

Supplementary information file

Evans et al (2024) “Plastid ancestors lacked a complete Entner-Doudoroff pathway, limiting plants to glycolysis and the pentose phosphate pathway”

Supplementary table 1. List of organisms used for phylogenetic analysis

Supplementary table 2. Occurrence of EDD homologs in genomes of plastid sister groups

Supplementary table 3. PCR primers used in this study.

Supplementary figure 1. Amino acid alignment of representative DHAD and EDD sequences

Supplementary figure 2. Species phylogeny of the Cyanobacteriota as previously reported by Moore et al.

Supplementary figure 3. SDS-PAGE analysis of purified preparations of dihydroxy acid dehydratase (DHAD), 6-phosphogluconate dehydratase (EDD) and 2-keto-3-deoxy-6-phosphogluconate aldolase (EDA).

Supplementary figure 4. Analytical detection of 2-keto-3-deoxygluconate-6-phosphate (KDPG) by liquid chromatography – tandem mass spectrometry (LC-MS/MS).

Supplementary figure 5. In vitro enzyme assay of purified soybean 2-keto-3-deoxygluconate-6-phosphate aldolase (GmEDA) showing aldolase activity towards substrates of the Calvin-Benson Cycle.

Supplementary figure 6. MS/MS product ion spectra of important metabolites used in this study.

Supplementary references

Supplementary table 1. List of organisms used for phylogenetic analysis

Organism	Phylum	Source	Sequence ID	Gene name
<i>Escherichia coli</i> DH1	Bacteria	1	ACX39449.1	EDD
<i>Rhizobium leguminosarum</i>	Bacteria	1	WP_116272381.1	EDD
<i>Zymomonas mobilis</i>	Bacteria	2	AEH62744.1	EDD
<i>Salmonella enterica</i>	Bacteria	1	AGS29956.1	EDD
<i>Acinetobacter baumannii</i>	Bacteria	1	WP_031976293.1	EDD
<i>Pseudomonas fluorescens</i>	Bacteria	1	SNY07451.1	EDD
<i>Sinorhizobium meliloti</i>	Bacteria	1	SDY78282.1	EDD
<i>Agrobacterium tumefaciens</i>	Bacteria	1	KJX89456.1	EDD
<i>Xanthomonas campestris</i>	Bacteria	1	CEM58200.1	EDD
<i>Cereibacter sphaeroides</i> KD131	Bacteria	1	ACM00817.1	EDD
<i>Dinoroseobacter shibae</i>	Bacteria	1	ABV93511.1	EDD
<i>Chloroflexota bacterium</i>	Bacteria	1	TME26042.1	EDD
<i>Anaerolineae bacterium</i>	Bacteria	1	CAG0952899.1	EDD
<i>Phaeobacter gallaeciensis</i>	Bacteria	1	AHD08411.1	EDD
<i>Acidiphilium cryptum</i> JF-5	Bacteria	1	ABQ29690.1	EDD
<i>Empedobacter haloabium</i>	Bacteria	1	TXE25295.1	EDD
<i>Cryomorphaceae bacterium</i>	Bacteria	1	MCF8220629.1	EDD
<i>Fluviicola</i> sp.	Bacteria	1	PHR35100.1	EDD
<i>Luteimonas aquatica</i>	Bacteria	1	WP_242107788.1	EDD
<i>Stenotrophomonas maltophilia</i>	Bacteria	1	WP_111190739.1	EDD
<i>Xylella fastidiosa</i>	Bacteria	1	WP_023907921.1	EDD
<i>Shigella flexneri</i>	Bacteria	1	EFY2669691.1	EDD
<i>Pseudoxanthomonas</i> sp.	Bacteria	1	WP_141453109.1	EDD
<i>Crocinitomicaceae bacterium</i>	Bacteria	1	MBT6514710.1	EDD
<i>Flavobacteriales bacterium</i>	Bacteria	1	MAC86373.1	EDD
<i>Amycolatopsis decaplanina</i>	Bacteria	1	WP_007029288.1	EDD
<i>Amycolatopsis keratiniphila</i>	Bacteria	1	WP_016332713.1	EDD
<i>Nocardioides</i> sp. MAH-18	Bacteria	1	WP_157345696.1	EDD
<i>Rhodococcus</i> sp. JVH1	Bacteria	1	WP_009475031.1	EDD
<i>Nocardioides panaciterrulae</i>	Bacteria	1	WP_218852091.1	EDD
<i>Cryobacterium glucosi</i>	Bacteria	1	WP_166791009.1	EDD
<i>Prauserella endophytica</i>	Bacteria	1	WP_137095839.1	EDD
<i>Prauserella flavalba</i>	Bacteria	1	WP_110340288.1	EDD
<i>Longispora fulva</i>	Bacteria	1	WP_203787466.1	EDD
<i>Hoyosella altamirensis</i>	Bacteria	1	WP_064440734.1	EDD
<i>Mycobacterium tuberculosis</i>	Bacteria	1	SGD88927.1	EDD
<i>Shigella dysenteriae</i>	Bacteria	1	EGE2518530.1	EDD
<i>Hoyosella subflava</i>	Bacteria	1	WP_237707834.1	EDD
<i>Enterobacteriaceae</i>	Bacteria	1	WP_001069487.1	EDD
<i>Prauserella flavalba</i>	Bacteria	1	WP_110340288.1	EDD
<i>Tenacibaculum</i> sp. KUL152	Bacteria	1	GFD89636.1	EDD
<i>Chryseobacterium</i> sp. ISL-80	Bacteria	1	MBT2704122.1	DHAD
<i>Rhizobium leguminosarum</i>	Bacteria	1	WP_129418895.1	DHAD

<i>Escherichia coli</i>	Bacteria	1	YP_026248.1	DHAD
<i>Mycobacterium tuberculosis</i>	Bacteria	1	SGC94481.1	DHAD
<i>Anaerolineae bacterium</i>	Bacteria	1	MCU0501630.1	DHAD
<i>Nocardioides sp. Iso805N</i>	Bacteria	1	WP_017934119.1	DHAD
<i>Streptomyces cinereus</i>	Bacteria	1	GGL96102.1	DHAD
<i>Enterobacteriaceae</i>	Bacteria	1	WP_001127401.1	DHAD
<i>Shigella boydii</i>	Bacteria	1	EAA4813663.1	DHAD
<i>Chamaesiphon polymorphus</i>	Bacteria	1	WP_106308272.1	DHAD
<i>Euryarchaeota archaeon</i>	Archaea	3	MBE02805.1	EDD
<i>Candidatus Woesearchaeota archaeon</i>	Archaea	1	MBM3202515.1	EDD
<i>Euryarchaeota archaeon</i>	Archaea	4	MBE01529.1	DHAD
<i>Candidatus Nitrosopumilus salaria</i>	Archaea	1	WP_008300640.1	EDD
<i>Ferroglobus placidus</i>	Archaea	1	WP_012966058.1	DHAD
<i>Candidatus Pacearchaeota archaeon</i>	Archaea	1	RLG12971.1	DHAD
<i>Candidatus Bathyarchaeota archaeon</i>	Archaea	5	HDD72113.1	DHAD
<i>Crenarchaeota archaeon</i>	Archaea	6	NPA70637.1	DHAD
<i>Ignicoccus hospitalis</i>	Archaea	1	WP_012123105.1	DHAD
<i>Desulfurococcaceae archaeon</i>	Archaea	1	MCC6057874.1	DHAD
<i>Thaumarchaeota archaeon</i>	Archaea	1	RLG07732.1	DHAD
<i>Archaeoglobi archaeon</i>	Archaea	1	MBC7109545.1	DHAD
<i>Thermoprotei archaeon</i>	Archaea	1	RLF07682.1	DHAD
<i>Methanoregulaceae archaeon</i>	Archaea	1	RPI39965.1	DHAD
<i>Archaeoglobi archaeon</i>	Archaea	3	MBC7114307.1	DHAD
<i>Methanosarcina barkeri</i>	Archaea	1	WP_048108734.1	DHAD
<i>Archaeoglobus profundus</i>	Archaea	1	WP_012940844.1	DHAD
<i>Sulfolobus sp.</i>	Archaea	7	PVU75057.1	DHAD
<i>Sulfolobaceae</i>	Archaea	1	WP_009990927.1	DHAD
<i>Aeropyrum pernix</i>	Archaea	1	WP_241759698.1	DHAD
<i>Stygiolobus caldivivus</i>	Archaea	1	WP_221290504.1	DHAD
<i>Acidianus hospitalis</i>	Archaea	1	WP_013777133.1	DHAD
<i>Candidatus Wolframiraptor sp.</i>	Archaea	1	RLF99771.1	DHAD
<i>Synechococcus elongatus</i>	Cyanobacteria	1	WP_011243210.1	DHAD
<i>Synechococcus moorigangaii CMS01</i>	Cyanobacteria	8	MBV5257376.1	EDD
<i>Nostoc sp. 3335mG</i>	Cyanobacteria	1	PXA97457.1	EDD
<i>Leptolyngbya valderiana BDU 20041</i>	Cyanobacteria	1	OAB60313.1	EDD
<i>Leptolyngbya sp. 15MV</i>	Cyanobacteria	9	QYU66659.1	EDD
<i>Planktothrix paucivesiculata</i>	Cyanobacteria	1	WP_083620833.1	DHAD
<i>Pseudanabaena sp. lw0831</i>	Cyanobacteria	1	WP_201322407.1	DHAD
<i>Oscillatoriales bacterium UBA8482</i>	Cyanobacteria	1	HBW56505.1	DHAD
<i>Dendronalium phyllosphericum</i>	Cyanobacteria	1	WP_214433426.1	DHAD
<i>Nostocales cyanobacterium</i>	Cyanobacteria	1	TAE57107.1	DHAD
<i>Gracilaria domingensis</i>	Cyanobacteria	1	KAI0558924.1	DHAD
<i>Phormidium sp. ETS-05</i>	Cyanobacteria	1	WP_199245539.1	DHAD
<i>Chondrus crispus</i>	Cyanobacteria	1	XP_005718290.1	DHAD
<i>Adonisia turfae</i>	Cyanobacteria	1	WP_163667036.1	DHAD
<i>Anabaena minutissima</i>	Cyanobacteria	1	WP_190699939.1	DHAD
<i>Nostoc parmelioides</i>	Cyanobacteria	1	WP_190569724.1	DHAD

