

Table 5. Protein-coding genes in sequenced chloroplast genomes and identified nuclear homologues thereof

	Z	O	N	O	P	M	E	C	N	M	O	G	P	C	C	P	E	S	Nuclear homologue identified	Comments, previous names, probable or known function
	e	r	i	e	i	a	u	h	e	e	d	u	o	a	y	l	p	y		
	a	y	c	n	n	r	g	l	p	s	o	i	r	n	a	a	i	n		
<i>accA</i>	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	-	-	+	AF056969	Acetyl-CoA carboxylase α SU
<i>accB</i>	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	-	-	+	sp Q42533	Acetyl-CoA carboxylase biotin carrier SU
<i>accD</i>	-	-	+	+	+	+	-	+	+	-	-	-	+	+	-	-	+	+		<i>ycf11</i> ; Acetyl-CoA carboxylase carboxytransferase β SU (<i>zfpA</i>)
<i>acpP</i>	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	-	-	+	sp P11829	Acyl carrier protein (<i>acpA</i> , <i>acpC</i>)
<i>apcA</i>	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	-	-	+		Allophycocyanin α SU
<i>apcB</i>	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	-	-	+		Allophycocyanin β SU
<i>apcD</i>	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	-	-	+		Allophycocyanin γ SU
<i>apcE</i>	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	-	-	+		Phycobilisome core linker prot.
<i>apcF</i>	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	-	-	+		Allophycocyanin 18 kDa β SU
<i>argB</i>	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	-	-	+	pir T46192	<i>N</i> -Acetylglutamate kinase
<i>atpA</i>	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+		ATP synthase CF1 α SU
<i>atpB</i>	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+		ATP synthase CF1 β SU
<i>atpD</i>	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	-	-	+	pir T08083	ATP synthase CF1 δ SU
<i>atpE</i>	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+		ATP synthase CF1 ϵ SU
<i>atpF</i>	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+		ATP synthase CF0 SU I
<i>atpG</i>	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	-	-	+	pir T05402	ATP synthase CF0 SU II
<i>atpH</i>	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+		ATP synthase CF0 SU III
<i>atpI</i>	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-	+		ATP synthase CF0 SU IV
<i>bas1</i>	-	-	-	-	-	-	-	-	-	-	+	-	+	-	-	-	-	+	sp Q96291	<i>ycf42</i> ; Peroxiredoxin (thiol-specific antioxidant prot)
<i>bioY</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	+		<i>Cpa</i> ORF188, <i>Syn</i> slr1365 biotin synthase
<i>carA</i>	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	-	-	+	AY046004	Carbamoyl phosphate synthetase small SU
<i>cbxX</i>	-	-	-	-	-	-	-	-	-	-	+	+	+	+	-	-	-	(+)	AF360342	Homologues in red type Calvin cycle operons of proteobacteria (<i>cfxQ</i> , <i>cbxX</i>) ^a
<i>ccsA</i>	+	+	+	+	+	+	-	+	+	+	+	+	+	+	+	-	-	+		<i>ycf5</i> ; Heme attachment to plastid cyt <i>c</i>
<i>cemA</i>	+	+	+	+	+	+	-	+	+	+	-	+	+	+	+	-	-	+		<i>ycf10</i> , <i>hbp</i> ; Envelope membrane prot.
<i>chlB</i>	-	-	-	-	+	+	-	+	+	+	-	-	+	-	+	-	-	+		Protochlorophyllide reductase ChlB SU
<i>chlI</i>	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	-	-	+	sp P93162	Protochlorophyllide reductase magnesium chelatase SU
<i>chlL</i>	-	-	-	-	+	+	-	+	+	+	-	-	+	-	+	-	-	+		Protochlorophyllide reductase ATP-bind. SU (<i>frxC</i>)
<i>chlN</i>	-	-	-	-	+	+	-	+	+	+	-	-	+	-	+	-	-	+	AF124161	Protochlorophyllide reductase ChlN SU, Note: the gene annotated as chlN in <i>Can</i> is moeB
<i>clpC</i>	-	-	-	-	-	-	-	-	-	-	+	+	+	+	-	+	-	+	sp P35100	Clp protease ATP-bind. SU
<i>clpP</i>	+	+	+	+	+	+	-	+	+	+	-	-	-	-	+	-	+	+	AAK39833	Clp protease proteolytic SU, nucleomorph-encoded in <i>Guillardia</i> ^b
<i>cobA</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	+	AB011477	<i>Syn</i> sll0378 uroporphyrin-III C-methyltransferase
<i>cpcA</i>	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	-	-	+		Phycocyanin α SU
<i>cpcB</i>	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	-	-	+		Phycocyanin β SU
<i>cpcG</i>	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	-	-	+		Phycobilisome rod-core linker polypeptide
<i>cpeA</i>	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	+		Phycoerythrin α SU
<i>cpeB</i>	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	-	-	+		Phycoerythrin β SU
<i>crtE</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	+	AC007260	Geranylgeranyl pyrophosphate synthetase hom.
<i>cvuI</i>	-	-	-	-	-	-	+	+	-	-	-	-	-	-	-	-	-	-		Intron encoded DNA endonuclease, sim. <i>Chlamydomonas I-CreI</i> (P05725) (<i>I-cvuI</i>); mitoch. hom. in <i>Acanthamoeba</i> (S46447)
<i>cysA</i>	-	-	-	-	-	+	-	+	+	+	-	-	-	-	-	-	-	+		Probable transport prot. (<i>mbpX</i>)
<i>cysT</i>	-	-	-	-	-	+	-	+	+	+	-	-	-	-	-	-	-	+		Probable transport prot. (<i>mbpY</i>)
<i>desA</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	+	pir T07742	Fatty acid ω -6 desaturase; Δ -12 desaturase, <i>Syn</i> sll1468
<i>dfr</i>	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	-	-	+		<i>ycf26</i> ; <i>Syn</i> sll0698 two-component sensor kinase (drug sensory prot. A)
<i>dnaB</i>	-	-	-	-	-	-	-	-	-	-	+	+	+	+	-	-	-	+		DNA-replication helicase SU

