PLANT SCIENCES

A unique flavoenzyme operates in ubiquinone biosynthesis in photosynthesis-related eukaryotes

Jing-Jing Xu[1]†, Xiao-Fan Zhang[2]†, Yan Jiang[1,3], Hang Fan[1], Jian-Xu Li[4], Chen-Yi Li[4], Qing Zhao[1], Lei Yang[1], Yong-Hong Hu[1], Cathie Martin[5], Xiao-Ya Chen[1,4]*

Coenzyme Q (CoQ) is an electron transporter in the mitochondrial respiratory chain, yet the biosynthetic pathway in eukaryotes remains only partially resolved. *C*6-hydroxylation completes the benzoquinone ring full substitution, a hallmark of CoQ. Here, we show that plants use a unique flavin-dependent monooxygenase (CoqF), instead of di-iron enzyme (Coq7) operating in animals and fungi, as a *C*6-hydroxylase. CoqF evolved early in eukaryotes and became widely distributed in photosynthetic and related organisms ranging from plants, algae, apicomplexans, and euglenids. Independent alternative gene losses in different groups and lateral gene transfer have ramified CoqF across the eukaryotic tree with predominance in green lineages. The exclusive presence of CoqF in Streptophyta hints at an association of the flavoenzyme with photoautotrophy in terrestrial environments. CoqF provides a phylogenetic marker distinguishing eukaryotes and represents a previously unknown target for drug design against parasitic protists.

INTRODUCTION

Ubiquinone, also known as coenzyme Q (CoQ), is an essential component of oxidative respiration present in eukaryotes at all trophic levels and also in some proteobacteria. Structurally, CoQ is composed of a benzoquinone head group bearing a polyisoprenoid tail of varied number of isoprenoid units between species, 6 (CoQ_6) in *Saccharomyces cerevisiae*, 8 (CoQ_8) in *Escherichia coli*, 9 (CoQ_9) in *Arabidopsis thaliana* and rice, and 10 (CoQ_{10}) in tobacco and humans. In eukaryotes, CoQ functions primarily as a mobile electron carrier in the mitochondrial aerobic respiratory chain that generates adenosine triphosphate (ATP), the cell's energy source. In addition, CoQ acts as a cofactor in multiple pathways, and in reduced form, it is a strong antioxidant in nearly all cellular membranes (*1*). In humans, CoQ deficiency impairs energy production and causes severe health problems (*2*). Although plants are autotrophic, CoQ is involved in many aspects of growth and development and indispensable for embryo development (*3*, *4*). Despite its paramount importance, the CoQ biosynthetic pathway in eukaryotes has not been elucidated fully.

Biochemical characterizations of CoQ biosynthetic enzymes have been focused on model organisms such as *E. coli* and *S. cerevisiae*. The *E. coli* proteins of CoQ biosynthesis are typically named with a prefix "Ubi," while "Coq" or "COQ" has been used to refer to the yeast and human proteins. Beginning with the generation of isopentenyl diphosphate unit and the aromatic ring precursor in the cytosol, CoQ biosynthesis proceeds at the inner mitochondrial membrane. Chemically, a key feature of ubiquinone is the fully substituted *para*-quinone ring, which is afforded by series of decorations of the prenylated aromatic head (Fig. 1). First, a hydroxyl group is added by Coq6 to *C*5 (*5*), followed by *O*-methylation by Coq3 (*6*). Enzymes catalyzing the *C*1-decarboxylation and the *C*1-hydroxylation remain undefined in eukaryotes. Following the Coq5-mediated *C*-methylation at *C*2 to form demethoxy–coenzyme Q (DMQ) (*7*), Coq7 (*8*) in yeasts and mammals adds a hydroxyl group to *C*6, which is, finally, *O*-methylated by Coq3 again (Fig. 1).

The eukaryotic CoQ biosynthetic genes that have been characterized to date are conserved across phylogenetic linages (table S1) (*9*). However, plants lack a protein homologous to Coq7 (*10*), the penultimate enzyme that saturates the *para*-quinone ring (Fig. 1) and a central regulator of the CoQ biosynthetic pathway (*11*, *12*). Coq7 is a di-iron protein and conserved in a range of eukaryotes, including *S. cerevisiae* (*8*), *Caenorhabditis elegans* (*13*), and human (*14*), as well as in some proteobacteria (*15*). In *E. coli*, the *C*6-hydroxylation is catalyzed by a flavin-dependent monooxygenase, UbiF (*15*). However, searching plant genomes did not result in identification of a UbiF homolog. It is likely that plants use a different, but as yet unidentified, enzyme to catalyze this reaction.

Here, we report that CoqF, a unique flavin-dependent monooxygenase with specific residues, is responsible for *C*6-hydroxylation of the ubiquinone aromatic ring in *A. thaliana*. CoqF emerged during the rise of major eukaryotic groups to substitute the di-iron Coq7. Besides being ubiquitous in plants, CoqF is widely distributed in green algae and several other major groups of eukaryotes, including *Plasmodium falciparum* and other apicomplexan parasites, but not in Metazoa and fungi.

RESULTS

A flavoenzyme from *Arabidopsis* is a functional *C*6-hydroxylase of CoQ
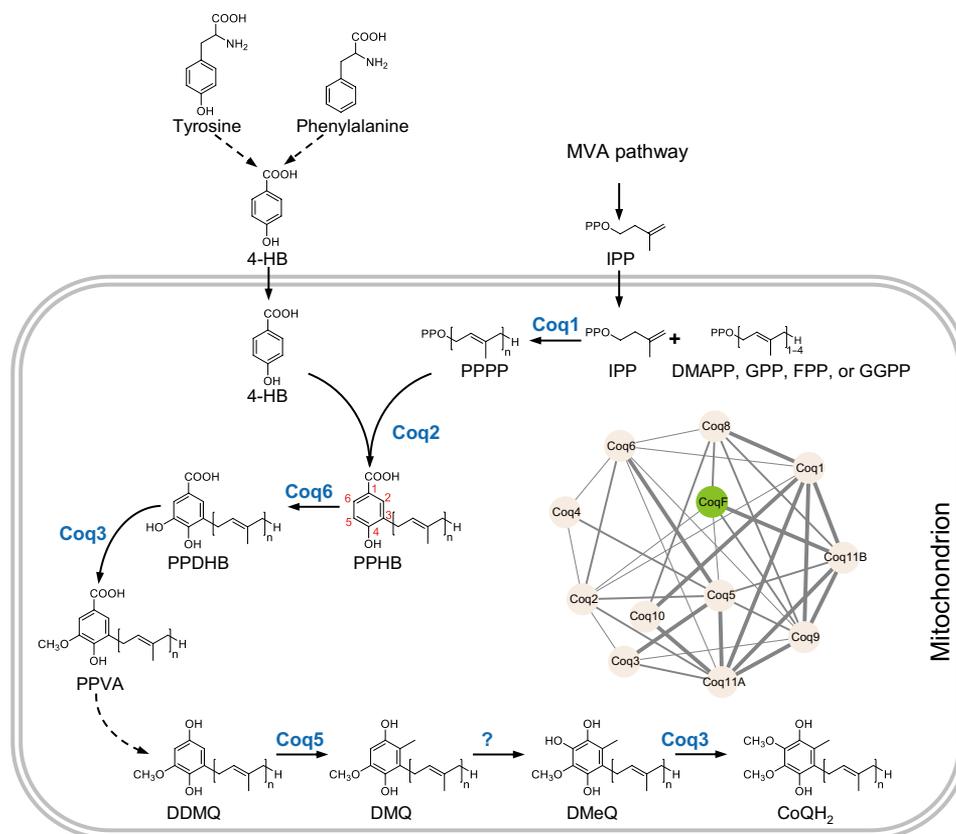
Arabidopsis genes involved in CoQ biosynthesis have been shown to form a coexpression network (*4*). To identify candidate genes encoding the CoQ *C*6-hydroxylase in plants, we mined the ATTED-II database (http://atted.jp/) and found that each of the identified CoQ biosynthetic pathway genes coexpressed with at least three other *Coq* genes (Fig. 1). From the top candidates pooled from the coexpression analysis, four genes (*At1g24340*, *At1g64950*, *At3g14690*,

Fig. 1. The CoQ biosynthetic pathway of plant and the coexpression network. The inset shows coexpression network, which was reconstituted from the top 1000 coexpressed genes from ATTED-II (http://atted.jp/) using *Arabidopsis* CoQ biosynthetic genes *Coq1* (At2g34630), *Coq2* (At4g23660), *Coq3* (At2g30920), *Coq4* (At2g03690), *Coq5* (At5g57300), *Coq6* (At3g24200), *Coq8* (At4g01660), *Coq9* (At1g19140), *Coq10* (At4g17650), *Coq11A* (At5g10730), and *Coq11B* (At5g15910) as queries. The network was drawn on the basis of rank of correlation. The top 100, top 101 to 500, and top 501 to 1000 coexpressed genes from every gene were used as bold, normal, and thin edges, respectively. Abbreviations: 4-HB, 4-hydroxybenzoate; IPP, isopentenyl pyrophosphate; DMAPP, dimethylallyl pyrophosphate; GPP, geranyl pyrophosphate; FPP, farnesyl pyrophosphate; GGPP, geranylgeranyl pyrophosphate; MVA, mevalonic acid; PPPP, polyprenyl-pyrophosphate; PPHB, polyprenyl-hydroxybenzoate; PPDHB, polyprenyl-dihydroxybenzoate; PPVA, polyprenyl-vanillic acid; DDMQ, demethoxy-demethyl-coenzyme Q; DMQ, demethoxy-coenzyme Q; DMeQ, demethyl-coenzyme Q.

and *At5g07990*), which were annotated as monooxygenases, coexpressed with more than three *Coq* genes. Of these putative enzymes, At1g24340 was predicted by TargetP (*16*) to contain an N-terminal mitochondrial targeting signal (MTS) and coexpressed with *Coq2*, *Coq5*, *Coq8*, *Coq9*, and *Coq11B* (Fig. 1). In yeast, Coq9 is an isoprene lipid–binding protein proposed to deliver DMQ to Coq7 (*17*), and Coq11 coprecipitates with Coq7 (*18*). *At1g24340* encodes a protein of 709 amino acid residues belonging to the flavin-dependent monooxygenase family. Notably, this protein is extended at its C terminus by a long sequence of ~226 amino acids (fig. S1). Although several enzymes of the CoQ biosynthetic pathway are flavin-dependent monooxygenases, including the *C5*-hydroxylase Coq6 and UbiI, the *C6*-hydroxylase UbiF, the *C1*-hydroxylase UbiH (table S1), and the promiscuous UbiM and UbiL (*15*), they share little sequence identity with At1g24340 outside of the flavin adenine dinucleotide (FAD)–binding and DG fingerprint motifs (*19*) and are all smaller (<510 amino acids) without the C-terminal extension (fig. S1).

