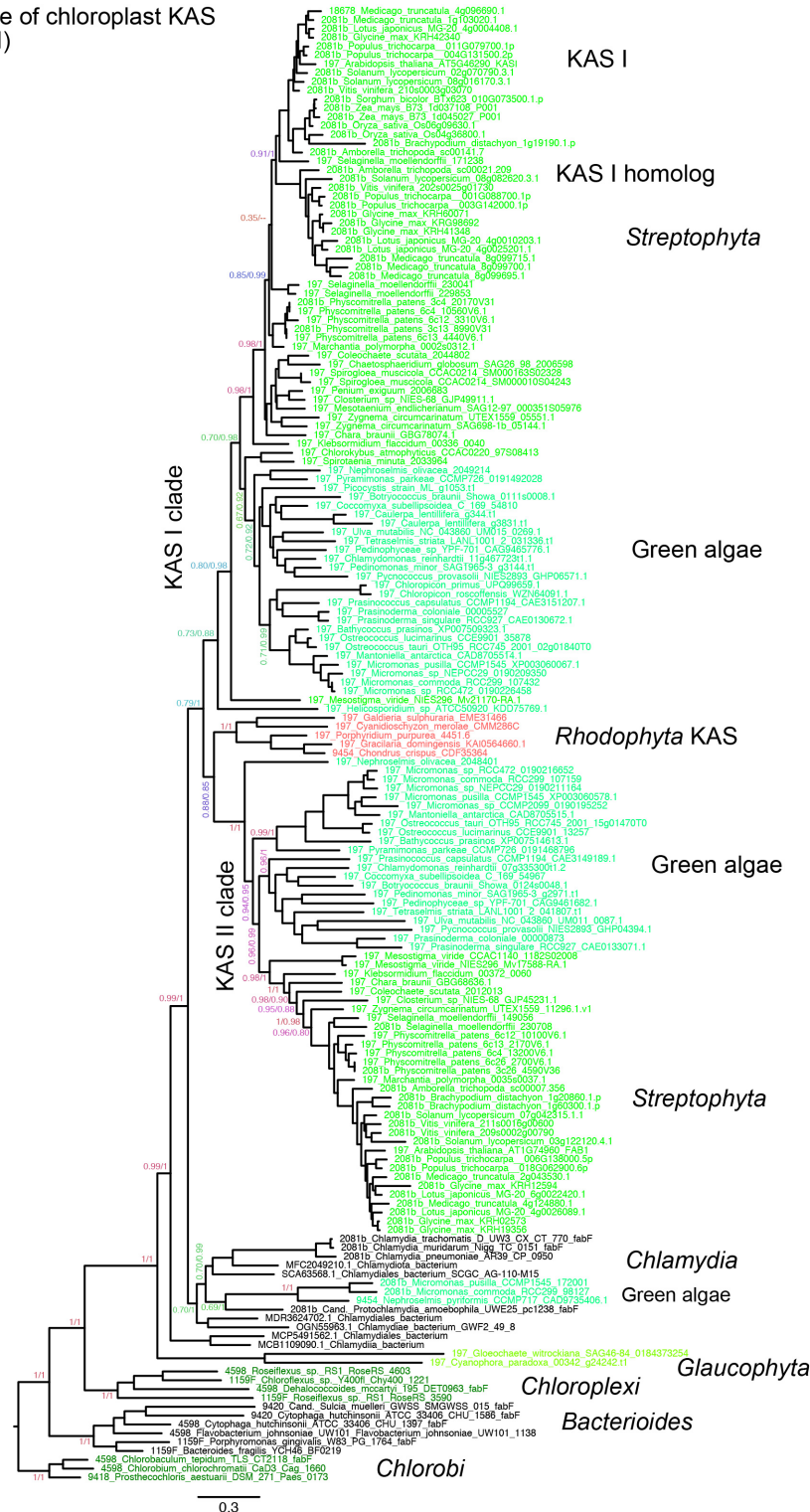
ATHT_ATIG36160=ACCI	RTKELLECGRRLDOKLISLAKKLLDAK	S-EAYANIELLQOLKAREKOLLPVVIGIAIKFAELHDSMRMAAKGVKISVVEVSGSRSPFFKKLNRRLABSSLVKNVR	2149
MeVI_Mv06825-RA.2	REPOLLEMRRLDPTMREAPPVLL	KQALPSKHRVALHFAELHDTFVRMKAKGVITGIVPWASAREFFYWRLLRRRLVEESFIRRLRTLDNRTPPAATPGHSFSAISSS	2517
CPA_00385_g24767.t1			2041
CPA_00350_g24319.t1			582
CPA_00350_g24318.t1	RKRDLOAAMARILDPRIRLNABLRDGGKDAEG	APLTFQRKREIKALIBORESELLPMYNQAVQFADLHDTGPRMKAKGVIEQVVPWKAAREFFYWRLLRRRLAEKAVRGDIEHAR	1683
GLWI_0184350418	RKRDLLAMORLDDOKLKDLYTALKEPVS	GLEGEARRRELKRDIBORESELLPIYNQIAVQFADLHDTGPRMKAKGVIEAVVPWPKSSREFFYWRIRRRVAVELGVQDIA	2168
CyGL_2003998			145
CyGL_2005688			883
CyGL_2005689			194
CyGL_2005687			1619