<i>dnaK</i>	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	-	-	+	pir T08151	Hsp70-type chaperone
<i>dsbD</i>	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	+	AF225913	<i>Por</i> ORF240, <i>Syn</i> sll0621; thiol:disulfide interchange prot. DsbD
<i>fabH</i>	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	+	sp Q07510	β -ketoacyl-acyl carrier prot. synthase III
<i>fdx</i>	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	-	-	+		<i>Por</i> ORF75a, <i>Syn</i> sss3184, 2[4Fe-4S] ferredoxin
<i>frtC</i>	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	-	-	+	AC006955	<i>frtB</i> , Ferredoxin-thioredoxin reductase β SU
<i>ftsI</i>	-	-	-	-	-	-	-	-	+	+	-	-	-	-	-	-	-	+		penicillin-binding protein
<i>ftsH</i>	-	-	-	-	-	-	-	-	-	-	+	+	+	+	-	-	-	+	AC004669	<i>ycf25</i> , Cell division prot., ATPase, protease
<i>ftsW</i>	-	-	-	-	-	-	-	-	+	+	-	-	-	-	+	-	-	+		Putative cell/organelle division prot.
<i>glnB</i>	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	-	-	+	AF095455	Nitrogen regulatory prot. P-II
<i>gltB</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	-	+	AL396716	Glutamate synthase (GOGAT)
<i>glmS</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	+		<i>Syn</i> sll0220; L-glutamine:D-fructose-6-phosphate amidotransferase; <i>Syn</i> sp P72720
<i>groEL</i>	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	-	-	+	sp P08926	Chaperonin, 60 kDa
<i>groES</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+	Q02073	Chaperonin, 10 kDa
<i>hemA</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+	P42804	Glutamyl-tRNA reductase (GLUTR)
<i>hisH</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	sp Q9SZ30	Histidinol-phosphate aminotransferase; <i>Syn</i> slr0084
<i>hlpA</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+		DNA-binding protein HU Homolog
<i>ilvB</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	sp P09114	Acetohydroxyacid synthase large SU
<i>ilvH</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	AB005242	Acetohydroxyacid synthase small SU
<i>infA</i>	+	+	+	-	+	+	-	+	+	+	-	-	-	-	-	-	-	+		Translational initiation factor 1
<i>infB</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	AF367343	Translational initiation factor 2
<i>infC</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	sp P36177	Translational initiation factor 3
<i>lipB</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	AC007519	<i>Syn</i> slr0994; lipoate-protein ligase B
<i>lpxA</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	AL161575	<i>Syn</i> sll0379; acyl-[acyl-carrier-protein]:UDP-N-acetylglucosamine o-acyltransferase
<i>lpxC</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	AC004133	<i>Syn</i> sll1508 37, <i>Syn</i> sll1508; UDP-3-O-acyl N-acetylglucosamine deacetylase
<i>matK</i>	+	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	+		<i>ycf14</i> Probable intron maturase
<i>menA</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+		<i>Syn</i> slr1518; 1,4-dihydroxy-2-naphthoate octaprenyltransferase
<i>menB</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	AC002292	<i>Syn</i> sll1127; naphthoate synthase