We transferred At1g24340 into a Δ*ubiF* mutant strain of *E. coli* constructed by CRISPR-Cas9, which was unable to grow on a respiratory medium containing succinate because of the defect in synthesizing CoQ_8 (Fig. 2A). At1g24340 rescued the growth of Δ*ubiF* on succinate, and consequently, we named this AtCoqF. Analysis by

liquid chromatography–multiple reaction monitoring–mass spectrometry (LC-MRM-MS) showed that AtCoqF restored CoQ_8 biosynthesis in this mutant strain, demonstrating a *C6*-hydroxylase activity (Fig. 2, B and C). AtCoq6, encoding the *C5*-hydroxylase, did not complement the Δ*ubiF* mutant (Fig. 2, A to C).

We assayed AtCoqF activity in *S. cerevisiae*. The yeast strain Δ*coq7* cannot grow on nonfermentative medium (Fig. 2D) (*20*). Attempts to rescue growth on ethanol-glycerol medium with the mitochondrion-targeted AtCoqF were, at first, unsuccessful. This was not completely unexpected because the CoQ biosynthetic enzymes are known to form a multisubunit complex in yeast (*21*). AtCoqF (709 amino acids) is much larger than yeast Coq7 (233 amino acids), and the two enzymes use different cofactors. It has been reported that Coq8 has adenosine triphosphatase (ATPase) activity and interacts with cardiolipin-containing membranes and analogs of CoQ biosynthetic precursors in yeast and other organisms (*22*), which may help stabilize the CoQ biosynthetic complex (*23*). When *Coq8* was overexpressed in yeast, AtCoqF indeed restored the Δ*coq7* growth. Analysis of the lipid extracts of each strain indicated that AtCoqF, helped by Coq8, recovered the CoQ_6 biosynthesis in Δ*coq7* cells to over 60% of the wild-type (WT) level (Fig. 2, E and F). In contrast, AtCoqF failed to complement the *S. cerevisiae* Δ*coq6*
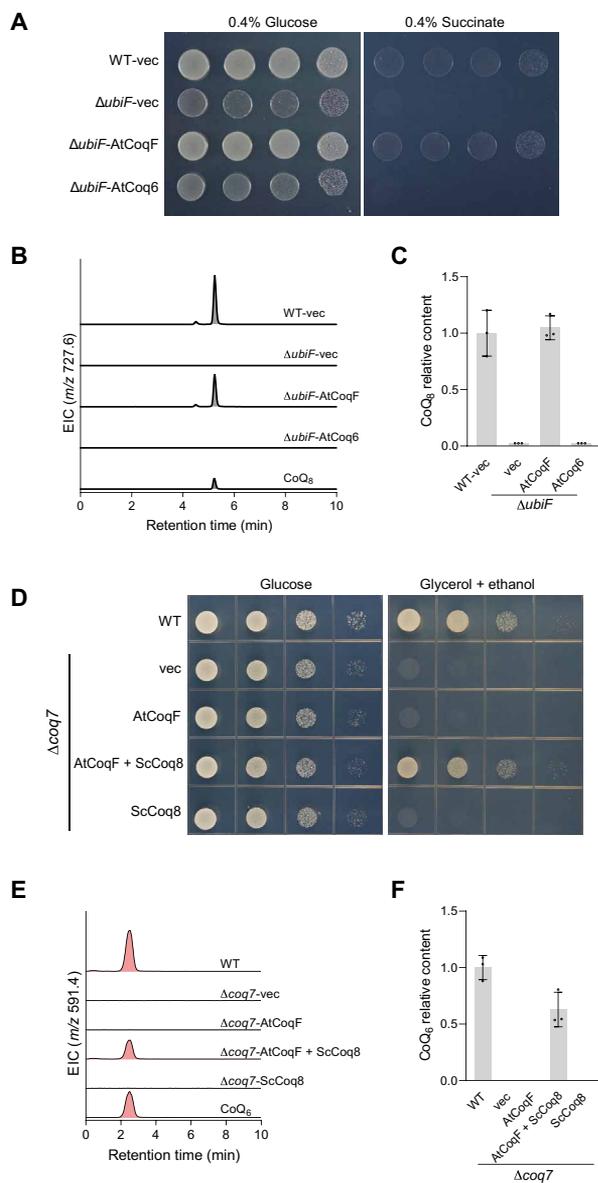
Fig. 2. AtCoqF rescues CoQ biosynthesis in C6-hydroxylase mutants of *E. coli* and *S. cerevisiae*. (**A**) The *E. coli* wild-type (WT) strain MG1655 was transformed with pTrc99a empty vector (vec). The *E. coli* Δ*ubiF* mutant strain, impaired in C6-hydroxylation, was transformed with the empty pTrc99a plasmid, the plasmid harboring *AtCoqF* or *AtCoq6*, respectively. Following growth in LB medium containing 0.4% glucose, serial dilutions were spotted onto M9 minimal medium supplemented with either 0.4% glucose or 0.4% succinate. The plates were incubated for 12 hours (glucose) or 24 hours (succinate) at 37°C. (**B**) Extracted ion chromatograms (EIC) of CoQ_8 [mass/charge ratio (*m/z*) 727.6] extracts from *E. coli* cells grown in LB medium with 0.4% glucose. (**C**) Quantification of the CoQ_8 content in (B), analyzed by LC-MRM-MS. Data are means of three biological replicates ± SE. (**D**) Growth of *S. cerevisiae* strains transformed with the indicated plasmids on SC-Ura-His medium with glucose and nonfermentable carbon source ethanol/glycerol: WT strain BY4742 transformed with the empty vectors or the *coq7* mutant strain transformed with the empty vectors (vec), pRS426-*AtCoqF* and empty pRS423 (AtCoqF), pRS426-*AtCoqF* and yeast *Coq8* on pRS423 (AtCoqF+ScCoq8), and empty pRS426 plus yeast *Coq8* on pRS423 (ScCoq8). Serial dilutions were spotted and incubated for 1 day (glucose) or 3 days (ethanol/glycerol) at 30°C. (**E**) EIC of CoQ_6 (*m/z* 591.4) in extracts of yeast cells grown in SC-Ura-His liquid medium with 2% glucose. (**F**) Quantification of CoQ_6 contents in (E). Data are means of three biological replicates ± SE.

strain lacking a functional *C*5-hydroxylase (fig. S2). These results indicated that in heterologous systems, AtCoqF functions specifically as a *C*6-hydroxylase of the ubiquinone head ring.

High *AtCoqF* expression and CoQ accumulation in seeds

To determine AtCoqF subcellular localization experimentally, stably transformed *Arabidopsis* plants carrying the *AtCoqF-GFP* fusion were generated. Observation by confocal microscopy colocalized the GFP signal with a mitochondrion-specific dye MitoTracker Red (fig. S3). In addition, we fused *E. coli* UbiF and human COQ7 to the AtCoqF MTS at their N-termini and green fluorescent protein (GFP) at their C-termini, respectively. When transiently expressed in *Nicotiana benthamiana* leaves, GFP-tagged MTS-UbiF or MTS-HsCOQ7 also localized in mitochondria (Fig. 3A), further confirming that AtCoqF is targeted to mitochondrion via the N-terminal MTS. Detection by LC-MRM-MS showed that the major form (CoQ_9) and minor form (CoQ_{10}) CoQ were distributed throughout the *A. thaliana* plant with a higher level in seeds, where the *AtCoqF* transcripts also were most abundant (Fig. 3, B and C).

CoqF is required for embryo development

To identify the function of AtCoqF in planta, two transfer DNA (T-DNA) lines of *Arabidopsis* with insertion in *At1g24340*, *coqf-1* (emb2421), and *coqf-2* (SALK_073461), were selected. The former, with the T-DNA disrupting the last exon, was previously isolated as an embryo-defective mutant on the SeedGenes project (http://seedgenes.org/), whereas *coqf-2* has the insertion in the first intron (Fig. 4A). No seedlings homozygous for either mutation could be recovered from seeds of heterozygous plants after self-pollination, and the segregation ratio for WT to heterozygous plants was close to the expected 1:2 (table S2). Aborted seeds (white or retarded in growth) were found in green siliques of heterozygous plants of both mutants (Fig. 4B), suggesting that the homozygous *coqf* mutants were embryo lethal, as reported for another CoQ biosynthetic gene, *AtPPT1*, which encodes the aromatic head polyprenyl diphosphate transferase (*3*).

We then transformed heterozygous *coqf-1* plants with *AtCoqF* under the control of *Arabidopsis Ubiquitin 10* (*Ubi10*) promoter, which is active in the embryo at an early stage and onward (*24*). Homozygous *coqf-1* plants were then isolated in the T1 generation (fig. S4). The T3 plants were subsequently screened for plants heterozygous for *coqf-1* and homozygous for *pUbi10::AtCoqF*, and all seeds in siliques were normal (Fig. 4B), further supporting that AtCoqF is absolutely required for embryo development.

In leaves, the *coqf-1*$^{−/−}$/*pUbi10::AtCoqF*$^{+/+}$ transgenic plants accumulated CoQ_9 comparable to the WT plant (Fig. 4, C to E). To overcome the lack of homozygous mutant material deprived of CoqF activity due to embryo abortion, and to see whether this flavin-dependent enzyme is functionally conserved in other plant species, we performed virus-induced gene silencing (VIGS) to suppress *CoqF* expression in *N. benthamiana*. The content of CoQ_{10} in VIGS-*NbCoqF* leaves was reduced by 85% compared with control infiltrations, where the CoqF substrate DMQ_{10} accumulated (Fig. 4, F to H). These data confirmed that CoqF catalyzes DMQ hydroxylation in the CoQ biosynthetic pathway in *N. benthamiana* and in *Arabidopsis*.