CyGL_2005686			1619
CyGL_2005690			1051
CyGL_2008724	RRRDLVVMHRLDOKLMEINAGLKSE	PMEEKKRLKKAIESREEKLLNVYNQIAVQFADLHDTGPRMKAKNVIRAVVPWKAAREFFYWRLLRRRLAEAIRSDISRAR	454
CyGL_2008727	RRRDLVTAHRLDOKLMEINAGLKTE	TIEKKKRLKKAIESREEKLLNVYNQIAVQFADLHDTGPRMKAKNVIRAVVPWKAAREFFYWRLLRRRLAEAIRSDIS	128
GINO_2058125			1055
GINO_2058049	NRRELDAHMLRDLPLRELDARLKGQ	GVLMEEKKREIKRLVBORESELLPMYNQAVQFADLHDTGPRMKAKGVITKVPWKAAREFFYWRLLRRRL	650
GLWI_018434346			172
GLWI_0184370582			361
GLWI_0184345692	RRDRDKRMHRLDHVLIDDLKLAQLAKGAVSSSSSSSSSSSDSIDREKKRIISLITNRERLLPTVQGVALKFAELHDTGPRMKAKGVISEVLONKRAHSPFFYWRLLRRRLAEIRKISNAOCHCFNNQVREFILRWMDPEEA	2560.2560.2570.2580.2580.2590.2600.2600.2610.2620.2630.2640.2650.2660.2670.2680.2690.2690.2700	764