<i>Trichormus azollae</i>	Cyanobacteria	1	WP_013189959.1	DHAD
<i>Nostoc sp. UHCC 0702</i>	Cyanobacteria	1	WP_206268843.1	DHAD
<i>Cylindrospermopsis raciborskii</i>	Cyanobacteria	1	WP_006277049.1	DHAD
<i>Gloeobacter violaceus</i>	Cyanobacteria	1	WP_011143160.1	DHAD
<i>Synechococcus sp. JA-2-3B'a(2-13)</i>	Cyanobacteria	1	WP_011433574.1	DHAD
<i>Pseudanabaena sp. PCC 7367</i>	Cyanobacteria	1	WP_015164194.1	DHAD
<i>Pseudanabaena sp. PCC 6802</i>	Cyanobacteria	1	WP_019501081.1	DHAD
<i>Gloeomargarita sp. SKYB120</i>	Cyanobacteria	1	MCS7293203.1	DHAD
<i>Gloeomargarita lithophora</i>	Cyanobacteria	10	WP_071455593.1	DHAD
<i>Calothrix membranacea</i>	Cyanobacteria	1	WP_190617033.1	DHAD
<i>Picochlorum sp. BPE23</i>	Algae	1	KAI8114006.1	hypothetical protein
<i>Chlamydomonas eustigma</i>	Algae	11	GAX72671.1	hypothetical protein
<i>Chlorella sorokiniana</i>	Algae	12	PRW61595.1	DHAD
<i>Polytomella parva</i>	Algae	1	QKY14824.1	DHAD
<i>Micractinium conductrix</i>	Algae	13	PSC73296.1	DHAD
<i>Coccomyxa sp. Obi</i>	Algae	1	BDA41163.1	DHAD
<i>Chlamydomonas reinhardtii</i>	Algae	14	XP_001693179.1	uncharacterized protein
<i>Volvox africanus</i>	Algae	15	GIL51840.1	hypothetical protein
<i>Ostreococcus tauri</i>	Algae	1	OUS43982.1	DHAD/EDD
<i>Micromonas pusilla CCMP1545</i>	Algae	1	XP_003055297.1	predicted protein
<i>Mougeotiopsis calospora</i>	Algae	16	PRJNA849386	TSA
<i>Mougeotia sp</i>	Algae	17	PRJNA543475	TSA
<i>Zygnema circumcarinatum</i>	Algae	18	PRJNA399177	TSA
<i>Spirogyra pratensis</i>	Algae	17	PRJNA543475	TSA
<i>Spirogyra pratensis</i>	Algae	19	PRJNA242253	TSA
<i>Spirogyra pratensis</i>	Algae	20	PRJNA335946	TSA
<i>Chara braunii</i>	Algae	21	GBG73628.1	hypothetical protein
<i>Sphagnum magellanicum</i>	Bryophytes	1	KAH9533862.1	hypothetical protein
<i>Sphagnum fallax</i>	Bryophytes	1	KAH8933315.1	hypothetical protein
<i>Physcomitrium patens</i>	Bryophytes	1	XP_024389039.1	DHAD
<i>Ceratodon purpureus</i>	Bryophytes	1	KAG0623704.1	hypothetical protein
<i>Tieghemiomyces parasiticus</i>	Fungi	1	KAJ1927328.1	DHAD
<i>Dimargaris cristalligena</i>	Fungi	1	KAJ1988464.1	DHAD
<i>Hyaloraphidium curvatum</i>	Fungi	1	KAI9021181.1	DHAD
<i>Anaeromyces robustus</i>	Fungi	1	ORX87668.1	DHAD
<i>Phlyctochytrium arcticum</i>	Fungi	1	KAI9102494.1	DHAD
<i>Thamnocephalis sphaerospora</i>	Fungi	1	RKP07992.1	DHAD
<i>Viridothelium virens</i>	Fungi	1	KAF2233468.1	DHAD
<i>Fonsecaea erecta</i>	Fungi	1	XP_018695233.1	DHAD
<i>Crucibulum laeve</i>	Fungi	1	TFK40805.1	DHAD
<i>Polychytrium aggregatum</i>	Fungi	1	KAI9202125.1	DHAD
<i>Blastomyces parvus</i>	Fungi	1	PGH09273.1	DHAD
<i>Schizosaccharomyces pombe</i>	Fungi	22	NP_593729.1	Putative DHAD
<i>Saccharomyces cerevisiae S288C</i>	Fungi	1	NP_012550.1	DHAD
<i>Elsinoe ampelina</i>	Fungi	1	KAF2221669.1	DHAD
<i>Manihot esculenta</i>	Plant	1	XP_021594917.1	DHAD

<i>Dioscorea alata</i>	Plant	¹	KAH7669517.1	DHAD
<i>Nicotiana tabacum</i>	Plant	¹	XP_016464930.1	DHAD
<i>Capsicum annuum</i>	Plant	¹	XP_016573626.2	DHAD
<i>Populus alba</i>	Plant	¹	XP_034922072.1	DHAD
<i>Jatropha curcas</i>	Plant	¹	XP_012092653.1	DHAD
<i>Beta vulgaris subsp. vulgaris</i>	Plant	¹	XP_019104709.1	DHAD
<i>Arachis duranensis</i>	Plant	¹	XP_015958659.1	DHAD
<i>Arachis hypogaea</i>	Plant	¹	XP_025693829.1	DHAD
<i>Citrus sinensis</i>	Plant	¹	XP_006489797.2	DHAD
<i>Solanum lycopersicum</i>	Plant	¹	XP_004240214.1	DHAD
<i>Coffea arabica</i>	Plant	¹	XP_027116753.1	DHAD
<i>Glycine max</i>	Plant	²³	KAH1217623.1	DHAD
<i>Catharanthus roseus</i>	Plant	¹	KAI5650029.1	hypothetical protein
<i>Phaseolus vulgaris</i>	Plant	¹	XP_007147759.1	hypothetical protein
<i>Salix suchowensis</i>	Plant	¹	KAG5227902.1	DHAD
<i>Ziziphus jujuba var. spinosa</i>	Plant	¹	XP_015887506.1	DHAD
<i>Zingiber officinale</i>	Plant	²³	XP_042382229.1	DHAD

DHAD, dihydroxy-acid dehydratase; EDD, 6-phosphogluconate dehydratase;
TSA, transcriptome sequence assembly

Supplementary table 2. Occurrence of EDD homologs in genomes of plastid sister groups^a