We asked whether the *C*6-hydroxylases of other types could function in *Arabidopsis*. The coding sequences of *E. coli* UbiF and human COQ7 were placed behind the MTS of AtCoqF and driven by the *Ubi10* promoter (fig. S4A). Plants homozygous for *coqf-1*
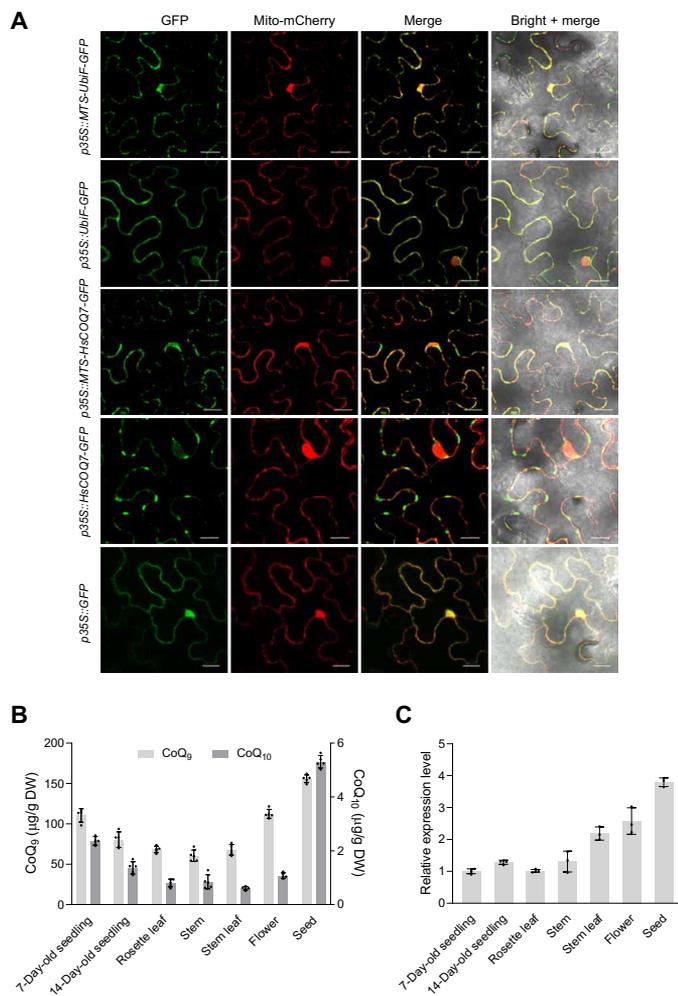
Fig. 3. AtCoqF is localized in mitochondria and highly expressed in seeds. (**A**) AtCoqF contains an N-terminal MTS. The MTS from AtCoqF was fused to the N termini of UbiF or HsCOQ7 and expressed transiently with the mCherry-fused mitochondrial marker (CD3-991) in *N. benthamiana* leaves. Free GFP was used as a control. Merged panels show the overlay of GFP and mCherry fluorescence. Viewed with confocal microscope. Scale bars, 20 μm. (**B**) CoQ contents in seedling, rosette leaf, stem, and cauline leaf of 5-week-old plants, flower (fully opened), and mature seed. Data are means of four to five biological replicates ± SE. (**C**) Relative expressions of *AtCoqF* in the organs as in (B) analyzed by qRT-PCR. Data are means of three to four biological replicates ± SE.

were obtained in the T1 generation for both constructs, indicating that the embryo-lethal phenotype was rescued (fig. S4B). Furthermore, the two alien proteins supported CoQ_9 biosynthesis in the homozygous *coqf-1* background to considerable levels (Fig. 4, I to L). Thus, both UbiF and HsCOQ7 could be accepted by *Arabidopsis* mitochondria for the synthesis of ubiquinone. These data, together with the results from *S. cerevisiae*, demonstrate that CoqF and Coq7 are functional equivalents in the mitochondrial CoQ biosynthetic pathway and can replace each other.

CoqF and homologous proteins belong to an isolated subfamily of flavoenzymes

To investigate how widespread CoqF proteins are in nature, a protein sequence similarity network (SSN) of AtCoqF was generated with the Enzyme Function Initiative–Enzyme Similarity Tool (EFI-EST) (https://efi.igb.illinois.edu/), which uses sequences from the UniProt database. The SSN showed two major clusters (Fig. 5A). Cluster I, to which AtCoqF belongs, comprises sequences from plants, green algae, and SAR (Stramenopiles, Alveolata, and Rhizaria), whereas cluster II is composed of bacterial and fungal proteins. Accordingly, phylogenetic trees were constructed on the basis of representative sequences selected from the two clusters (Fig. 5B and fig. S5).

The proteins of the CoqF subfamily as a whole are distant from other flavin-dependent enzymes involved in CoQ biosynthesis (fig. S6). To test whether the two clusters had diverged functionally, genes encoding proteins of 12 representative species were introduced into the Δ*ubiF* mutant strain of *E. coli*. We found that the cluster I proteins selected from different eukaryotic groups all restored the CoQ_8 biosynthesis in *E. coli* Δ*ubiF* cells, suggesting that these organisms have the ability to use CoqF to synthesize ubiquinone. In contrast, neither bacterial nor fungal proteins of cluster II conferred any CoqF activity in *E. coli* (Fig. 5B). Moreover, cluster II proteins have already been identified as 2,4-dichlorophenol 6-monooxygenases (tfdB) (*25*) and 2-hydroxybiphenyl-3-monooxygenase (hbpA) (*26*) (Fig. 5A), but AtCoqF was inactive toward these phenolics (fig. S7). These data indicate that the cluster II enzymes act in other pathways, and at least some of them are scavengers of phenolic compounds. Furthermore, analysis by TargetP indicated that the fungal proteins from cluster II lack an MTS, whereas the cluster I proteins have MTSs (Fig. 5C).

Multisequence alignment of the proteins identified two phenylalanine residues, Phe[121] and Phe[283], which are highly specific to functional CoqFs (Fig. 5D and fig. S5). Structural modeling of AtCoqF with a partial substrate [DMQ with two isoprene units (DMQ_2)] in the active site revealed that the side chains of Phe[121] and Phe[283] make important hydrophobic interactions with the aromatic head group to stabilize the substrate (Fig. 5E). Another amino acid, His[88], is most likely the residue involved in substrate deprotonation, similar to His[48] in hbpA (*27*). To analyze these three residues, AtCoqF mutant variants were generated and expressed in the Δ*ubiF* strain of *E. coli* (Fig. 5F). Substitution of any of these three residues (His[88], Phe[121], and Phe[283]) with an alanine reduced the CoQ_8 levels substantially in Δ*ubiF* cells. Exchanging Phe[121] and Phe[283] with the respective residues in fungal or bacterial class II homologs impaired CoqF activity to differing degrees, and when both were replaced, the mutant enzymes were almost completely inactive (Fig. 5F). We propose that acquisition of these two phenylalanine residues in the ancestral protein and the addition of an N-terminal mitochondria–targeting peptide paved the way to the birth of a functional ubiquinone *C*6-hydroxylase in eukaryotes, probably earlier than 1.7 billion years ago (Fig. 5B).

Ramification of CoqF in eukaryotes

To examine the distribution of CoqF and Coq7 in eukaryotes, we analyzed the protist (Ensembl Protists), algal (PhycoCosm), and plant (Ensembl Plants) genomes, as well as the EukProt database (*28*), which contained 742 eukaryotic species representing known eukaryotic diversity. Both CoqF and Coq7 are widely distributed in eukaryotes (Fig. 6A), covering 253 and 236 species in these datasets, respectively (table S3).

The current eukaryote tree is divided into the two major domains of Diaphoretickes (plants, algae, SAR, etc.) and Amorphea (animals,
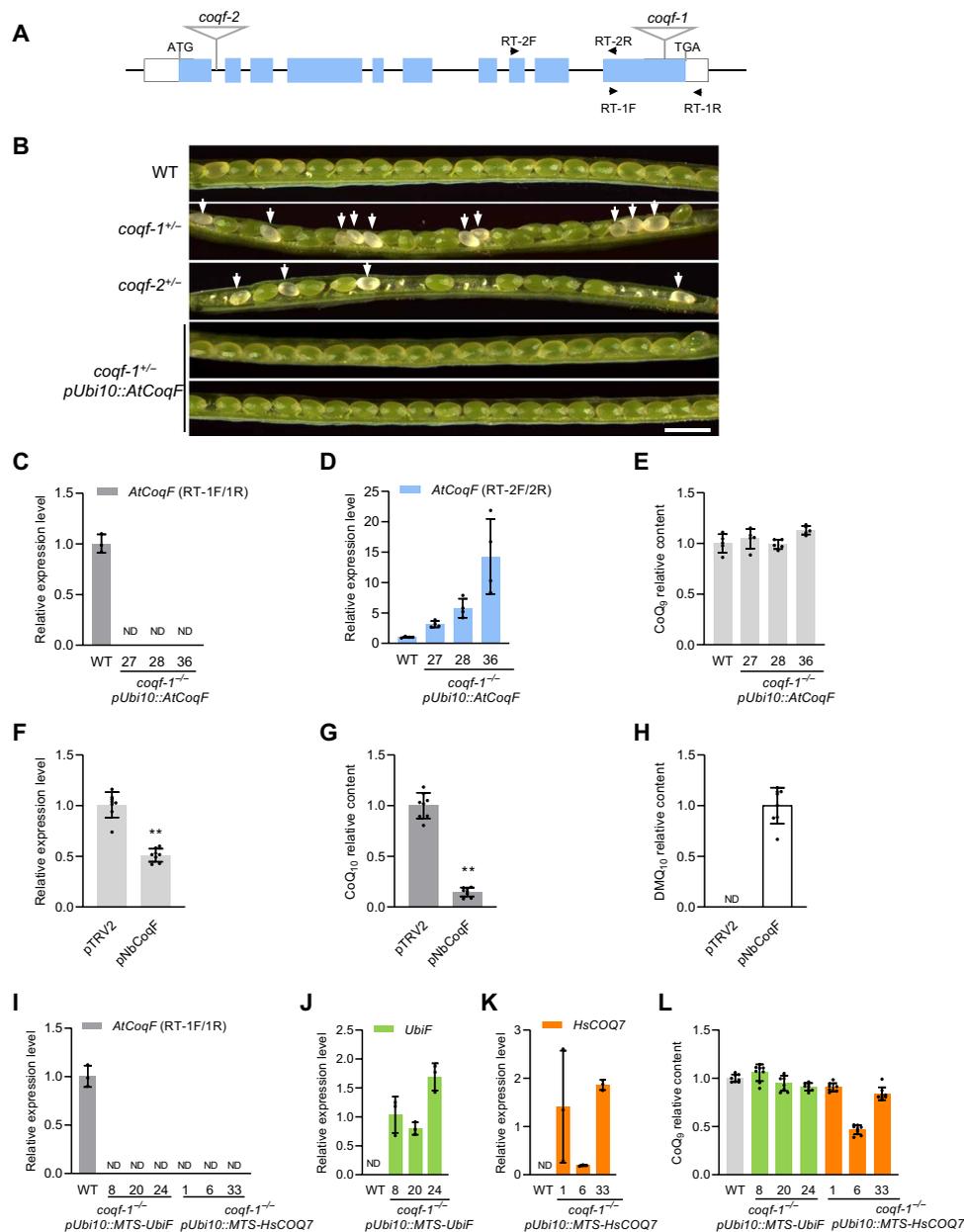
Fig. 4. CoqF is involved in plant CoQ biosynthesis. (**A**) Graphical representation of *AtCoqF* and the two T-DNA insertions. Exons and untranslated regions are indicated by blue and white boxes, respectively. The endogenous *AtCoqF* transcripts were detected by primers RT-1F and RT-1R, and transcriptions of both *AtCoqF* and *pUbi10::AtCoqF* transgene were detected by primers RT-2F and RT-2R. (**B**) Opened siliques of WT, heterozygous *AtCoqF* mutants, and *coqf-1*[+/−] plants transformed with *pUbi10::AtCoqF*. White arrows identify abnormal seeds, undeveloped seeds not indicated. Scale bar, 1 mm. (**C** and **D**) Transcript level of *AtCoqF* (C) and of *AtCoqF* plus *pUbi10::AtCoqF* (D) in T3 complementation lines. Data are means of three to four biological replicates ± SE. (**E**) Relative content of CoQ_9 in T3 plants of *pUbi10::AtCoqF* in the homozygous *coqf-1* background, extracted from rosette leaves of 4-week-old plants. Data are means of five biological replicates ± SE. (**F**) Relative expression of *NbCoqF* in *N. benthamiana* leaves following VIGS (pNbCoqF), 2 weeks after treatment. pTRV2, empty vector control. (**G** and **H**) Relative contents of CoQ_{10} (G) and DMQ_{10} (H) in *NbCoqF*-silenced leaves, compared to pTRV2 control. **$P < 0.01$, determined by two-tailed Student's *t* test. Data are means of eight biological replicates ± SE. (**I** to **L**) *AtCoqF* (I), *UbiF* (J), and *HsCOQ7* (K) expression in T2 plants of *pUbi10::MTS-UbiF* or *pUbi10::MTS-HsCOQ7* in the homozygous *coqf-1* background, measured by qRT-PCR. Data are means of three biological replicates ± SE. (L) Relative content of CoQ_9 in T2 plants of *pUbi10::MTS-UbiF* or *pUbi10::MTS-HsCOQ7* in the homozygous *coqf-1* background. Data are means of eight biological replicates ± SE. ND, not detected.