<i>menC</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	(+)	AC011665	<i>O</i> -succinylbenzoate-CoA synthase, <i>Ath</i> nuclear homologue menC-menD fusion
<i>menD</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	AC011665	<i>Syn</i> sll0603; 2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate synthase, <i>Ath</i> nuclear homologue menC-menD fusion
<i>menE</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+		<i>Syn</i> slr0492; <i>O</i> -succinylbenzoic acid-CoA ligase
<i>menF</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	AF367342	<i>Syn</i> slr0817; menaquinone-specific isochorismate synthase
<i>minD</i>	-	-	-	-	-	-	-	+	+	+	-	+	-	-	-	-	-	+	AB030278	Septum-site determining prot.
<i>minE</i>	-	-	-	-	-	-	-	+	-	-	-	+	-	-	-	-	-	+		<i>Syn</i> ssl0546; <i>Chl minE</i> has very weak sim. to <i>Syn</i> ssl0546
<i>mntA</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+		<i>Cpa</i> ORF244, manganese transport system ATP-bind. prot.
<i>mntB</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+		<i>Cpa</i> ORF299, manganese transport system membrane prot.
<i>moeB</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	AF124161	<i>Por</i> ORF382, molybdopterin biosynthesis prot., <i>Syn</i> sll1536; Note: annotated as chlN in published <i>Can</i> sequence
<i>nadA</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+		Quinolate synthase
<i>nblA</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+		<i>ycf18</i> ; Phycobilisome degradation prot. <i>Syn</i> sp P73890
<i>ndhA</i>	+	+	+	+	-	+	-	-	+	+	-	-	-	-	-	-	-	+		NADH-plastoquinone oxidoreductase SU 1
<i>ndhB</i>	+	+	+	+	-	+	-	-	+	+	-	-	-	-	-	-	-	+		NADH-plastoquinone oxidoreductase SU 2
<i>ndhC</i>	+	+	+	+	-	+	-	-	+	+	-	-	-	-	-	-	-	+		NADH-plastoquinone oxidoreductase SU 3
<i>ndhD</i>	+	+	+	+	-	+	-	-	+	+	-	-	-	-	-	-	-	+		NADH-plastoquinone oxidoreductase SU 4
<i>ndhE</i>	+	+	+	+	-	+	-	-	+	+	-	-	-	-	-	-	-	+		NADH-plastoquinone oxidoreductase SU 4L
<i>ndhF</i>	+	+	+	+	-	+	-	-	+	+	-	-	-	-	-	-	-	+		NADH-plastoquinone oxidoreductase SU 5
<i>ndhG</i>	+	+	+	+	-	+	-	-	+	+	-	-	-	-	-	-	-	+		NADH-plastoquinone oxidoreductase SU 6
<i>ndhH</i>	+	+	+	+	-	+	-	-	+	+	-	-	-	-	-	-	-	+		NADH-plastoquinone oxidoreductase 49 kDa SU
<i>ndhI</i>	+	+	+	+	-	+	-	-	+	+	-	-	-	-	-	-	-	+		NADH-plastoquinone oxidoreductase SU I
<i>ndhJ</i>	+	+	+	+	-	+	-	-	+	+	-	-	-	-	-	-	-	+		NADH-plastoquinone oxidoreductase SU J
<i>ndhK</i>	+	+	+	+	-	+	-	-	+	+	-	-	-	-	-	-	-	+		NADH-ubiquinone oxidoreductase SU K (<i>psbG</i>)
<i>ntcA</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+		<i>ycf28</i> ; Global nitrogen transcriptional regulator
<i>odpA</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	AC007323	Pyruvate dehydrogenase E1 component, α SU
<i>odpB</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	AF167983	Pyruvate dehydrogenase E1 component, β SU
<i>pbsA</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+		Heme oxygenase, gene more similar to nuclear homologues of animals than of plants
<i>petA</i>	+	+	+	+	+	+	-	+	+	+	+	+	+	+	+	-	-	+		Apocytochrome <i>f</i>