Organism	DHAD	EDD	EDA	Reference
Recently diverged plastid sister groups				
<i>Gloeomargarita lithophora</i>	WP_071455593	n.d. ^b	WP_071454108	10,24
<i>Gloeobacter kilaueaensis</i>	WP_023175106	n.d.	WP_023172264	24
<i>Synechococcus elongatus</i>	WP_208679212	n.d.	WP_208678531	24
<i>Synechocystis</i> sp. 6803	WP_010874288	n.d.	AGF52924	
<i>Synechocystis</i> sp.	WP_194072189	n.d.	WP_237269654	24
<i>Pseudanabaena</i> sp	WP_094530374	n.d.	PZU96335	10
<i>Arthrospira</i> sp.	WP_006616382	n.d.	MBS0018420	25
<i>Lyngbya</i> sp.	WP_265233541	n.d.	MBW4695200	25
<i>Oscillatoria</i> sp.	WP_272061443	n.d.	WP_017718781	25
<i>Trichodesmium</i> sp.	MDJ0516496	n.d.	MCL2928218	25
<i>Prochlorococcus marinus</i>	WP_269610147	n.d.	WP_011124409	25
<i>Anabaena</i> sp.	WP_066383956	n.d.	WP_190367598	25
<i>Nodularia spumigena</i>	WP_042202732	n.d.	WP_063872911	25
<i>Cylindrospermopsis raciborskii</i>	WP_071241830	n.d.	WP_009344397	25
<i>Cyanothece</i> sp.	NEQ98939	n.d.	WP_107670816	25
<i>Acaryochloris marina</i>	WP_299492222	n.d.	WP_212665323	25
<i>Geitlerinema</i> sp.	WP_015171406	n.d.	WP_015173050	25
<i>Prochlorothrix hollandica</i>	WP_026099305	n.d.	WP_017713606	25
EDD homologs in cyanobacteria				
<i>Nostoc</i> sp.	WP_015141122	PXA97457	PXA92509	25
<i>Synechococcus moorigangaii</i> CMS01	MBV5262251	MBV5257376	MBV5257377	
<i>Leptolyngbya valderiana</i> BDU 20041	OAB57411	OAB60313	OAB60315	
<i>Leptolyngbya</i> sp.	WP_198806192	QYU66659	WP_204152276	25

^a All sequences obtained from NCBI (<https://www.ncbi.nlm.nih.gov/>)^b n.d., not detected

Supplementary table 3. PCR primers used in this study. Primers include attB overhang for gateway cloning

Primer name	Sequence (5'→3')
GmKDPG1_attB1	GGGGACAAGTTTGTACAAAAAAGCAGGCTATGGGCACTGCAACTGCAACC
GmKDPG1_attB1_pm	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCACAGTTGACAGAACCCTCTGC
GmKDPG1_attB2	GGGGACCACTTTGTACAAGAAAGCTGGGTATAAATTCTATTCACAGATTTGTT ACCC
GmKDPG1stp_attB2	GGGGACCACTTTGTACAAGAAAGCTGGGTCTAATAAATTCTATTCACAGATTT GTTACCC
GmDHAD_attB1	GGGGACAAGTTTGTACAAAAAAGCAGGCTATGCAGTCCACACTCTTCAACCC
GmDHAD_attB1_pm	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCGTGAAGCTGAACAAGTACAGC TCC
GmDHAD_attB2	GGGGACCACTTTGTACAAGAAAGCTGGGTCTCGTCTGTTACGCATCCAC
GmDHADstp_attB2	GGGGACCACTTTGTACAAGAAAGCTGGGTCTACTCGTCTGTTACGCATCCAC
EcKDPG_attB1	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGAAAACTGGAAAACA
EcKDPGstp_attB2	GGGGACCACTTTGTACAAGAAAGCTGGGTCTTACAGCTTAGCGCCTTC
EcEDD_attB1	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGAATCCACAATTGTTA
EcEDDstp_attB2	GGGGACCACTTTGTACAAGAAAGCTGGGTCTTAAAAAGTGATACAGGT
EcDHAD-attB1	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCACATGCCTAAGTACCGTTCCG CC
EcDHADstp-attB2	GGGGACCACTTTGTACAAGAAAGCTGGGTGTTAACCCCCCAGTTTCGATTTA TCG

Supplementary figure 1. Amino acid alignment of representative DHAD and EDD sequences

DHAD_Ec	-----M V V I W E V L V R C G P T G M	16
DHAD_Se	-----M A G A R A L W R A T --G M	13
DHAD_Ab	-----M A G A R A L W R A T --G M	13
DHAD_Tm	-----M R S D V I K K G L E R V P H S L K A L --G I	24
DHAD_Syne	-----M P Q Y R S R T T T Y G R N M A G A R A L W R A T --G M	27
DHAD_No	-----M P A Y R S R T T T H G R N M A G A R A L W R A T --G M	27
DHAD_At	M Q A T I F S P R A T L F P C K P L L P S H N V N S R R P S I S C A S Q S V T A D P S P P I T D T N K L N K Y S S R I T E P K S Q G G S Q A I L H G V --G L	78
DHAD_Gm	M Q S T L F N P T H S L I P T S P H S I R S N S G H A S L S V R A S I A V E T P -----T E T V K L N K Y S S R I T E P K S Q G A S Q A V L Y G V --G L	71
EDD_No	-----M T V R Q S I R D V T D R I A A -----R S R D S R R D Y L K R I E A A R E A G V Y E S S L S C G --N L	47
EDD_Synm	-----M P A P L H P V I E E V T A R I E A -----R S R S T R A A Y L S L I A E A R K D G P O K I L S C G --N L	49
EDD_Wa	-----M I V H P T L L A V T D R I R A -----R S Q H R R D A Y L K R V Q A A Y Q S G V T R H R I N C T --N I	47
EDD_Ec	-----M N P Q L L R V T N R I I E -----R S R E T R S A Y L A R I E Q A K T S T V H S Q L A C G --N L	45
EDD_Zm	-----M T D L H S T V E K V T A R V I E -----R S R E T R K A Y L D L I Q Y E R E K G V D P N L S C S --N L	48
EDD_Se	-----M N P N L L R V T Q R I V E -----R S Q Q T R E A Y L A R I E Q A K T A T V H S Q L A C G --N L	45
EDD_Ab	-----M D L P N P I L A K V T E R V I A -----R S Q K T R S A Y L Q R I E H A Q G K F P A R G A L S C A --N L	48
EDD_Agro	-----M S A D S R I Q A I T A R I V E -----R S K P Y R E T Y L E R L R L Q V S K G V H S V L S C G --N L	47

DHAD_Ec	TD-----A D F G K P I T A V V N S F T Q F V P G H V H L R D L G K I V A E Q I E A A G G V A K E F N T I -A V D D G I A M G H G G M L Y S	82
DHAD_Se	TD-----S D F G K P I T A V V N S F T Q F V P G H V H L R D L G K I V A E Q I E A S G G V A K E F N T I -A V D D G I A M G H G G M L Y S	79
DHAD_Ab	KD-----E D F G K P I T A V V N S F T Q F V P G H V H L K D L G Q I V A A E I Q A A G G V A K E F N T I -A V D D G I A M G H G G M L Y S	79
DHAD_Tm	TD-----D E M R R P F T G I V S S W N E I I P G H V H L D K V V E A V K A G V R M A G G V P F V F P T I -G I C D G I A M D H R G M K F S	90
DHAD_Syne	KD-----E D F E K P I T A V A N S F T Q F V P G H V H L K D L G Q I V A R E I E R A G G V A K E F N T I -A V D D G I A M G H G G M L Y S	93
DHAD_No	KD-----S D F G K P I T A V A N S F T Q F V P G H V H L K D L G Q I V A R E I E A A G G V A K E F N T I -A V D D G I A M G H G G M L Y S	93
DHAD_At	SD-----D D L L K P Q T G I S S V W Y E G N T C N M H L L K L S E A V K E G V E N A C M V G F R F N T I -G V S D A I S M C T R G M C F S	144
DHAD_Gm	SE-----D D M A K P Q V G V S V W Y E G N T C N M H L L H L S E A V R D G V A A A C M V P F R F N T V -G V S D A I S M C T R G M C Y S	137
EDD_No	A H G F A A C T P A E K A A L A G N K T L N I G I V T S Y N D M L S A H Q P Y Q Y P D I I K E A A R E M C A T A Q V A G G V P A M C D G V T Q S Q P G M D L S	127
EDD_Synm	A H G F A A -S G E D K E K L S L V N P N I A I V T A Y N D M L S A H Q P L G A Y P D I I K A A A R E A C A T A Q I A G G V P A M C D G V T Q S Q A G M B L S	128
EDD_Wa	A H A Y A A L P V C D K G L F K A E V A P N I G I V T A F N D M L S A H Q P Y E N Y P A L I K D E V R K V S A T A Q V A G G V P A M C D G V T Q S Q T G M B L S	127
EDD_Ec	A H G F A A C Q P E D K A S L K S M L R N N I A I I T S Y N D M L S A H Q P Y E H Y P E I I R K A L H E A N A V G Q V A G G V P A M C D G V T Q S Q D G M B L S	125
EDD_Zm	A H G F A A M N G -D K P A L R D F N R M N I G V V T S Y N D M L S A H E P Y Y R P E Q M K V F A R E V S A T V Q V A G G V P A M C D G V T Q S Q P G M B E S	127
EDD_Se	A H G F A A C Q P E D K A S L K S M L R N N I A I I T S Y N D M L S A H Q P Y E H Y P Q I I R Q A L H S V N A V G Q V A G G V P A M C D G V T Q S Q D G M B L S	125
EDD_Ab	A H G F A G M E D N E K L I K V G R E P N I G I V S S Y N E M L S A H A P Y P F T P D I I K T A A R E N G G V A Q F A G G V P A M C D G I T Q S Q P G M B L S	128
EDD_Agro	A H G F A V C S P A D K I L A G D R V P N I G I I T A Y N D M L S A H Q P Y E T F P A I I R D A A K E A G G I A Q V A G A V P A M C D G V T Q S Q P G M B L S	127