fungi, Amoebozoa, etc.), plus several additional clades (*29*). Coq7 was found across all major lineages of eukaryotes (Fig. 6A, data files S1 and S2), suggesting that it was inherited from the common ancestor of eukaryotes. Distinctly, CoqF is primarily distributed in

Diaphoretickes (Fig. 6, A and B). In several groups (9 of 40 in Fig. 6A), mainly those of algae and SAR, both CoqF and Coq7 are present but usually in different sublineages. In general, eukaryotic cells seem to support just one type of *C6*-hydroxylase, with the
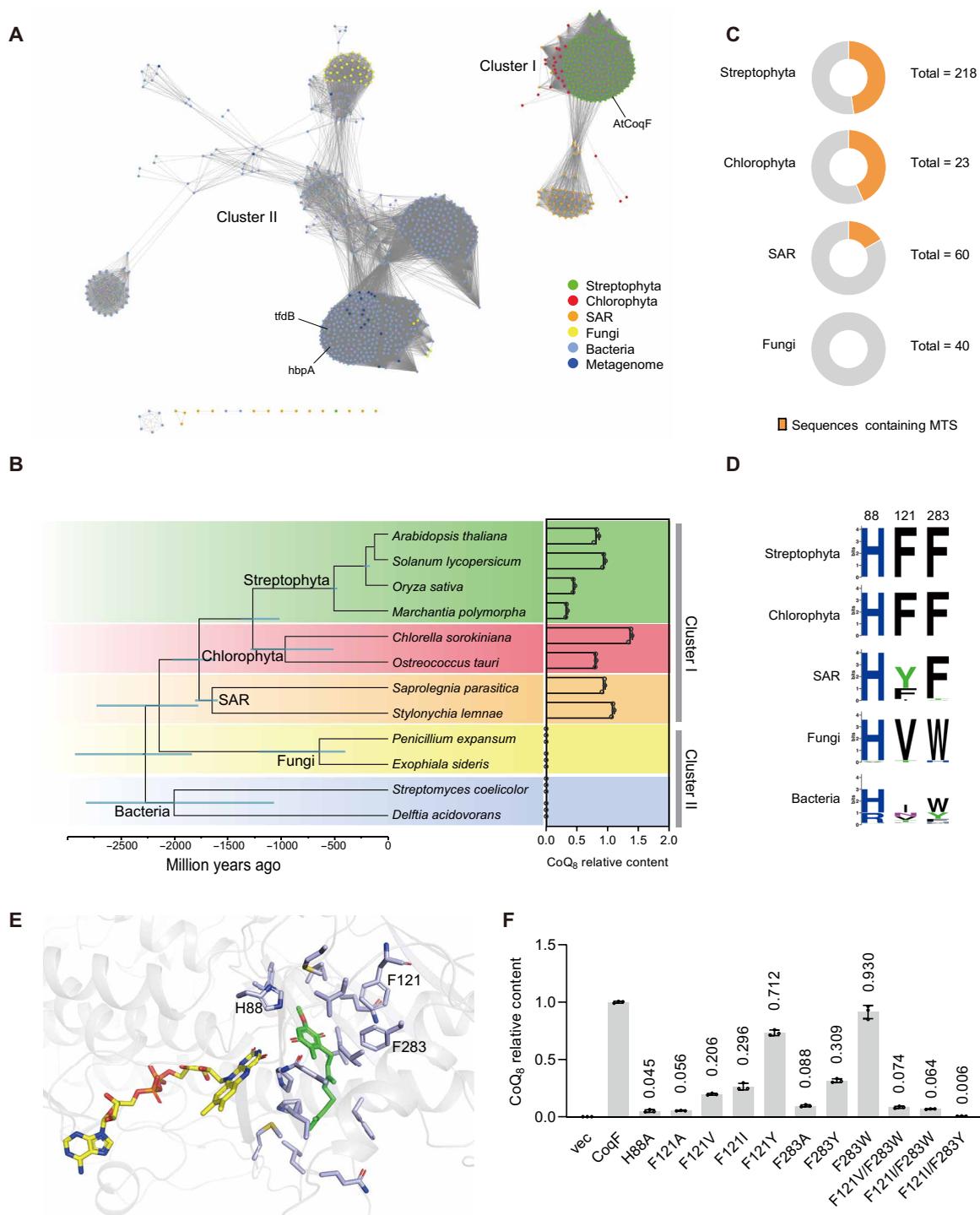
Fig. 5. Characterization of CoqF proteins. (**A**) Sequence similarity network (SSN) of CoqF calculated by EFI-EST (threshold score 85). Nodes are colored by taxonomy. AtCoqF, tfdB of *Delftia acidovorans*, and hbpA of *Pseudomonas nitroreducens* are indicated. (**B**) A time-calibrated phylogenetic tree of CoqF homologs in clusters I and II selected from SSN. The tree was constructed by IQ-TREE under LG+F+R4 and BEAST2. Bars at nodes are the 95% highest probability density. *CoqFs* from each species were introduced into *E. coli* Δ*ubiF* strain, and the CoQ$_8$ content in WT strain was set to 1. Data are means of three biological replicates ± SE. (**C**) MTSs of proteins in the SSN, predicted by TargetP. The absence of MTS from some proteins could be caused by incomplete sequences deposited or because not all MTSs could be recognized by TargetP. (**D**) Distributions of amino acid residues corresponding to AtCoqF His[88], Phe[121], and Phe[283]. Logos were generated using the WebLogo (http://weblogo.berkeley.edu/). (**E**) Docking of DMQ$_2$ in the active site of CoqF. DMQ$_2$ (green) and FAD (yellow) are shown as stick models. (**F**) Activity of CoqF variants. Quantification of the CoQ$_8$ content in *E. coli* Δ*ubiF* strains grown in LB medium with 0.4% glucose. Data are means of three biological replicates ± SE.
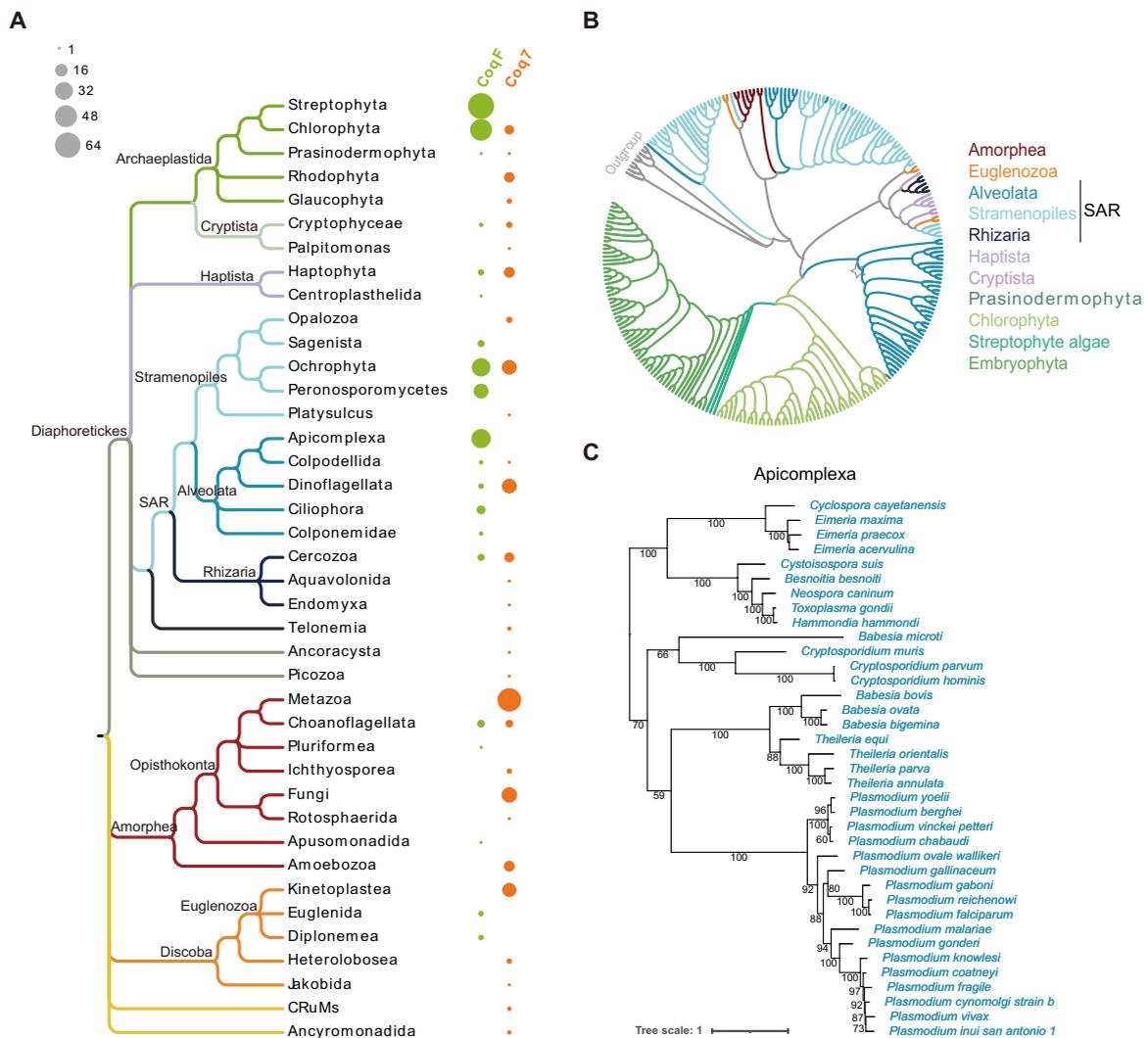
Fig. 6. Distribution of CoqF and Coq7 in eukaryotes. (**A**) Topology of the eukaryotic tree, drawn as proposed (*28*, *59*), showing CoqF and Coq7 distributions. Circle sizes represent the number of species containing the protein. The figure was created with EvolView (https://evolgenius.info/evolview/). (**B**) Maximum likelihood phylogenetic tree of eukaryotic CoqF proteins. The tree was derived from an MAFFT alignment and constructed using IQ-TREE under the LG+F+ R9 model selected by ModelFinder. The tree was rooted using cluster II proteins from the SSN as outgroup. A fully annotated version of this tree is shown in data file S3. The apicomplexan clade shown in (C) is indicated by a star. (**C**) The branch of the CoqF phylogenetic tree formed by apicomplexan species, which are mostly pathogenic parasites.

exception of 17 species in the dataset, which harbor genes encoding both enzymes (table S3), although artifacts introduced during sampling cannot be excluded completely at this time, qualifying this conclusion.