<i>rpl11</i>	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	-	-	+	AB042934	Ribosomal protein L11
<i>rpl12</i>	-	-	-	-	-	-	+	+	+	-	+	+	+	+	+	-	-	+	sp P02398	Ribosomal protein L12 ^g
<i>rpl13</i>	-	-	-	-	-	-	-	-	-	-	+	+	+	+	-	-	-	+	sp P12629	Ribosomal protein L13
<i>rpl14</i>	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+		Ribosomal protein L14
<i>rpl16</i>	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+		Ribosomal protein L16
<i>rpl18</i>	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	-	-	+	AF336922	Ribosomal protein L18
<i>rpl19</i>	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	-	-	+	AF250384	Ribosomal protein L19
<i>rpl20</i>	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+		Ribosomal protein L20
<i>rpl21</i>	-	-	-	-	-	+	-	-	-	-	+	+	+	+	+	-	-	+	sp P51412	Ribosomal protein L21
<i>rpl22</i>	+	+	+	+	+	+	+	-	-	+	+	+	+	+	+	-	-	+		Ribosomal protein L22
<i>rpl23</i>	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	-	+		Ribosomal protein L23
<i>rpl24</i>	-	-	-	-	-	-	-	-	-	-	+	+	+	+	-	-	-	+	sp P11893	Ribosomal protein L24
<i>rpl27</i>	-	-	-	-	-	-	-	-	-	-	+	+	+	+	-	-	-	+	sp P30155	Ribosomal protein L27
<i>rpl28</i>	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+	-	-	+	P30596	Ribosomal protein L28
<i>rpl29</i>	-	-	-	-	-	-	-	-	-	-	+	+	+	+	-	-	-	+	AF147725	Ribosomal protein L29
<i>rpl31</i>	-	-	-	-	-	-	-	-	-	-	+	+	+	+	-	-	-	+	AC023754	Ribosomal protein L31
<i>rpl32</i>	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-	+		Ribosomal protein L32
<i>rpl33</i>	+	+	+	+	+	+	-	-	-	+	+	+	+	+	+	-	+	+		Ribosomal protein L33
<i>rpl34</i>	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	-	-	+		Ribosomal protein L34
<i>rpl35</i>	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	-	-	+	sp P23326	Ribosomal protein L35
<i>rpl36</i>	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+		Ribosomal protein L36
<i>rpoA</i>	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	+		RNA polymerase α SU, (formerly <i>ycf67</i> in <i>Eug</i> , see ref. 3)
<i>rpoB</i>	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+		RNA polymerase β SU
<i>rpoC1</i>	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+		RNA polymerase β' SU
<i>rpoC2</i>	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+		RNA polymerase β'' SU ^h
<i>rps1</i>	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	+	P29344	Ribosomal protein S1
<i>rps2</i>	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+		Ribosomal protein S2
<i>rps3</i>	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+		Ribosomal protein S3
<i>rps4</i>	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+		Ribosomal protein S4
<i>rps5</i>	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	-	+	X93156	Ribosomal protein S5
<i>rps6</i>	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	-	-	+	AF250383	Ribosomal protein S6
<i>rps7</i>	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+		Ribosomal protein S7
<i>rps8</i>	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+		Ribosomal protein S8
<i>rps9</i>	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	-	-	+	AB022676	Ribosomal protein S9
<i>rps10</i>	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	-	-	+	AP000375	Ribosomal protein S10
<i>rps11</i>	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+		Ribosomal protein S11
<i>rps12</i>	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+		Ribosomal protein S12
<i>rps13</i>	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	-	-	+	sp P42732	Ribosomal protein S13
<i>rps14</i>	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+		Ribosomal protein S14
<i>rps15</i>	+	+	+	+	+	+	-	-	-	+	-	-	-	-	-	-	-	+		Ribosomal protein S15
<i>rps16</i>	+	+	+	+	-	-	-	-	-	+	+	+	+	+	+	-	-	+		Ribosomal protein S16
<i>rps17</i>	-	-	-	-	-	-	-	-	-	-	+	+	+	-	+	+	-	+	sp P17092	Ribosomal protein S17
<i>rps18</i>	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+		Ribosomal protein S18
<i>rps19</i>	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+		Ribosomal protein S19
<i>rps20</i>	-	-	-	-	-	-	-	-	-	-	+	+	+	-	+	-	-	+	AP001299	Ribosomal protein S20
<i>secA</i>	-	-	-	-	-	-	-	-	-	-	+	+	+	+	-	-	-	+	pir S72453	Preprotein-translocase SU A
<i>secY</i>	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	-	-	+	sp Q38885	Preprotein-translocase SU Y
<i>syfB</i>	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	+		Phenylalanine tRNA synthetase
<i>syh</i>	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	+		Histidine tRNA synthetase
<i>thdF</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+	AC009243	<i>Syn</i> sll1615 ThdF thiophene and furan oxidation protein
<i>thiG</i>	-	-	-	-	-	-	-	-	-	-	+	-	+	+	-	-	-	+		ThiG protein, thiamine biosynthesis
<i>trpA</i>	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	-	-	+	AL139656	Tryptophane synthase α SU
<i>trpG</i>	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	-	-	+	AC004133	Anthranilate synthase component II, glutamine amidotransferase
<i>trxA</i>	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	-	-	+	sp 48384	Thioredoxin