DHAD_Ec	L P S R E I I A D S V E Y M V N A H C A D A M V C I S N C D K I T P G M I M A A R L N I P V I E V S G G G M E A S K T K L S D Q I I K L D L V D A M I Q G A	161
DHAD_Se	L P S R E I I A D S V E Y M V N A H C A D A M V C I S N C D K I T P G M I M A S R L N I P V I E V S G G G M E A S K T K L S D K I I K L D L V D A M I T Q G A	158
DHAD_Ab	L P S R E I I A D S V E Y M V N A H C A D A M V C I S N C D K I T P G M I M A A R L N I P V I E V S G G G M E A S K V K F R G D K A I D L V D A M V A A A	158
DHAD_Tm	L P S R E I I A D S I E I V A S G F P F D G L V F V P N C D K I T P G M M M A M G R L N I P S V L I S G G G M L A G R Y N -G R D I D L I T V F E A V G G Y K	168
DHAD_Syne	L P S R D I I A D S V E Y M V N A H C A D A L V C I S N C D K I T P G M I M A A R L N I P A V E V S G G G M E A S K V I L N G E E R H L D L V D A M V V A A	172
DHAD_No	L P S R E I I A D S V E Y M V N A H C A D A M V C I S N C D K I T P G M I M A A R L N I P A V E V S G G G M E A S K V V V K G E V A L D L V D A M V A A A	172
DHAD_At	L Q S R D I I A D S I E T V M S A Q W Y D G N I S I P G C D K N P G T I I A M G R L N R P C I M V Y G G T I K P G H F Q -D K T Y D I V S A F Q S Y G E F V	222
DHAD_Gm	L Q S R D I I A D S I E T V M A A Q W Y D G N I S I P G C D K N P G T I I A M G R L N R P S I M V Y G G T I K P G H F E -G N T F D I V S A F Q S Y G E Y V	215
EDD_No	L D S Y T D E E V K T I E A F E R S A C P T G G S C S G M F T A N S M N C L T E A L G S L P C N G S I V A T H A N R K K F L K A G Q L I V E A K R Y Y E Q D D A	200
EDD_Synm	L P S R D I A I M A T A V S L S H D M F D A A I C L G V C D K I V P G L I I G A L R F G H I P V E L P A G E M R S G L P N -T E K A R I R Q L -----Y A	201
EDD_Wa	L P S R D I A I L S T A V A L S H H V F D A N I F L G V C D K I V P G L I I G A L R F G H I P A V E V P A G E M T T G I S N -A D K A R V I R Q L -----Y A	200
EDD_Ec	L D S R E V I A M S A A V G L S H N M F D G A I F L G V C D K I V P G L I I A A A S F G H I P A V E V P S G G M A S G L P N -K E K V I R I R Q L -----Y A	198
EDD_Zm	L P S R D I A I L A T S V S L S H G M F E G A A L I C I C D K I V P G L I I G A L R F G H I P T I L V P S G G M T T G I P N -K E K I R I R Q L -----Y A	200
EDD_Se	L D S R E V I A M S A A V G L S H N M F D G A I F L G V C D K I V P G L I I A A A S F G H I P A I E V P S G G M A S G L P N -K E K V I R I R Q L -----Y A	198
EDD_Ab	L P S R E T I A M S T A I A L S H N M F D A A I C L G V C D K I V P G L I I G A L Q F G Y L P T I E V P A G M S S G L S N -D D K A K I R Q Q -----F A	201
EDD_Agro	L P S R D A I A M A A G I G L S H N M F D A A Y L G V C D K I V P G L I I A A L A F G H I P A V E V P A G E M T S G L P N -D E K S R I R Q L -----Y A	200

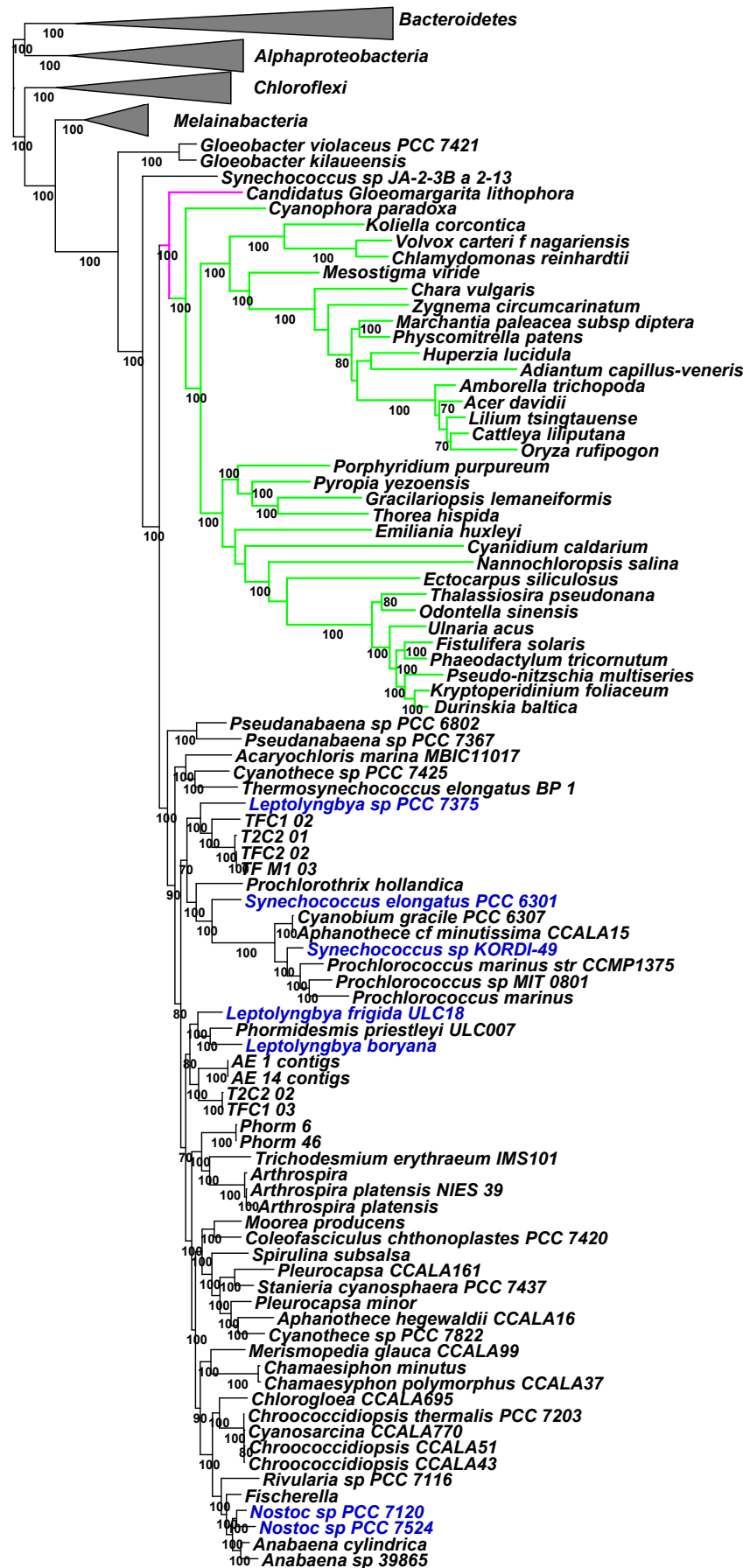
DHAD_Ec	D P K V S D S Q S D Q V E R S A C P T G G S C S G M F T A N S M N C L T E A L G S L P C N G S I L A T H A D R K Q F L N A G K R I V E T K R Y Y E Q N D E	241
DHAD_Se	D P K V S D D Q S N Q V E R S A C P T G G S C S G M F T A N S M N C L T E A L G S L P C N G S I L A T H A D R K Q F L N A G K R I V E T K R Y Y E Q N D E	238
DHAD_Ab	D D S Y T D E E V K T I E A F E R S A C P T G G S C S G M F T A N S M N C L T E A L G S L P C N G S I V A T H A N R K K F L K A G Q L I V E A K R Y Y E Q N D A	238
DHAD_Tm	V G K V D E E T I K A I E D L A C P G A G S C A G L F T A N T M N S A E A L G S L A P R E N G T V P A V H A K R L R A K E A G M L V V E V K R -----	241
DHAD_Syne	D D R E S D E D V A T I E R S A C P T G G S C S G M F T A N S M N C L T E A L G S L P C N G S I L A T H G D R K E F L E A G R L A V K A K R Y Y E Q D D E	252
DHAD_No	D D S Y T D E E V K T I E R S A C P T G G S C S G M F T A N S M N C L T E A L G S L P C N G S T L A T H A D R E K F R E A G H L I V D A K R Y Y E Q D D A	252
DHAD_At	S G S I S D E Q R K T V L H H S C P G A G A C G G M Y T A N T M A S A I E A M G S L F Y S S S P A E D P I K L D E C R L A G K Y L L E L K M -----	295
DHAD_Gm	S G S I N D D Q R Q N V I R N S C P G A G A C G G M Y T A N T M A S A I E A M G S L F Y S S S T P A E D P I K L D E C R L A G K Y L L E L K M -----	288
EDD_No	E G K V G R A E I L L E A S K S Y H S A G T C T F Y G T A N S N Q M M E V M G H L P G S A F V N P G T P I R D A T R E A N R A L A T A Q -----G N	275
EDD_Synm	E G K A G R D Q L L D A E S A S Y H S P G T C T F Y G T A N S N Q M M E V M G H L P G S A F V N P G T P I R E A V K A T A R R A A K T A L -----G Q	276
EDD_Wa	E G K A D R E A L L E S M A S Y H S A G T C T F Y G T A N S N Q M M E V M G H L P G S S F V N P G T P I R E A T R H A R L A A H A E S -----K	274
EDD_Ec	E G K V D R M A L L E S A A S Y H A P G T C T F Y G T A N T N Q M V E V M G H L P G S S F V H P D A P I R D A T A A A R Q V T R M T G N -----G N	273
EDD_Zm	Q G K I G Q K E I L D M E A A C Y H A B G T C T F Y G T A N T N Q M V E V M G H M P G S A F V T P G T P I R Q A T R A V H R V A E L G W K -----G D	275
EDD_Se	E G K V D R M A L L E S A A S Y H A P G T C T F Y G T A N T N Q M V E V M G H L P G S S F V H P D A P I R E A T A A A R Q V T R M T G N -----G N	273
EDD_Ab	T G Q V G R D A L L E A S A A Y H G G G T C T F Y G T A N S N Q M M E V M G H L P S A A F V H P H T P I R D A T A E A K R V L D E T A E -----R G	276
EDD_Agro	E G K V G R A E I L L E A S K S Y H G G G T C T F Y G T A N S N Q M M E V M G H M P G S S F V N P G T P I R D A T R E A N R A L A T A Q -----G N	275