Most organisms carrying CoqF can be assigned to phototrophic lineages of primary endosymbiosis (Archaeplastida) or secondary endosymbiosis (Cryptophyceae, Haptophyta, Ochrophyta, Myzozoa, and Chlorarachniophyceae), including the two secondary plastid-bearing algal species (*Guillardia theta* and *Bigelowiella natans*) that have sequenced genomes (*30*). However, the parasitic apicomplexans of Alveolata also have CoqF, as do the phytopathogenic oomycetes (Peronosporomycetes) of Stramenopiles (table S3 and data file S3). Notably, apicomplexan species are unique in having vestigial and nonphotosynthetic plastids called apicoplasts (*31*), and those parasites infecting human and animals, such as *P. falciparum*, *Cryptosporidium parvum*, *Neospora caninum*, and *Toxoplasma gondii*,

all have CoqF instead of Coq7, which is present in their hosts (Fig. 6C and table S3).

The phylogenetic tree of CoqFs (Fig. 6B and data files S3 and S4) complies largely with the branches of Diaphoretickes proposed recently (*29*), suggesting that CoqF evolved shortly after Diaphoretickes diverged. Exceptionally, in other domains of eukaryotes dominated by Coq7, eight species of Amorphea and six species of Discoba were found to have CoqF (Fig. 6, A and B; table S3; and data file S3). In Discoba, the Euglenida species, such as *Euglena gracilis* and *Euglena mutabilis*, are photosynthetic. The presence of CoqF outside Diaphoretickes may be the result of eukaryote-to-eukaryote lateral gene transfer, possibly facilitated by endosymbiosis. In the CoqF phylogenetic tree, the sequences from Amorphea and Discoba are indeed dispersed and discretely branched with SAR and nongreen algal homologs (Fig. 6B and data file S3), suggesting that gene transfer occurred more than once.

Both types of the CoQ *C6*-hydroxylase are present in Euglenozoa. Of the three subgroups of Euglenozoa, euglenids and diplonemids have CoqF, whereas kinetoplastids have Coq7 (Fig. 6A). While euglenozoans share the characteristic discoidal mitochondrial cristae, kinetoplastids are distinct in having a complex mitochondrial genome structure (*32*). The use of two different enzymes for a key step of CoQ biosynthesis adds another difference between these morphologically distinctive mitochondria.

The rise of CoqF in Viridiplantae

The Archaeplastida comprises Glaucophyta, Rhodophyta (red algae), and Viridiplantae (green algae and plants). In contrast to the wide occurrence of CoqF in Viridiplantae, Coq7 is present in Rhodophyta and Glaucophyta (Fig. 6A, table S3, and data file S1), including *Cyanophora*, the glaucophyte model taxon (*33*). This distribution pattern suggests a complete or at least a tendency for loss of CoqF from Rhodophyta and Glaucophyta, two sister groups of Viridiplantae. Notably, in the phylogenetic tree of Coq7, the red algal proteins form a monophyletic group, which is distantly separated from other algal Coq7s but otherwise embedded in the Amorphea branches (data file S1), arguing strongly for an origin by lateral gene transfer. Possibly, the Rhodophyta ancestor regained Coq7 from a protozoan invader and then discarded its own enzyme.

The Viridiplantae have undergone alternative losses and CoqF proliferation. In Chlorophyta, CoqF is found in 47 species and Coq7 in 9 species, and 1 species (*Caulerpa lentillifera*) may have both. While Coq7 is relatively frequent in extant representatives of the early branches of the Chlorophyta (prasinophytes), including all five species of Pyramimonadophyceae examined, CoqF is predominant in core chlorophytes (Fig. 7A), which flourished from the Late Cryogenian and Early Ediacaran periods ca. 650 million years ago (*34*). In Prasinodermophyta, both CoqF (*Prasinoderma coloniale*) and Coq7 (*Prasinococcus capsulatus*) are found (Fig. 7A) and are allied with Mamiellophyceae or Pyramimonadophyceae counterparts

(data files S1 and S3), supporting it as a sister group to Chlorophyta (*35*), probably derived from an ancestor that harbored both enzymes. However, these interpretations must be considered preliminary because, currently, the available sequence data are limited.

Streptophyte algae are considered a sister group to Embryophyta (*36*). Searches of the Joint Genome Institute (JGI) Algal multiomics resource PhycoCosm revealed the presence of CoqF in extant streptophyte algae (Fig. 6B and table S3). As both Coq7 and CoqF species are present in Chlorophyta (Fig. 7A), Coq7 was most likely eliminated from Streptophyta before terrestrialization. Similar to the exclusive occurrence of Coq7 in Metazoa (animals), CoqF is ubiquitous in Embryophyta (land plants) (Fig. 7B), exemplifying the association of the two types of CoQ biosynthetic pathways with different trophic styles of eukaryotes in the terrestrial world.

Analysis of the numbers of *CoqF* genes in crop species showed that there is one in the diploid *Oryza sativa*, *Sorghum bicolor*, and *Solanum lycopersicum*, two in the palaeopolyploid *Glycine max*, and three in the hexaploid *Triticum aestivum* (Fig. 7B). Thus, CoqF is likely a single-copy housekeeping gene in the diploid genomes of plants, well suited for serving as a phylogenetic maker.

DISCUSSION

Previous reports have shown that an ancestral Coq7 was likely present in the common ancestor of alpha-, beta-, and gamma-proteobacteria (*15*). Given the hypothesis that the mitochondrion was derived from an alpha-proteobacterium (*37*), the eukaryotic common ancestor may have acquired a *Coq7* gene from alpha-proteobacteria through lateral gene transfer, possibly as a consequence of endosymbiosis, which has been maintained in almost all major eukaryotic lineages. We propose that the last eukaryotic common ancestor had both Coq7 and ancestral CoqF homologs, and the latter gained the CoqF activity in the phyletic group, which later branched into Archaeplastida, Cryptista, Haptista, and SAR (Figs. 5B and 6A).
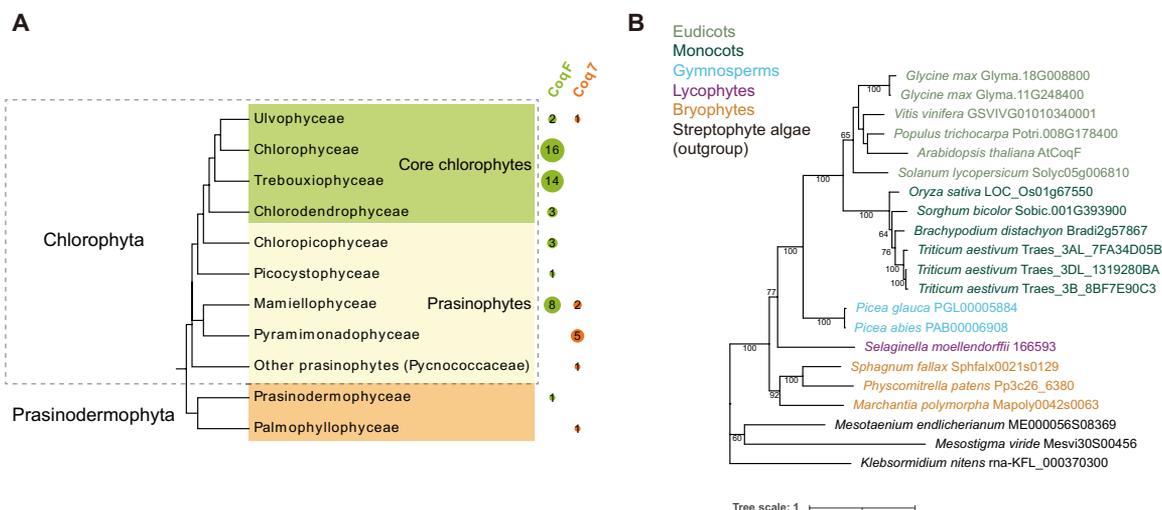
Fig. 7. The CoqF proliferation in Viridiplantae. (**A**) Distribution of CoqF and Coq7 in Chlorophyta and Prasinodermophyta. CoqF is increasingly abundant in core chlorophytes. Topology of the tree is drawn as proposed (*60*). (**B**) Maximum likelihood phylogenetic tree of CoqF in Embryophyta. The protein sequences of gymnosperms were obtained from the PLAZA project (https://bioinformatics.psb.ugent.be/plaza/), and the others from Phytozome (https://phytozome.jgi.doe.gov/). The tree was derived from an MAFFT alignment and constructed using IQ-TREE with a JTT+I+G4 model. Bootstrap values (on the basis of 1000 replicates) ≥50 are shown. CoqFs of streptophyte algae were used as outgroup.

The complementary distributions of CoqF and Coq7 across these eukaryotic groups suggest that both enzymes coexisted in cells before alternative eliminations, which occurred independently in each major lineage more than once, and may have been associated with or stimulated by physiological innovations that served as a driving force behind the diversification of eukaryotes.