<i>ycf70</i>	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	footnote i, <i>Ory</i> ORF91 JQ0209, <i>Zea</i> ORF69 S58539
<i>ycf71</i>	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	footnote i, <i>Ory</i> ORF82 JQ0268, <i>Zea</i> ORF75 S58641
<i>ycf72</i>	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	footnote i, <i>Ory</i> ORF137 JQ0269, <i>Zea</i> ORF137 S58640
<i>ycf73</i>	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	footnote i, <i>Ory</i> ORF249 JQ0274, <i>Zea</i> ORF173 S58632
<i>ycf74</i>	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	footnote i, ORF in <i>trnA</i> intron <i>Ory</i> ORF109 JQ0280, <i>Zea</i> ORF49 S58608
<i>ycf75</i>	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	footnote i, <i>Ory</i> ORF63a JQ0283, <i>Zea</i> ORF63 S58623
<i>ycf76</i>	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	footnote i, <i>Ory</i> ORF85b JQ0278, <i>Zea</i> ORF85 S58627
<i>ycf77</i>	-	-	-	-	-	-	+	+	+	-	-	-	-	-	-	-	-	-	-	AY064138 footnote i, <i>Chl</i> ORF1720, in positions 950-1350 residual sim. to FtsH ^j
<i>ycf78</i>	-	-	-	-	-	-	+	+	+	-	-	-	-	-	-	-	-	-	-	footnote i, <i>Chl</i> ORF819, hom. in <i>Chlamydomonas reinhardtii</i> cpDNA X92726 ^k
<i>ycf80</i>	-	-	-	-	-	-	-	-	-	-	+	+	+	-	-	-	-	-	-	NC_003075 footnote i, <i>Por</i> ORF450, <i>Can</i> ORF253; Tic22-like protein; nucleomorph-encoded hom. in <i>Guillardia</i>
<i>ycf81</i>	-	-	-	-	-	-	-	+	+	-	-	-	-	-	-	-	-	-	+	<i>Nep</i> ORF138, <i>Syn</i> slr1972
<i>ycf82</i>	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	+	<i>Syn</i> slr1508 (sim. lipopolysaccharide-synthesis and glycosyltransferases)
<i>ycf83</i>	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	+	<i>Syn</i> slr0204
<i>ycf84</i>	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	+	<i>Syn</i> slr0882 (hyp. integral membrane protein)