DHAD_Ec	S A L P -R N I A S K A X F E N A M T L D I A M G G S T N T V L H L L A A A Q E A E I D F T M S I D K L S R K V P O L C K V A P S T Q K Y H M E D V H R A G G	320
DHAD_Se	S A L P -R N I A S K A A F E N A M T L D I A M G G S T N T V L H L L A A A Q E A E I D F T M S I D K L S R K V P O L C K V A P S T Q K Y H M E D V H R A G G	317
DHAD_Ab	S I L P -R S I A T K A A F K N A M T L D I A M G G S T N T V L H L L A A A N E A E V D F T M D I D E L S R R V P L S K V A P A K Q D V H M E D V H R A G G	317
DHAD_Tm	D V K P -R D I T L D S F M N A V M V D L A T G G S T N T V L H L L A I A E S P E I D F D I K L F D E L S R K I P H I C N I S P V -G P Y H I Q D L D D A G G	319
DHAD_Syne	S V L P -R S I A S F K A F E N A I C L D I A M G G S T N T V L H L L A A A H E A C V D F T M K I D R L S R K I P N L C K V A P S T Q K Y H M E D V H R A G G	331
DHAD_No	T A L P -R G I A T F A A F E N A M S L D I A M G G S T N T V L H L L A A A Q E G G V E F T M A I D R L S R R V P L S K V A P A K S D V H M E D V H R A G G	331
DHAD_At	D L K P -R D I T P K S L R N A M V S V M A L G G S T N A V L H L L A I A R S V G L E I T L D F F Q K V S D A V P L A D L K P S -G K Y V M E D I H K I G G	373
DHAD_Gm	D L K P -R D I T R K S L R N A M V I V M A L G G S T N A V L H L L A I A K A S V G I D I L T D F F Q K V S D E V P L A D L K P S -G K Y V M E D V H K I G G	366
EDD_No	D Y T P I G H I I D E K A I V N G L V G L H A T G G S T N H T M H L V A A A A A G L Q V T W D M S L S D A T P L A R V Y P N -G V A D V N H F H A A G G	354
EDD_Synm	D Y R P I G E A I D A R A I N A I V G L N A T G G S T N H T I H L V A A R A A G I L I N W D F F E D I S H V T P L L T K I Y P N -G P A D V N H F H A A G G	355
EDD_Wa	A A L S I C D I V D E R V I V N A M V G L L A T G G S T N H T I H L V A A R A A G I L I N W D F F S L S A D V P L A R V Y P N -G K A D I N H F H A A G G	353
EDD_Ec	E W M P I G K M I D E K V V V N G I V A L L A T G G S T N H T M H L V A A R A A G I Q I N W D F S D L S D V P L A R I Y P N -G P A D I N H F Q A A G G	352
EDD_Zm	D Y R P I G K I I D E K S I V N A I V G L L A T G G S T N H T M H I A A R A A G I V I N W N D F H D L S E V V P L A R I Y P N -G P R D I N E F Q A A G G	354
EDD_Se	T W M P L G M I I D E K S I V N G I V A L L A T G G S T N H T M H L V A A R A A G I L I N W D F F S L S E V V P L A R I Y P N -G P A D I N H F Q A A G G	352
EDD_Ab	N Y T P I G H V I D E K A I I N G I V A L L A T G G S T N H T L H L I A A R A A G I L I D W D F F D E L S A V V P L A K I Y P N -G K A D V N H F Q A A G G	355
EDD_Agro	E F T P A G E M I D E R S V V N G V V G L H A T G G S T N H T M H L I A A R A A G I I L T W Q I S I L S D I V P L A R V Y P N -G L A D V N H F H A A G G	354