CoqF belongs to the flavin-dependent monooxygenase family, while Coq7 is a ferritin-like protein that contains a di-iron center. Starting ~2.4 billion years ago, the thriving of photosynthetic organisms led to elevated levels of atmospheric oxygen and reduced iron availability in oceans (38). Substitution of the di-iron enzyme in an essential pathway with a flavin-dependent monooxygenase in some eukaryotes, especially in photoautotrophs, could have been an evolutionary response to iron limitation, as suggested for ferredoxin replacement by flavodoxin in some algae, including dinoflagellates and diatoms (39). In land plants, CoqF is the sole C6-hydroxylase. Compared to di-iron monooxygenase, flavin-containing proteins might be less vulnerable to damage by reactive oxygen species (ROS) that are often very high in photosynthetic cells (40). The CoqF distribution pattern reflects that of another flavin-dependent enzyme present in mitochondria, L-galactonolactone dehydrogenase (GLDH), the terminal enzyme of ascorbic acid (vitamin C) biosynthesis, which is present in eukaryotes of photosynthetic lineages that can make ascorbic acid without producing ROS as by-products of synthesis (41).

Unlike plants, which generate a large proportion of their ATP via photophosphorylation, animals rely on oxidative respiration for energy. Consequently, maintaining a sufficient level of CoQ could be a selective advantage for animals during competition. In humans, CoQ deficiency is inevitably associated with aging and deteriorating health. Ironically, at least in experimental conditions, plants are more tolerant of alterations in the Krebs cycle and even perform better under mild reductions in mitochondrial succinate dehydrogenase activity that reduces ubiquinone to ubiquinol (42, 43). Arabidopsis mutants displaying ~60% decrease in CoQ content appeared normal in growth and development (44). However, CoQ is absolutely required for embryogenesis in Arabidopsis and, hence, the completion of the plant life cycle (3, 4). Whether or not the use of a flavin-dependent enzyme for the final step in CoQ ring substitution facilitated the transition of plant ancestors from water to land, the tendency of CoqF dominance among green eukaryotes is clear.

The CoQ biosynthetic pathway of apicomplexans provides important drug targets against parasites (45). CoqF is mechanistically different from Coq7 in mammalian hosts and structurally distinct among flavoenzymes, rendering it an ideal target for the development of drugs for the treatment of malaria, cryptosporidiosis, neosporosis, toxoplasmosis, and other parasitic diseases caused by apicomplexan protists, which remain important health threats to both humans and livestock.

MATERIALS AND METHODS
Plant materials and chemicals
All A. thaliana lines used are in the Columbia-0 (Col-0) background. The T-DNA insertional mutants of SALK_073461 and emb2421 were obtained from the Nottingham Arabidopsis Stock Center. Primers used in this investigation are listed in table S4.

Arabidopsis and N. benthamiana plants were grown in greenhouse at 22°C under long-day condition (16 hours light and 8 hours dark). CoQ_4, CoQ_6, CoQ_8, CoQ_9, CoQ_{10}, 2,4-dichlorophenol, 3,5-dichlorocatechol, 2-hydroxybiphenyl, and 2,3-dihydroxybiphenyl standards were purchased from Sigma-Aldrich.

E. coli strain and complementation assays
E. coli MG1655 was used for genome deletion to obtain the $\Delta ubiF$ strain by CRISPR-Cas9 as described (46). Briefly, MG1655 carrying the pCas plasmid was transformed with pTargetT containing the N20 sequence, the single-guide RNA sequence, and the donor DNA used as the genome-editing template. Cells were incubated overnight at 30°C on LB medium containing kanamycin (50 mg/liter), spectinomycin (50 mg/liter), and 0.4% [w/v] glucose. Mutations were identified by colony polymerase chain reaction (PCR) and confirmed by DNA sequencing.

The coding region of AtCoqF was amplified from A. thaliana cDNAs and cloned into pTrc99a vector. The coding regions of OsCoqF (LOC_Os01g67550) and SlCoqF (Solyc05g006810) were amplified from O. sativa and S. lycopersicum cDNAs, respectively. Other sequences tested in Fig. 5B were generated by gene synthesis (GenScript). AtCoqF variants with point mutations were generated by site-directed mutagenesis using the primers shown in table S4 and confirmed by DNA sequencing. The plasmids were transferred into the $\Delta ubiF$ strain. Serial dilutions were spotted onto M9 minimal medium with 0.4% glucose or 0.4% succinate as the only carbon sources, as indicated specifically.

Yeast strain and complementation assays
Yeast strain S. cerevisiae BY4742 and derivates were used. $\Delta coq7$ (Y12381) was purchased from Euroscarf. AtCoqF was cloned into a multicopy yeast shuttle vector pRS426 controlled by the yeast Coq8 promoter and terminator. For heterologous expression in S. cerevisiae, AtCoq6 was codon-optimized through gene synthesis (GenScript) and cloned into pRS426 controlled by the yeast Coq8 promoter and terminator. The MTS (residues 1 to 34) of yeast Coq3 was fused to the N-terminal of AtCoqF and AtCoq6, respectively. pRS423 was used to clone the yeast Coq8 controlled by its native promoter and terminator. Yeast cells were transformed using Frozen-EZ Yeast Transformation kit (ZYMO RESEARCH). Transformants were selected at 30°C on SC-Ura-His (0.67% yeast nitrogen base with ammonium sulfate, and complete amino acid supplement lacking uracil and histidine) plates with 2% glucose. For complementation analyses, serial dilutions of yeast cells were dropped onto SC-Ura-His agar media containing either 2% glucose or 3% glycerol/2.5% ethanol as carbon sources or as indicated specifically.

Analysis of CoQ contents
E. coli cells were grown at 37°C on LB medium containing 0.4% glucose as a carbon source until OD_{600} (optical density wavelength of 600 nm) reached ~2. Cells were centrifuged, washed with water, and the pellet mass was determined. Samples (10 mg wet weight) were extracted with 0.2 ml of isopropanol containing CoQ_4 (0.1 mg/liter) (internal standard) in a sonicator bath for 2 hours.

Yeast cells were grown at 30°C on SC-Ura-His medium with 2% glucose and collected as stated above. Samples (10 mg wet weight) were lysed by the addition of 0.2 ml of methanol followed by vortexing for 90 s. After centrifugation at 12,000 rpm for 10 min, the pellets were extracted with 0.2 ml of isopropanol containing CoQ_4 (0.1 mg/liter) (internal standard) in a sonicator bath for 2 hours and centrifuged at 12,000 rpm for 10 min. The supernatant from two extractions were combined for LC-MS analysis.

For *Arabidopsis* and *N. benthamiana*, samples were harvested, frozen in liquid nitrogen, and freeze dried. Ten milligrams of freeze-dried samples were extracted with 1 ml of isopropanol in a sonicator bath for 2 hours.

CoQ was detected by LC-MRM-MS in positive ionization mode using an Agilent 1260 high-performance liquid chromatography (HPLC) and 6460 Triple Quadrupole LC-MS system. An Agilent XDB-C18 column (2.1 × 50 mm, 3.5-μm particles) was used at 40°C, flow rate of 0.4 ml/min, and with a binary gradient system consisting of solvent A (isopropanol) and solvent B (acetonitrile/H_2O, 7:1, 10 mM ammonium acetate). The percentage of solvent A was increased linearly from 35 to 85% over 6 min. Transitions used for quantification were as follows: CoQ_4, 455.3 > 197.1; CoQ_6, 591.4 > 197.1; CoQ_8, 727.6 > 197.1; CoQ_9, 795.6 > 197.1; CoQ_{10}, 863.7 > 197.1; and DMQ_{10}, 833.7 > 167.0.

In vivo assays for tfdB and hbpA enzyme activities
E. coli–optimized *Delftia acidovorans tfdB* (UniProtKB Q8KN28) and *Pseudomonas nitroreducens hbpA* (UniProtKB O06647) were synthesized (GenScript) and subcloned into pTrc99a vector. The constructs and the plasmid harboring *AtCoqF* were transferred into the *E. coli* MG1655 strain. The strains were initially grown in LB medium containing ampicillin (100 mg/liter) at 37°C to an OD_{600} of 0.4. Next, 2,4-dichlorophenol (0.25 mM) or 2-hydroxybiphenyl (0.25 mM) was added. Cells were then grown for 20 hours at 25°C, harvested by centrifugation, washed with water, and the pellet mass was determined. Samples were extracted with methanol in a sonicator bath for 2 hours.

Samples were analyzed by LC-MRM-MS in negative ionization mode using 1260 HPLC and 6460 Triple Quadrupole LC-MS system. An Agilent XDB-C18 column (2.1 × 50 mm, 3.5-μm particles) was used at 30°C, flow rate of 0.4 ml/min, and with a binary gradient system consisting of solvent A (water with 0.1% formic acid) and solvent B (acetonitrile). The percentage of solvent B was increased linearly from 20 to 70% over 7 min. Transitions used for quantification were as follows: 2,4-dichlorophenol, 161.0 > 125.0; 3,5-dichlorocatechol, 177.0 > 77.0; 2-hydroxybiphenyl, 169.1 > 115.0; and 2,3-dihydroxybiphenyl, 185.1 > 129.1.

Protein subcellular localization
For GFP fusion, the *AtCoqF* coding region was PCR amplified from cDNAs without the stop codon and cloned into the Gateway vector pGWB505 for C-terminal fusion. The construct was introduced into *Agrobacterium tumefaciens* GV3101, and *Arabidopsis* WT plants were transformed by floral dip (*47*). To image mitochondria, transgene seedlings were incubated with 20 nM MitoTracker Red (Thermo Fisher Scientific) for 60 min at room temperature.

The *E. coli UbiF* was PCR amplified from genomic DNA, and the human *COQ7* (UniProtKB Q99807-2) was generated by gene synthesis (GenScript), both without the stop codon. UbiF or HsCOQ7 was placed behind the MTS of AtCoqF (residues 1 to 41) and cloned into pGWB505 vector. As a mitochondrial marker, binary vector CD3-991 (*48*) was coinfiltrated with GFP constructs. After transformation with the constructs, the *A. tumefaciens* GV3101 (pSoup-p19) cells were infiltrated into *N. benthamiana* leaves. Three days later, epidermal cells were observed using an Olympus FV10i confocal laser scanning microscope. Excitation wavelengths were 473 nm for GFP, 578 nm for MitoTracker Red, and 580 nm for mCherry. Emissions were collected at 490 to 540 nm for GFP, 598 nm for MitoTracker Red, and 610 nm for mCherry.

Generation of *Arabidopsis* complementation lines
AtCoqF was introduced into pCAMBIA1300 vector between the *Arabidopsis Ubi10* promoter and the nopaline synthase (*nos*) terminator (*49*). UbiF or HsCOQ7 behind the AtCoqF MTS were similarly cloned into pCAMBIA1300. The constructs were then used to transform the heterozygous *coqf-1* plants of *A. thaliana*.