Plus and minus signs indicate presence and absence, respectively of the gene in the given chloroplast genome. A plus sign in parentheses (+) indicates that the gene is not found in BLAST searches in the *Synechocystis* genome, but was found in other cyanobacterial or related genomes and was thus scored as being present in the ancestral plastid genome. Synonyms are indicated in parentheses, except in the case of ORF designations. Accession numbers refer to Protein Information Resource (PIR), GenBank, or SwissProt data bases. Reading frames designated as pseudogenes in *Epifagus* cpDNA (5) and black pine cpDNA (6) were scored as absent here. The last revision of *ycf* nomenclature ended with *ycf79* (7). *Epifagus* and *Plasmodium* were excluded from the phylogeny and hence from the gene loss analysis.

Abbreviations and Genome Accession Numbers: *Zea*, *Zea mays* (X86563); *Ory*, *Oryza sativa* (X15901); *Nic*, *Nicotiana tabacum* (S54304); *Spi*, *Spinacia oleracea* (AJ400848); *Oen*, *Oenothera elata* (NC_002693); *Pin*, *Pinus thunbergii* (D17510); *Mar*, *Marchantia polymorpha* (X04465); *Eug*, *Euglena gracilis* (NC_001603); *Chl*, *Chlorella vulgaris* (NC_001865); *Nep*, *Nephroselmis olivacea* (NC_000927); *Mes*, *Mesostigma viride* (NC_002186); *Odo*, *Odontella sinensis* (Z67753); *Gui*, *Guillardia theta* (NC_000926); *Por*, *Porphyra purpurea* (U38804); *Can*, *Cyanidium caldarium* (AF022186, Z36235, Z70297); *Pla*, *Plasmodium falciparum* (X95275, X95276); *Cpa*, *Cyanophora paradoxa* (NC_001675); *Epi*, *Epifagus virginiana* (M81884); *Syn*, *Synechocystis* PCC 6803 (NC_000911); sim., similar to; hom., homologue; SU, subunit; bind., binding; prot., protein; RT, reverse transcriptase.

^a*cbbX* is located in proteobacterial *cbb* operons (*rbcLr-rbcSr-cbbX*), not to be confused with *rbcX* in cyanobacterial *cbb* operons (*rbcLg-rbcX-rbcSg*; ref. 8). Orthology of *Arabidopsis* nuclear AF360342 uncertain.

^b*Cyanophora* cpDNA encodes two distinct *clpP* genes that are 36% identical, whereby *Synechocystis* also possesses two distinct *clpP* genes that are however 74% identical.

^c*petL*, a 31 amino acid long component of the cytochrome *b6/f* complex, is too short to detect its *Synechocystis* homologue (gi1653694) among the highest BLAST scores, but it was counted as being of cyanobacterial origin.

^d*pgmA* has related but paralogous nuclear homologues for cytosolic phosphoglycerate mutase in many higher plants (e.g. *Nic* P35494).

^cThere is still a problem with *psbW* nomenclature. The gene designated in previous studies of chloroplast genomes (including ref. 9) as *psbW* shares no similarity with the nuclear-encoded gene for the psbW protein functionally characterized from spinach chloroplasts (10). The spinach gene designation has priority, but the gene data bases use the psbW designation. The chloroplast-encoded protein designated as PsbW is similar to *Synechocystis* sll1398 (GenBank 1651690), annotated as a 13 kDa PSII protein. PsbW was previously described as *ycf79*.