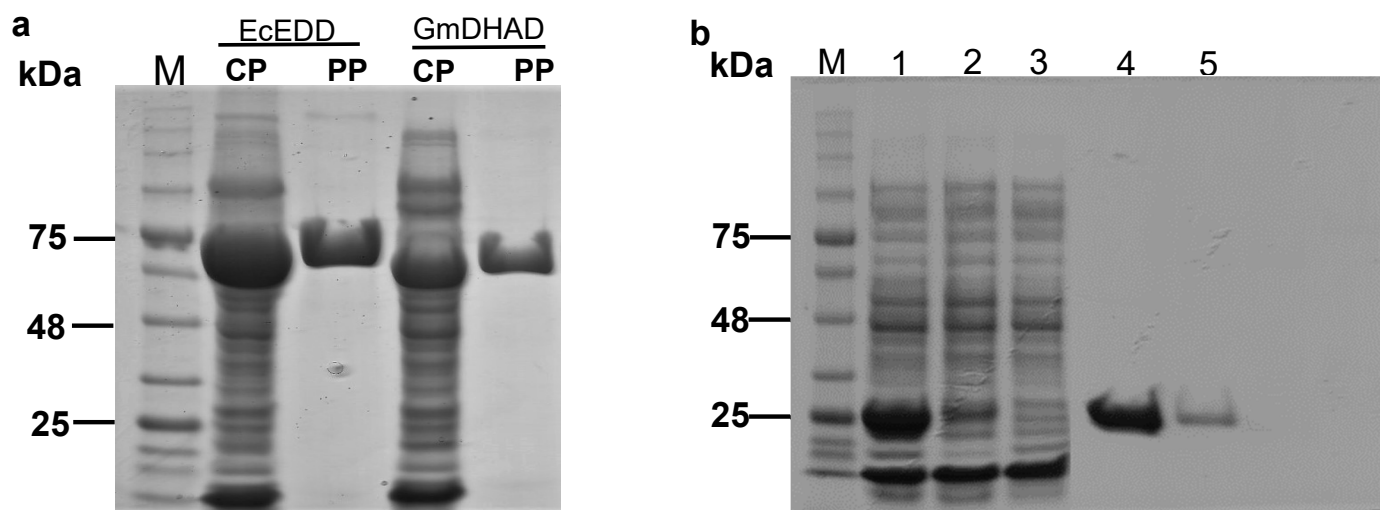
DHAD_Ec	RSLEHA	Y	SKDGG	LAVL	YGNFA	ENGCI	VKTAG	Y	DDSI	LKFTG	PAKV	YES	ODDA	VEIL	GKKV	V	AGDV	VVIR	EGEK	CG	475										
DHAD_Se	RSLEYA	Y	SKDGG	LAVL	YGNFA	ENGCI	VKTAG	Y	DDSI	LKFTG	PAKV	YES	ODDA	VEIL	GKKV	V	EGDV	VVIR	EGEK	CG	472										
DHAD_Ab	RNAEHA	Y	SKDGG	LAVL	YGNIA	LDGCI	VKTAG	Y	DESIL	KFTGS	PARV	ES	ODAA	VEIL	GNK	V	AGDV	VVIR	EGEK	CG	472										
DHAD_Tm	RPFDNF	Y	HKEG	LGLIF	GNLA	PEGA	WAKI	SGV	PEKMM	HV	GP	AV	EDG	EAET	LIS	GK	V	KGDV	VVIR	EGEK	CG	437									
DHAD_Syne	RDIEHA	Y	SDGG	LAVL	YGNLA	EGCII	VKTAG	Y	DEIN	LVFSG	PAV	VCES	ODAE	VN	IL	GNK	V	EGDV	WLIR	EGEK	CG	487									
DHAD_No	RDVEHA	Y	SKDGG	LAVL	YGNLA	PEGCI	VKTAG	Y	DESIL	TFRGT	PARV	ES	ODAS	VAGI	LGNE	VK	EGDV	VVIR	EGEK	CG	486										
DHAD_At	RPLSN	PIK	TH	QIL	GLD	PA	DGS	WAKI	ITG	KEGLY	FSG	PA	LV	EGE	EM	IA	ISADP	SFK	GT	VVIR	EGEK	CG	492								
DHAD_Gm	RPVEN	PIK	THA	IL	Y	GLN	LA	PQGS	WAKI	ITG	KEGLY	FSG	PA	LV	EGE	EM	IA	ISD	PSFK	GK	VVIR	EGEK	CG	485							
EDD_No	TSARTP	Y	QOTG	LK	LIT	GNL	GR	Y	SVIK	SA	V	PEH	RV	VEA	PA	RV	HQD	GLIA	FA	KAGE	IT	GDV	IAV	VVR	EG	SG	KAI	G	486		
EDD_Synn	RPASDP	Y	PIAPD	GGLR	VVAG	L	GR	Y	AVVK	SA	V	PEH	RV	VEA	PA	RV	ITD	DL	LA	FA	KAGE	IT	RDV	IAV	VVR	EG	SG	KAI	G	487	
EDD_Wa	STVNEP	Y	SPNG	LAL	LE	GNL	GR	Y	WIKI	SA	V	AE	EN	HV	RA	PA	TV	DSQ	EL	IDA	FKRGE	IT	RDV	IAV	VVR	EG	SG	KAI	G	485	
EDD_Ec	ASFQEP	Y	SHHG	QTK	VL	SGN	LGR	Y	AVMK	SA	V	EN	QI	IEA	PA	TV	DSQ	EL	DA	FA	KAGE	IT	RDV	IAV	VVR	EG	SG	KAI	G	484	
EDD_Zm	RPVSNP	Y	AKD	GGLR	L	LE	GNL	GR	Y	AMK	SA	V	EN	QI	IEA	PA	TV	DSQ	EL	DA	FA	KAGE	IT	RDV	IAV	VVR	EG	SG	KAI	G	486
EDD_Se	ASFQKP	Y	SFHG	QTK	VL	SGN	LGR	Y	AVMK	SA	V	EN	QI	IEA	PA	TV	DSQ	EL	DA	FA	KAGE	IT	RDV	IAV	VVR	EG	SG	KAI	G	484	
EDD_Ab	RSIDAP	Y	QPDG	LGR	LMQ	L	GR	Y	WIKI	SA	V	AE	PH	RV	KV	PA	TV	DSQ	EL	DA	FA	KAGE	IT	RDV	IAV	VVR	EG	SG	KAI	G	487
EDD_Agro	RSIETP	Y	QSDT	GGLM	LIT	GNL	GR	Y	WIKI	SA	V	AE	PH	RV	KV	PA	TV	DSQ	EL	DA	FA	KAGE	IT	RDV	IAV	VVR	EG	SG	KAI	G	487

DHAD_Ec	AARREAQDARGDKWT-PK-NRERQVSFAIRA-YASL-----	ATSADKCAVRDCKSLGG-----	605
DHAD_Se	AARREAQGEARGDKAWT-PK-NRQRQVSFAIRA-YASL-----	ATSADKCAVRDCKSKLGG-----	602
DHAD_Ab	AHRRTIQEA-----KGWH-PKEERKRKYSKALKV-YAMH-----	TTSAAGKCAVRVL-----	595
DHAD_Tm	ERRM-----KEFT-PL-VKEVDSDY-LRR-YAFF-----	VQSAGCAIFRKP-----	554
DHAD_Syne	AHRRAAAMEARGDQAWT-PK-DRDRPISQAIIQA-YAAM-----	TTSAARGVRDLSQLGSR-----	619
DHAD_No	AERHAKMEAKGAKAWK-PF-GRKREVSFALRA-YAAM-----	TTNPAKCAVRDVQSVERD-----	617
DHAD_At	NDRR-----KKWTAPAYKVNRGV---LYK-YIKN-----	VQSASDCCVTDE-----	608
DHAD_Gm	EARR-----KKWTAPYKANRGA---LYK-YIKN-----	VTPASSGCVTDE-----	601
EDD_No	FSRT-----PAIE-DLRPQQFGMGRELEAGRNLL-----	VGVADRCASVFG-----	604
EDD_Sym	AARA-----AAIC-DLTASHRGLGRELFAGRQI-----	ASSAEBCALSIPLGCEPASPRPVTTETVA-----	621
EDD_Wa	AQRE-----VATA-LLSDNQFGMGRELFESGRL-----	ADTAEKCGCTFAFD-----	606
EDD_Ec	AARE-----PHIP-DLSASRVGTGRBLEFSALREK-----	LSGAEQCACITIF-----	603
EDD_Zm	NARP-----HAEKPAFRPGTGRBLEDTIRQN-----	AAKAEDCAVAIYAGAGI-----	607
EDD_Se	AARQ-----PHIP-DLSASRVGTGRBLEFGALREK-----	LSGAEQCACITIF-----	603
EDD_Ab	QSRP-----VAQPEHQAEENVGFRBLEFGVRAA-----	AAPEHCACSVFGALVGVEEPQAH-----	617
EDD_Agro	AKRE-----PARA-DSLNGEVMGRBLFAPSAAAMPAPIRAPAFSSIDILKGEDRCGFRLLTFCASSDCSFLPLTI-----		636