Quantitative reverse transcription PCR
Total RNAs were extracted using RNAprep Pure Plant Plus Kit (TIANGEN BIOTECH), and cDNAs were synthesized from 500 ng of RNAs using a PrimeScript RT reagent Kit with gDNA Eraser (Takara). A 20-times diluted cDNA sample was used for quantitative reverse transcription PCR (qRT-PCR) using the SYBR Green reagent (Takara) on Mastercycler ep Realplex2 (Eppendorf). *A. thaliana PP2AA3* (At1g13320) (*50*) or *N. benthamiana Ubiquitin C* (AB026056.1) (*51*) were used as internal references.

Virus-induced gene silencing
VIGS assays were performed using the tobacco rattle virus (TRV) vector system as described (*52*). Fragment of *NbCoqF* was amplified by PCR and cloned into pTRV2 vector. *A. tumefaciens* (GV2260) harboring pTRV1 and pTRV2 (or the pTRV2 containing the sequence for silencing) were infiltrated at a 1:1 ration into the two biggest true leaves of 3-week-old plants of *N. benthamiana*. Two weeks later, the leaves were harvested and analyzed.

Sequence similarity networks
The SSN of CoqF was constructed using the EFI-EST tool (https://efi.igb.illinois.edu/) with AtCoqF as the query sequence for a BLAST search of the UniProt database. The maximum BLAST sequences was 1000. Sequences less than 500 amino acids were removed. The network was visualized in Cytoscape software (https://cytoscape.org/) with an alignment score threshold of 85. The e-value cutoff for the analysis was 10^{-5}.

Phylogenetic analysis of flavin-dependent monooxygenases
CoqF homologs in clusters I and II were selected from the SSN. Other flavin-dependent monooxygenases of CoQ biosynthesis were from the literature (*15*). The sequences were then aligned using MAFFT L-INS-i v7.475 (*53*). Phylogenetic trees were constructed using the maximum likelihood method IQ-TREE v2.1.2 (www.iqtree.org/) under the best-fit model selected by ModelFinder (www.iqtree.org/ModelFinder/) with the minimum of Bayesian information criterion (BIC) score (-MFP -cmax 15). Branch supports were assessed with 1000 ultrafast bootstrap approximation (UFBoot) (*54*). The figures were created with iTOL (https://itol.embl.de/).

Distribution patterns of CoqF and Coq7
The predicted proteomes were downloaded from EukProt (https://doi.org/10.6084/m9.figshare.12417881.v2), Ensembl Protists (http://protists.ensembl.org), PhycoCosm (https://phycocosm.jgi.doe.gov), and Ensembl Plants (http://plants.ensembl.org). For CoqF, the sequences of Streptophyta (AtCoqF, A0A438I297, A0A2K1ZJ43, I1LNA7, A0A3Q7H5G3, A0A3P6EAF4, B9EVE7, A0A1B6QNM0, A0A0Q3N317, D8R9Z5, A0A2K1IC09, and A0A2R6X1P7), Chlorophyta (C1N4Y2, C1EJA0, A0A090M1G2, D8U132, A0A2P6TRB5, and E1Z7R8), and SAR (A0A078AZ44, A0A067C612, and A0A081AB79) in SSN were aligned with MAFFT L-INS-i v7.475, and the aligned

sequences were used as a seed alignment for building a hidden Markov model (HMM) using hmmer3 (*55*). For Coq7, the HMM profile PF03232 was downloaded from Pfam (http://pfam.xfam.org/). The CoqF and Coq7 HMM profiles were used to screen the proteomes using an e-value cutoff of 10^{-18}. For each species, only the best HMM result was used for further analysis. Protein sequences were manually searched for the presence of $EX_{n1}EXXHX_{n2}EX_{n3}EXXH$ motifs in Coq7. For CoqF, sequences lacking the three conserved residues were removed by manual inspection.

Before phylogenetic analyses, the regions without homology within sequences were identified and masked using PREQUAL v. 1.02 (https://github.com/simonwhelan/prequal) using a posterior probability threshold of 0.95. The filtered sequences were then aligned using MAFFT and trimmed by trimAl (v1.2.rev57) (*56*) to remove all positions with gaps over 90% (-gt 0.1). Phylogenetic trees were constructed using the maximum likelihood method IQ-TREE v2.1.2 under the best-fit model selected by ModelFinder with the minimum of BIC score (-MFP -cmax 15). Branch supports were assessed with 1000 UFBoot. The figures were created with iTOL.

Structural modeling and ligand docking
Homology models of AtCoqF protein were generated using SWISS-MODEL (https://swissmodel.expasy.org/) according to the template of HbpA (Protein Data Bank: 5BRT) (*27*). Ligand docking was performed using AutoDock4.2.6 with AutoDockTools (*57*).

Protein divergence time estimations
Sequences of CoqF homologs were downloaded from the UniProt database (https://uniprot.org/) for Bayesian molecular dating. Then, the amino acid alignments were produced using MAFFT L-INS-i v7.475. The best-fit model was found using ModelFinder embedded into IQ-TREE v2.1.2 (-m MF). Divergence times were estimated using the Bayesian evolutionary method integrated in BEAST2 packages (http://beast2.org/) by setting the following parameters: WAG with empirical frequencies (Gamma Category Count of 4 and proportion invariant of 0.11, according to SMS v1.8.4, www.atgc-montpellier.fr/sms/) as amino acid exchange rate matrices, a relaxed clock log normal as clock model, and a birth death model as tree priors.

In all cases, uniform prior distribution (min-max) was used to represent these ancient node constraints, and the root time was calibrated using a uniform distribution from 1.6 to 3.2 Ga. According to a recent molecular clock analysis (*58*), the following calibration points were applied: eudicotyledons (124 to 130 Ma), flowering plants (203 to 288 Ma), Embryophyta (470 to 540 Ma), Viridiplantae (1065 to 1524 Ma), Stramenopiles and Alveolate (1594 to 1901 Ma), and fungi (400- Ma). The posterior probabilities were estimated using Markov chain Monte Carlo (MCMC) with 50,000,000 chain length. The MCMC trace was summarized using Tracer v1.7.1 (http://beast.community/tracer) to assess the convergence and effective sample sizes of parameters. The maximum clade credibility tree was produced using TreeAnnotator with a burn in value set to 30% and visualized by FigTree v1.4.3 (http://tree.bio.ed.ac.uk/software/figtree/).

Statistical analysis
Significance was determined by an unpaired two-tailed Student's *t* test, *P <0.05, **P <0.01.

SUPPLEMENTARY MATERIALS
Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/7/50/eabl3594/DC1

View/request a protocol for this paper from *Bio-protocol*.

REFERENCES AND NOTES
1. Y. Wang, S. Hekimi, Understanding ubiquinone. *Trends Cell Biol.* **26**, 367–378 (2016).
2. J. Rahman, S. Rahman, Mitochondrial medicine in the omics era. *Lancet* **391**, 2560–2574 (2018).
3. K. Okada, K. Ohara, K. Yazaki, K. Nozaki, N. Uchida, M. Kawamukai, H. Nojiri, H. Yamane, The AtPPT1 gene encoding 4-hydroxybenzoate polyprenyl diphosphate transferase in ubiquinone biosynthesis is required for embryo development in *Arabidopsis thaliana*. *Plant Mol. Biol.* **55**, 567–577 (2004).
4. A. L. Ducluzeau, Y. Wamboldt, C. G. Elowsky, S. A. Mackenzie, R. C. Schuurink, G. J. Basset, Gene network reconstruction identifies the authentic *trans*-prenyl diphosphate synthase that makes the solanesyl moiety of ubiquinone-9 in *Arabidopsis*. *Plant J.* **69**, 366–375 (2012).
5. M. Ozeir, U. Mühlenhoff, H. Webert, R. Lill, M. Fontecave, F. Pierrel, Coenzyme Q biosynthesis: Coq6 is required for the C5-hydroxylation reaction and substrate analogs rescue Coq6 deficiency. *Chem. Biol.* **18**, 1134–1142 (2011).
6. W. W. Poon, R. J. Barkovich, A. Y. Hsu, A. Frankel, P. T. Lee, J. N. Shepherd, D. C. Myles, C. F. Clarke, Yeast and Rat Coq3 and *Escherichia coli* UbiG polypeptides catalyze both O-Methyltransferase steps in coenzyme Q biosynthesis. *J. Biol. Chem.* **274**, 21665–21672 (1999).
7. R. J. Barkovich, A. Shtanko, J. A. Shepherd, P. T. Lee, D. C. Myles, A. Tzagoloff, C. F. Clarke, Characterization of the COQ5 gene from *Saccharomyces cerevisiae* evidence for a c-methyltransferase in ubiquinone biosynthesis. *J. Biol. Chem.* **272**, 9182–9188 (1997).
8. B. N. Marbois, C. F. Clarke, The COQ7 gene encodes a protein in *Saccharomyces cerevisiae* necessary for ubiquinone biosynthesis. *J. Biol. Chem.* **271**, 2995–3004 (1996).
9. J. A. Stefely, D. J. Pagliarini, Biochemistry of mitochondrial coenzyme Q biosynthesis. *Trends Biochem. Sci.* **42**, 824–843 (2017).
10. T. Toda, K. Hayashi, Y. Ogiyama, K. Yokomi, T. Nakagawa, T. Kaino, M. Kawamukai, Functional conservation of coenzyme Q biosynthetic genes among yeasts, plants, and humans. *PLOS ONE* **9**, e99038 (2014).
11. J. M. Villalba, P. Navas, Regulation of coenzyme Q biosynthesis pathway in eukaryotes. *Free Radic. Biol. Med.* **165**, 312–323 (2021).
12. S. Padilla, U. C. Tran, M. Jiménez-Hidalgo, J. M. López-Martín, A. Martín-Montalvo, C. F. Clarke, P. Navas, C. Santos-Ocaña, Hydroxylation of demethoxy-Q6 constitutes a control point in yeast coenzyme Q6 biosynthesis. *Cell. Mol. Life Sci.* **66**, 173–186 (2009).
13. J. J. Ewbank, T. M. Barnes, B. Lakowski, M. Lussier, H. Bussey, S. Hekimi, Structural and functional conservation of the *Caenorhabditis elegans* Timing Geneclk-1. *Science* **275**, 980–983 (1997).
14. Z. Vajo, L. M. King, T. Jonassen, D. J. Wilkin, N. Ho, A. Munnich, C. F. Clarke, C. A. Francomano, Conservation of the *Caenorhabditis elegans* timing gene *clk-1* from yeast to human: A gene required for ubiquinone biosynthesis with potential implications for aging. *Mamm. Genomes* **10**, 1000–1004 (1999).
15. L. Pelosi, A.-L. Ducluzeau, L. Loiseau, F. Barras, D. Schneider, I. Junier, F. Pierrel, P. C. Dorrestein, Evolution of ubiquinone biosynthesis: Multiple proteobacterial enzymes with various regioselectivities to catalyze three contiguous aromatic hydroxylation reactions. *mSystems* **1**, e00091-16 (2016).
16. J. J. Almagro Armenteros, M. Salvatore, O. Emanuelsson, O. Winther, G. von Heijne, A. Elofsson, H. Nielsen, Detecting sequence signals in targeting peptides using deep learning. *Life Sci. Alliance* **2**, e201900429 (2019).
17. D. C. Lohman, D. Aydin, H. C. Von Bank, R. W. Smith, V. Linke, E. Weisenhorn, M. T. McDevitt, P. Hutchins, E. M. Wilkerson, B. Wancewicz, J. Russell, M. S. Stefely, E. T. Beebe, A. Jochem, J. J. Coon, C. A. Bingman, M. Dal Peraro, D. J. Pagliarini, An isoprene lipid-binding protein promotes eukaryotic coenzyme Q biosynthesis. *Mol. Cell* **73**, 763–774.e10 (2019).
18. C. M. Allan, A. M. Awad, J. S. Johnson, D. I. Shirasaki, C. Wang, C. E. Blaby-Haas, S. S. Merchant, J. A. Loo, C. F. Clarke, Identification of Coq11, a new coenzyme Q biosynthetic protein in the CoQ-synthome in *Saccharomyces cerevisiae*. *J. Biol. Chem.* **290**, 7517–7534 (2015).
19. M. H. Eppink, H. A. Schreuder, W. J. Van Berkel, Identification of a novel conserved sequence motif in flavoprotein hydroxylases with a putative dual function in FAD/NAD(P)H binding. *Protein Sci.* **6**, 2454–2458 (1997).
20. U. C. Tran, B. Marbois, P. Gin, M. Gulmezian, T. Jonassen, C. F. Clarke, Complementation of Saccharomyces cerevisiae coq7 mutants by mitochondrial targeting of the *Escherichia coli* UbiF polypeptide: Two functions of yeast Coq7 polypeptide in coenzyme Q biosynthesis. *J. Biol. Chem.* **281**, 16401–16409 (2006).
21. Y. Wang, S. Hekimi, The complexity of making ubiquinone. *Trends Endocrinol. Metab.* **30**, 929–943 (2019).