^fClosely related to the Calvin cycle transcriptional regulator *cbbR* (11).

^gChloroplast *rpl12* has in some cases also been designated as *rpl7*.

^h*rpoC2* is sometimes also designated as *rpoD* in cpDNA, but the gene designated *rpoD* in *Syn* is a sigma factor, the designation *rpoC2* for the chloroplast gene is less ambiguous and hence preferable.

ⁱNo similarity to prokaryotic genes was detected in BLAST searches.

^jDesignated as *ycf1* in the *Chlorella* genome (12), designated as *ycf2* in *Mesostigma* and *Nephroselmis*, but we found no significant similarity between this *Chlorella* gene and cpDNA genes previously designated as *ycf1* and *ycf2* (13). Thus, this gene is designated here as *ycf77*, even though the tree homologues have similarity only over short stretches. They are related to FtsH, but are not genuine orthologs of eubacterial FtsH.

^kWe found no significant similarity between cpDNA genes previously designated as *ycf2* (13) and the gene tentatively designated as *ycf2* in the *Chlorella* genome (12), that is hence designated as *ycf78* here.

^l*ycf46* has a related but paralogous nuclear homologue in *Ory* D17789 (14).

^m*ycf64* has similarity to glutaredoxin-like proteins, nuclear homologues are found in yeast and *Plasmodium*, a distant homologue is nuclear in *Arabidopsis* (e327479 in Z97339).

1. Hager, M., Biehler, K., Illerhaus, J., Ruf, S. & Bock, R. (1999) *EMBO J.* **18**, 5834-5842
2. Grimm, B., Kruse, E. & Kloppstech, K. (1989) *Plant Mol. Biol.* **13**, 583-593.
3. Sheveleva, E. V., Giordani, N. V. & Hallick, R. B. (2002) *Nucleic Acids Res.* **5**, 1247-1254
4. Bauer, C. C., Ramaswamy, K. S., Endley, S., Scappino, L. A., Golden, J. W. & Haselkorn, R. (1997) *FEMS Microbiol. Lett.* **151**, 23-30.
5. Wolfe, K. H., Morden, C. W. & Palmer, J. D. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 10648-10652.
6. Wakasugi, T., Tsudzuki, J., Ito, S., Nakashima, K., Tsudzuki, T. & Sugiura, M. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 9794-9798.
7. Stoebe, B., Hansmann, S., Goremykin, V., Kowallik, K. V. & Martin, W. (1999) in *Advances in Plant Molecular Systematics*, eds. Hollingsworth, C., Batemann, R. & Gornall, M. (Taylor & Francis, London), pp. 327-352.
8. Li, L.-A. & Tabita, F. R. (1997) *J. Bacteriol.* **197**, 3793-3796.
9. Martin, W., Stoebe, B., Goremykin, V., Hansmann, S., Hasegawa, M. & Kowallik, K. V. (1998) *Nature (London)* **393**, 162-165.
10. Lorkovic, Z. J., Schroder, W. P., Pakrasi, H. B., Irrgang, K. D., Herrmann, R. G. & Oelmuller, R. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 8930-8934.
11. Qian, Y. & Tabita, F. R. (1996) *J. Bacteriol.* **178**, 12-18.
12. Wakasugi, T., Nagai, T., Kapoor, M., Sugita, M., Ito, M., Ito, S., Tsudzuki, J., Nakashima, K., Tsudzuki, T., Suzuki, Y., *et al.* (1997) *Proc. Natl. Acad. Sci. USA* **94**, 5967-5972.
13. Reardon, E. M. & Price, C. A. (1995) *Plant Mol. Biol. Rep.* **13**, 320-326.
14. Suzuka, I., Koga-Ban, Y., Sasaki, T., Minobe, Y. & Hashimoto, J. (1994) *Plant Sci.* **103**, 33-40.