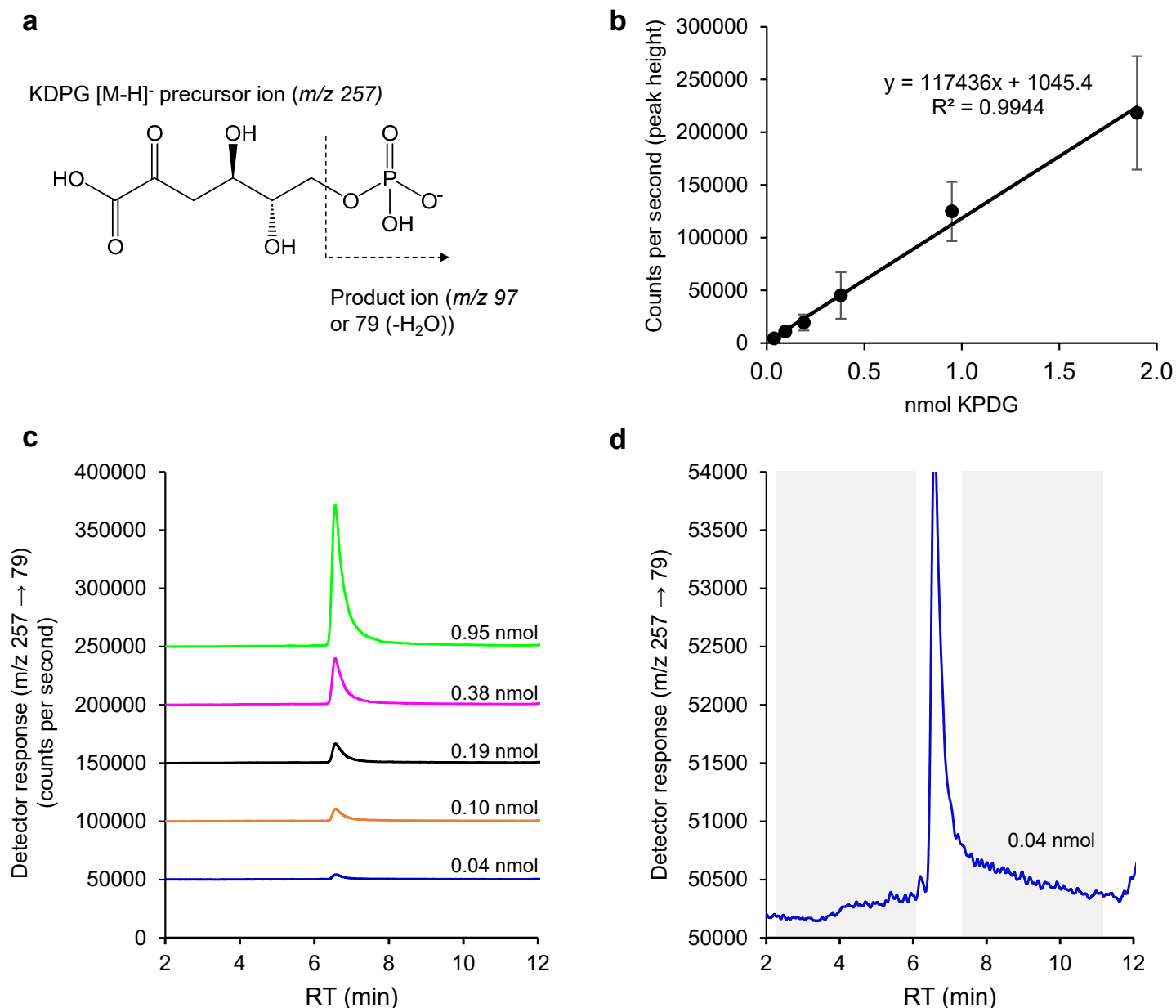
Supplementary figure 1. Amino acid alignment of representative DHAD and EDD sequences from prokaryotes and eukaryotes. Residues in dark and light gray represent 80% and 60% similarity cutoffs, respectively. *Ec*, *Escherichia coli*; *Se*, *Salmonella enterica*; *Ab*; *Acinetobacter baumannii*; *Tm*; *Thermotoga maritima*; *Syne*; *Synechococcus elongatus*; *No*; *Nostoc sp.*; *At*, *Arabidopsis thaliana*; *Gm*, *Glycine max*; *Synm*, *Synechococcus moorigangaii*; *Wa*, *Candidatus Woesearchaeota archaeon*; *Zm*, *Zymomonas mobilis*; *Agro*; *Agrobacterium tumefaciens*



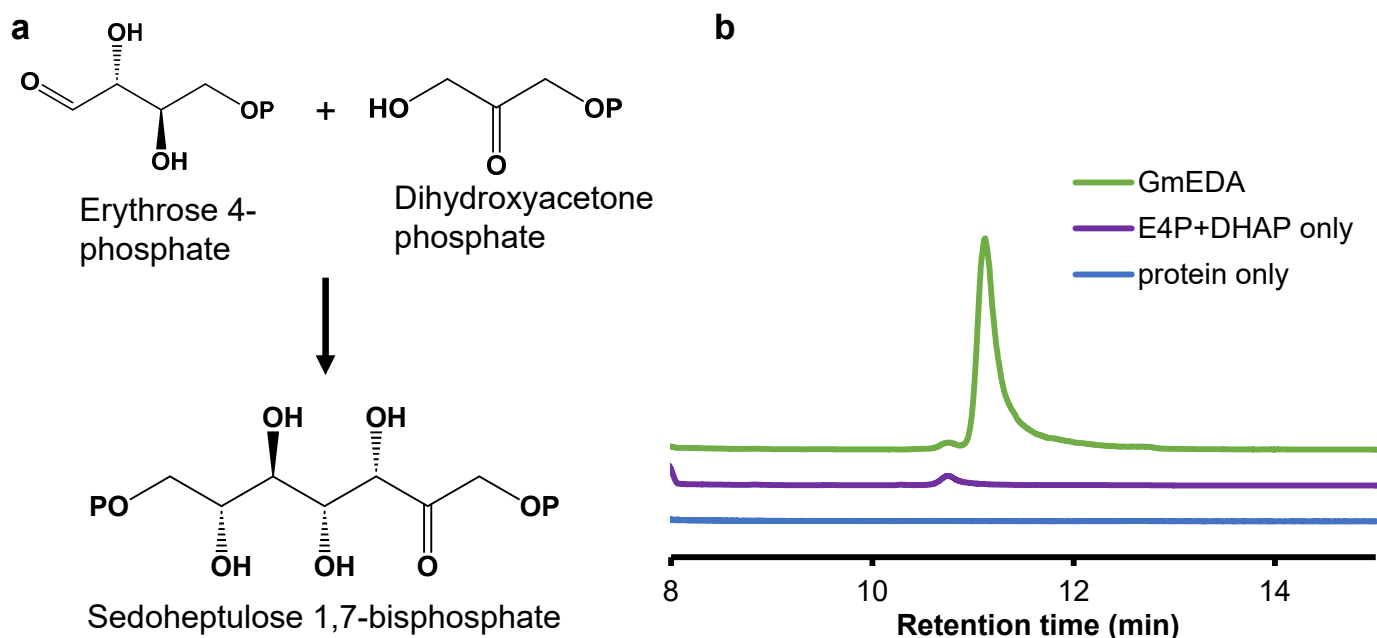
Supplementary figure 2. Species phylogeny of the Cyanobacteriota as previously reported by Moore et al.²⁴. This reconstruction emphasizes species encoding both genes of the Entner-Doudoroff pathway (6-phosphogluconate dehydratase (EDD) and 2-keto-3-deoxy-6-phosphogluconate aldolase (EDA) as shown in supplemental table S2 of this study. Branch lengths in the maximum likelihood phylogenetic tree were based on concatenated ribosomal protein sequences and were downloaded from doi: [10.6084/m9.figshare.7629383](https://doi.org/10.6084/m9.figshare.7629383). While EDA is widespread among cyanobacteria, the EDD gene is present only in a handful of species, none of which appeared in this species phylogeny. Blue text signifies their closest relatives among crown cyanobacteria. The pink branch indicates the closest sister group (*Gloeomargarita lithophora*) to the Archaeplastida (green branches).



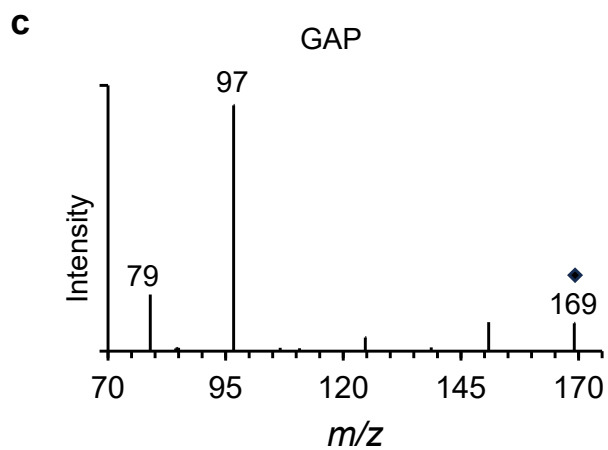
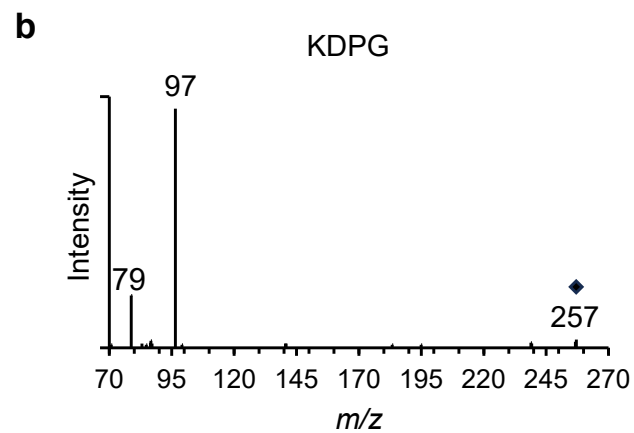
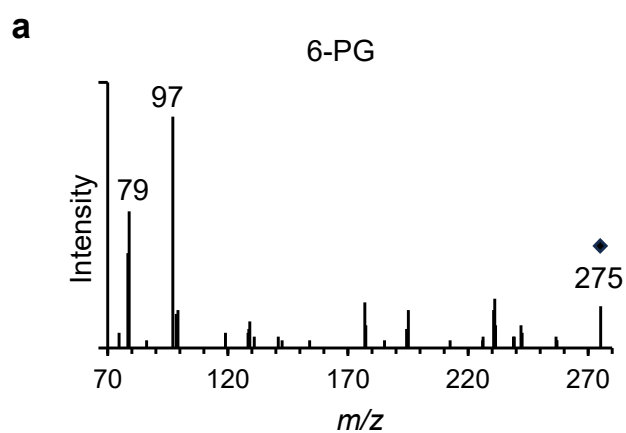
Supplementary figure 3. SDS-PAGE analysis of purified preparations of dihydroxy acid dehydratase (DHAD), 6-phosphogluconate dehydratase (EDD) and 2-keto-3-deoxy-6-phosphogluconate aldolase (EDA). Proteins from plant and bacterial sources were heterologously expressed in *Escherichia coli* BL21(DE3). **a**, Purification of *E. coli* EDD and *Glycine max* DHAD. CP, crude protein; PP, Ni-agarose purified protein. **b**, Purification of *E. coli* EDA and *G. max* EDA. 1, Crude protein extract of an IPTG-induced culture expressing EcEDA. 2, Crude protein extract of an IPTG-induced culture expressing GmEDA. 3, Uninduced control culture; 4, Ni-NTA purified EcEDA. 5, Ni-NTA purified GmEDA.