ML tree of acyl-carrier proteins (ACP)



Supplementary Fig. S2. Phylogenetic tree of acyl carrier proteins (ACP). The ML tree is shown. This is an unrooted tree intended to show the relatedness of sequence clades. Plants and algae contain two types of ACPs, one related to the ACPs of *Alphaproteobacteria* (mitochondrial ACP), and the other related to the ACPs of *Cyanobacteria* (chloroplast ACP). The ACP encoded by the chloroplast genome of *Cyanophora paradoxa* (not included in the analysis) is not related to cyanobacterial ACP, but is distantly related to all these sequences. In a previously published phylogenetic tree (Sato 2020a), *Cyanophora* ACP branched from within the clade of actinobacterial ACP. Therefore, it was not included in the current analysis to avoid deformation of phylogenetic tree. The values on each branch indicate branching confidence levels for the ML and BI trees. For simplicity, 1.00 is shown as "1". Confidence values are indicated only on large branches. For color codes, see the legend for Fig. 2.

ML tree of chloroplast KAS
(ML/BI)



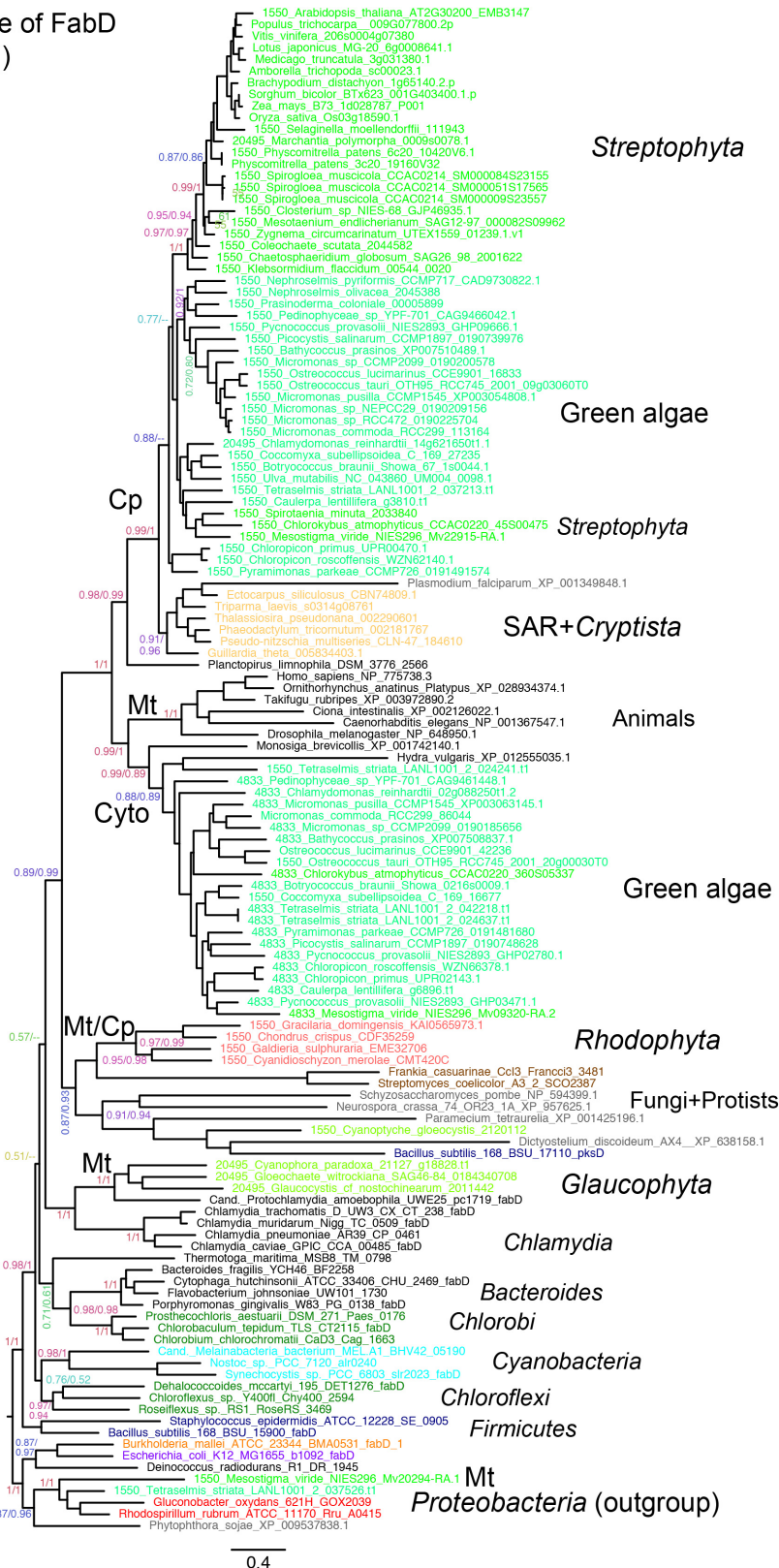
Supplementary Fig. S3. Phylogenetic tree of chloroplast KAS. Originally, a large phylogenetic tree including many more taxa was prepared using the homolog clusters in APclust2024, Gclust2024, and some other datasets. The sequences related to the chloroplast KAS clade were used to infer a more precise tree. This is an ML tree, with confidence values from ML and BI trees. The *Chlorobi* clade was taken as the outgroup based on the initial larger tree. For color codes, see the legend for Fig. 2.

ML tree of mitochondrial KAS
(ML/BI)



Supplementary Fig. S4. Phylogenetic tree of mitochondrial KAS. Originally, a large phylogenetic tree including many more taxa was prepared using the homolog clusters in APclust2024, Gclust2024, and some other datasets. The sequences related to the mitochondrial KAS clade were used to infer a more precise tree. This is an ML tree, with confidence values from ML and BI trees. The clade including *Betaproteobacteria* and *Gammaproteobacteria* was taken as the outgroup based on the initial larger tree. For color codes, see the legend for Fig. 2.

ML tree of FabD
(ML/BI)



Supplementary Fig. S5. Phylogenetic tree of FabD. This is an ML tree, with confidence values from ML and BI trees. The *Proteobacteria* clade was taken as the outgroup based on an initial larger tree. For color codes, see the legend for Fig. 2.

```

EC122136_fr1 -----
GIWI_FAS_0184342518 NTFEAYAQANKDSEFFERNYTPSEIQYCDAIQEG--RFSRYAGRWAAKEAVVKAI---TSASPOGPSL--WKGSAAPLKD 3895
CPA_FASc_0025_g7952.t1 SIFDL---ANED--FISRNFTDDEVAYCQSAFTDGERQARFAGRWAAKEAVVKAL---SSSAPDSQNL--WKGSAGAPLKN 3920
CPA_FASa_1335_g8213.t1 ATFD---ASET--MLERNFTAEELEHCEASSDA--PSRLAGKWAAKEAVAKAL---SSSAPDSAAAR--LRGSAAPLNE 3956
CPA_FASb_0912_g5450.t1 GLFRGRF-AGDL--FLERNFTPAELAHCARQDPD--ASSLAGRWAAKEAVVKALGAASEAAPATPANKGLRGSAAPLRE 3965
.....4330.....4340.....4350.....4360.....4370.....4380.....4390.....4400

                                .:: * *:***: * *:***:
EC122136_fr1 -----CHYGR-----GISEVVVTISHAGAYAVAQAIAR----- 28
GIWI_FAS_0184342518 IEVLPTDSGAPVVILHGHAKAVATAGITEVTVTISHAGAYAIQAQATAR----- 3944
CPA_FASc_0025_g7952.t1 IEIAHSSSGAPIVKLRAPAEVAKAVAVSEVIVTISHAGAYALAQAVAR----- 3969
CPA_FASa_1335_g8213.t1 IEILPSPSGAPQVRLHGYAREVAAAAGVAEIVVVSISHAGDYAVAHAVARCVAGLVTFRPGASAPGPPFLARQLHFGPAAS 4036
CPA_FASb_0912_g5450.t1 IEL-----AACGLAAVAVSIAHSGEYAVAHAA-----PAA- 3998
.....4410.....4420.....4430.....4440.....4450.....4460.....4470.....4480

EC122136_fr1 ----- 28
GIWI_FAS_0184342518 ----- 3944
CPA_FASc_0025_g7952.t1 ----- 3969
CPA_FASa_1335_g8213.t1 AASPRAFSSAPVPPPTAGSVVEOPPLARGFSPAPASAFSPVRLSGOOMKRGLLKVETAPLVAPPPRRPAPF 4106
CPA_FASb_0912_g5450.t1 -----PLP-----APAPSSKAAP-----RTQAA 4016
.....4490.....4500.....4510.....4520.....4530.....4540.....4550

```

Supplementary Fig. S6. Identification of *Glaucozystis* FAS I fragment in the alignment of *Glaucophyta* FAS I. Only the C-terminal region of the multiple alignment of the four FAS I of *Glaucophyta* is shown as the base. The transcript EC122136 of *Glaucozystis nostochinearum* from the EMBI database is 636 bases long. The short reading frame found in the 5' most region was added to the alignment.