Supplementary figure 4. Analytical detection of 2-keto-3-deoxygluconate-6-phosphate (KDPG) by liquid chromatography – tandem mass spectrometry (LC-MS/MS). **a**, KDPG precursor (Q_1) and product (Q_3) ions detected by multiple reaction monitoring. **b**, Calibration curve constructed with authentic KDPG used to calculate limits of detection and quantification (LOD and LOQ) ($n = 3$, error bars signify standard deviation). **c**, LC-MS/MS calibration peaks of KDPG at different concentrations used to construct the curve in **b**. Each chromatogram above 0.04 nmol is vertically offset by 50,000 counts per second for improved visibility. Separation was achieved using a Waters BEH amide HILIC chromatography column (2 × 150 mm, 2.5 μ m particle size). See Methods for complete analytical conditions. **d**, 100× zoom of **c** to show the region of the chromatogram used to calculate the background signal for LOD and LOQ calculations (shaded gray region). The lowest calibration level (0.04 nmol KDPG) is shown.



Supplementary figure 5. *In vitro* enzyme assay of purified soybean 2-keto-3-deoxygluconate-6-phosphate aldolase (GmEDA) showing aldolase activity towards substrates of the Calvin-Benson Cycle. **a.** Condensation of erythrose-D-4-phosphate (E4P) and dihydroxyacetone phosphate (DHAP) to produce D-sedoheptulose-1,7-bisphosphate (S1,7BP) as monitored by this assay (see Methods for experimental details). **b.** LC-MS/MS multiple reaction monitoring chromatogram of the GmEDA enzymatic reaction product (S1,7BP) based on its characteristic mass transition (m/z 369 $\rightarrow$ 79).



Supplementary figure 6. MS/MS product ion spectra of important metabolites used in this study. **a**, 6-phosphogluconate (6-PG); **b**, 2-keto-3-deoxy-6-phosphogluconate (KDPG); **c**, glyceraldehyde-3-phosphate (GAP). Spectra were obtained by direct infusion of authentic standards. Electrospray ionization was conducted in negative mode using nitrogen as collision gas (see Methods for additional details). The m/z 97 and 79 product ions represent a phosphate ion and its dehydrated phosphite form, respectively. Diamonds indicate the $[M-H]^-$ precursor ion. Each scan was repeated at least three times with identical results.

Supplementary References

1 May 2023 version of NCBI (<https://www.ncbi.nlm.nih.gov/>).

- 2 Pappas, Katherine M., et al. "Genome sequence of the ethanol-producing *Zymomonas mobilis* subsp. mobilis lectotype strain ATCC 10988." (2011): 5051-5052.
- 3 Liu, Yi-Fan, et al. "Anaerobic degradation of paraffins by thermophilic Actinobacteria under methanogenic conditions." *Environmental Science & Technology* (2020) 54.17:10610-10620.
- 4 Tully, Benjamin J., Elaina D. Graham, and John F. Heidelberg. "The reconstruction of 2,631 draft metagenome-assembled genomes from the global oceans." *Scientific data* (2018) 5:1-8.
- 5 Zhou, Liu, et al. "Genome-and community-level interaction insights into carbon utilization and element cycling functions of Hydrothermarchaeota in hydrothermal sediment." *Msystems* (2020) 5:e00795-19.
- 6 Hou, Jialin, et al. "Microbial succession during the transition from active to inactive stages of deep-sea hydrothermal vent sulfide chimneys." *Microbiome* (2020) 8:1-18.
- 7 Munson-McGee, Jacob H., et al. "Nanoarchaeota, their Sulfolobales host, and Nanoarchaeota virus distribution across Yellowstone National Park hot springs." *Applied and Environmental Microbiology* (2015) 81:7860-7868.
- 8 Singh, Tarkeshwar, and Punyasloke Bhadury. "Description of a new marine planktonic cyanobacterial species *Synechococcus moorigangaii* (Order Chroococcales) from Sundarbans mangrove ecosystem." *Phytotaxa* (2019) 393:263-277.
- 9 Brenes-Guillén, L., et al. "Draft Genome Sequences of Two Cyanobacteria *Leptolyngbya* spp. Isolated from Microbial Mats in Miravalles Thermal Spring, Costa Rica." *Microbiology Resource Announcements* (2021) 10.41:10-1128.
- 10 Ponce-Toledo, Rafael I., et al. "An early-branching freshwater cyanobacterium at the origin of plastids." *Current Biology* (2017) 27:386-391.
- 11 Hirooka, Shunsuke, et al. "Acidophilic green algal genome provides insights into adaptation to an acidic environment." *Proceedings of the National Academy of Sciences* (2017) 114:E8304-E8313.
- 12 Arriola, Matthew B., et al. "Genome sequences of *Chlorella sorokiniana* UTEX 1602 and *Micractinium conductrix* SAG 241.80: implications to maltose excretion by a green alga." *The Plant Journal* (2018) 93:566-586.
- 13 Fuentes-Ramírez, Emma O., et al. "The plastid proteome of the nonphotosynthetic chlorophyte alga *Polytomella parva*." *Microbiological Research* (2021) 243:126649.
- 14 Merchant, Sabeeha S., et al. "The *Chlamydomonas* genome reveals the evolution of key animal and plant functions." *Science* (2007) 318:245-250.
- 15 Yamamoto, Kayoko, et al. "Three genomes in the algal genus *Volvox* reveal the fate of a haploid sex-determining region after a transition to homothallism." *Proceedings of the National Academy of Sciences* (2021) 118:e2100712118.
- 16 Hess, Sebastian, et al. "A phylogenomically informed five-order system for the closest relatives of land plants." *Current Biology* (2022) 32:4473-4482.
- 17 De Vries, Jan, et al. "Heat stress response in the closest algal relatives of land plants reveals conserved stress signaling circuits." *The Plant Journal* (2020) 103:1025-1048.
- 18 De Vries, Jan, et al. "Embryophyte stress signaling evolved in the algal progenitors of land plants." *Proceedings of the National Academy of Sciences* (2018) 115:E3471-E3480.
- 19 Ju, Chuanli, et al. "Conservation of ethylene as a plant hormone over 450 million years of evolution." *Nature plants* (2015) 1:1-7.
- 20 Van de Poel, Bram, et al. "Transcriptome profiling of the green alga *Spirogyra pratensis* (Charophyta) suggests an ancestral role for ethylene in cell wall metabolism, photosynthesis, and abiotic stress responses." *Plant Physiology* (2016) 172:533-545.
- 21 Nishiyama, Tomoaki, et al. "The Chara genome: secondary complexity and implications for plant terrestrialization." *Cell* (2018) 174:448-464.
- 22 Wood, V., et al. "The genome sequence of *Schizosaccharomyces pombe*." *Nature* (2002) 415:871-880.
- 23 Kim, Myung-Shin, et al. "Genome assembly of the popular Korean soybean cultivar Hwangkeum." *G3* (2021) 11:jkab272.
- 24 Moore, Magnabosco, et al. "An expanded ribosomal phylogeny of Cyanobacteria supports a deep placement of plastids" *Frontiers in Microbiology* (2019) 10:1612
- 25 Ochoa de Alda, Esteban, et al. "The plastid ancestor originated among one of the major cyanobacterial lineages" *Nature Communications* (2014) 5:4937