22. A. G. Reidenbach, Z. A. Kemmerer, D. Aydin, A. Jochem, M. T. McDevitt, P. D. Hutchins, J. L. Stark, J. A. Stefely, T. Reddy, A. S. Hebert, E. M. Wilkerson, I. E. Johnson, C. A. Bingman, J. L. Markley, J. J. Coon, M. Dal Peraro, D. J. Pagliarini, Conserved lipid and small-molecule modulation of COQ8 reveals regulation of the ancient kinase-like UbiB family. *Cell Chem. Biol.* **25**, 154–165.e11 (2018).

23. L. X. Xie, M. Ozeir, J. Y. Tang, J. Y. Chen, S.-K. Jaquinod, M. Fontecave, C. F. Clarke, F. Pierrel, Overexpression of the Coq8 Kinase in *Saccharomyces cerevisiae* coq null mutants allows for accumulation of diagnostic intermediates of the Coenzyme Q6 biosynthetic pathway. *J. Biol. Chem.* **287**, 23571–23581 (2012).

24. N. Geldner, V. Dénervaud-Tendon, D. L. Hyman, U. Mayer, Y. D. Stierhof, J. Chory, Rapid, combinatorial analysis of membrane compartments in intact plants with a multicolor marker set. *Plant J.* **59**, 169–178 (2009).

25. H. Ren, Q. Li, Y. Zhan, X. Fang, D. Yu, 2,4-Dichlorophenol hydroxylase for chlorophenol removal: Substrate specificity and catalytic activity. *Enzyme Microb. Technol.* **82**, 74–81 (2016).

26. W. A. Suske, M. Held, A. Schmid, T. Fleischmann, M. G. Wubbolts, H. P. Kohler, Purification and characterization of 2-hydroxybiphenyl 3-monooxygenase, a novel NADH-dependent, FAD-containing aromatic hydroxylase from *Pseudomonas azelaica* HBP1. *J. Biol. Chem.* **272**, 24257–24265 (1997).

27. M. Kanteev, A. Bregman-Cohen, B. Deri, A. Shahar, N. Adir, A. Fishman, A crystal structure of 2-hydroxybiphenyl 3-monooxygenase with bound substrate provides insights into the enzymatic mechanism. *Biochim. Biophys. Acta* **1854**, 1906–1913 (2015).

28. D. J. Richter, C. Berney, J. F. H. Strassert, F. Burki, C. de Vargas, EukProt: A database of genome-scale predicted proteins across the diversity of eukaryotic life. *bioRxiv* 2020.2006.2030.180687 (2020).

29. S. M. Adl, D. Bass, C. E. Lane, J. Lukeš, C. L. Schoch, A. Smirnov, S. Agatha, C. Berney, M. W. Brown, F. Burki, P. Cárdenas, I. Čepička, L. Chistyakova, J. Del Campo, M. Dunthorn, B. Edvardsen, Y. Eglit, L. Guillou, V. Hampl, A. A. Heiss, M. Hoppenrath, T. Y. James, A. Karnkowska, S. Karpov, E. Kim, M. Kolisko, A. Kudryavtsev, D. J. G. Lahr, E. Lara, L. Le Gall, D. H. Lynn, D. G. Mann, R. Massana, E. A. D. Mitchell, C. Morrow, J. S. Park, J. W. Pawlowski, M. J. Powell, D. J. Richter, S. Rueckert, L. Shadwick, S. Shimano, F. W. Spiegel, G. Torruella, N. Youssef, V. Zlatogursky, Q. Zhang, Revisions to the classification, nomenclature, and diversity of eukaryotes. *J. Eukaryot. Microbiol.* **66**, 4–119 (2019).

30. B. A. Curtis, G. Tanifuji, F. Burki, A. Gruber, M. Irimia, S. Maruyama, M. C. Arias, S. G. Ball, G. H. Gile, Y. Hirakawa, J. F. Hopkins, A. Kuo, S. A. Rensing, J. Schmutz, A. Symeonidi, M. Elias, R. J. Eveleigh, E. K. Herman, M. J. Klute, T. Nakayama, M. Oborník, A. Reyes-Prieto, E. V. Armbrust, S. J. Aves, R. G. Beiko, P. Coutinho, J. B. Dacks, D. G. Durnford, N. M. Fast, B. R. Green, C. J. Grisdale, F. Hempel, B. Henrissat, M. P. Höppner, K. Ishida, E. Kim, L. Kořený, P. G. Kroth, Y. Liu, S. B. Malik, U. G. Maier, D. McRose, T. Mock, J. A. Neilson, N. T. Onodera, A. M. Poole, E. J. Pritham, T. A. Richards, G. Rocap, S. W. Roy, C. Sarai, S. Schaack, S. Shirato, C. H. Slamovits, D. F. Spencer, S. Suzuki, A. Z. Worden, S. Zauner, K. Barry, C. Bell, A. K. Bharti, J. A. Crow, J. Grimwood, R. Kramer, E. Lindquist, S. Lucas, A. Salamov, G. I. McFadden, C. E. Lane, P. J. Keeling, M. W. Gray, I. V. Grigoriev, J. M. Archibald, Algal genomes reveal evolutionary mosaicism and the fate of nucleomorphs. *Nature* **492**, 59–65 (2012).

31. G. I. McFadden, E. Yeh, The apicoplast: Now you see it, now you don't. *Int. J. Parasitol.* **47**, 137–144 (2017).

32. M. L. Povelones, Beyond replication: Division and segregation of mitochondrial DNA in kinetoplastids. *Mol. Biochem. Parasitol.* **196**, 53–60 (2014).

33. One Thousand Plant Transcriptomes Initiative, One thousand plant transcriptomes and the phylogenomics of green plants. *Nature* **574**, 679–685 (2019).

34. A. Del Cortona, C. J. Jackson, F. Bucchini, M. Van Bel, S. D'Hondt, P. Škaloud, C. F. Delwiche, A. H. Knoll, J. A. Raven, H. Verbruggen, K. Vandepoele, O. De Clerck, F. Leliaert, Neoproterozoic origin and multiple transitions to macroscopic growth in green seaweeds. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 2551–2559 (2020).

35. L. Li, S. Wang, H. Wang, S. K. Sahu, B. Marin, H. Li, Y. Xu, H. Liang, Z. Li, S. Cheng, T. Reder, Z. Cebi, S. Wittek, M. Petersen, B. Melkonian, H. Du, H. Yang, J. Wang, G. K. Wong, X. Xu, X. Liu, Y. Van de Peer, M. Melkonian, H. Liu, The genome of *Prasinoderma coloniale* unveils the existence of a third phylum within green plants. *Nat. Ecol. Evol.* **4**, 1220–1231 (2020).

36. S. Cheng, W. Xian, Y. Fu, B. Marin, J. Keller, T. Wu, W. Sun, X. Li, Y. Xu, Y. Zhang, S. Wittek, T. Reder, G. Gunther, A. Gontcharov, S. Wang, L. Li, X. Liu, J. Wang, H. Yang, X. Xu, P. M. Delaux, B. Melkonian, G. K. Wong, M. Melkonian, Genomes of subaerial zygnematophyceae provide insights into land plant evolution. *Cell* **179**, 1057–1067.e14 (2019).

37. A. J. Roger, S. A. Muñoz-Gómez, R. Kamikawa, The origin and diversification of mitochondria. *Curr. Biol.* **27**, R1177–R1192 (2017).

38. A. D. Anbar, Oceans. Elements and evolution. *Science* **322**, 1481–1483 (2008).

39. A. F. Lodeyro, R. D. Ceccoli, J. J. Pierella Karlusich, N. Carrillo, The importance of flavodoxin for environmental stress tolerance in photosynthetic microorganisms and transgenic plants. Mechanism, evolution and biotechnological potential. *FEBS Lett.* **586**, 2917–2924 (2012).

40. B. Nowicka, A. Trela-Makowej, D. Latowski, K. Strzalka, R. Szymańska, Antioxidant and signaling role of plastid-derived isoprenoid quinones and chromanols. *Int. J. Mol. Sci.* **22**, 2950 (2021).

41. G. Wheeler, T. Ishikawa, V. Pornsaksit, N. Smirnoff, Evolution of alternative biosynthetic pathways for vitamin C following plastid acquisition in photosynthetic eukaryotes. *eLife* **4**, e06369 (2015).

42. W. L. Araújo, A. Nunes-Nesi, S. Osorio, B. Usadel, D. Fuentes, R. Nagy, I. Balbo, M. Lehmann, C. Studart-Witkowski, T. Tohge, E. Martinoia, X. Jordana, F. M. Damatta, A. R. Fernie, Antisense inhibition of the iron-sulphur subunit of succinate dehydrogenase enhances photosynthesis and growth in tomato via an organic acid-mediated effect on stomatal aperture. *Plant Cell* **23**, 600–627 (2011).

43. D. Fuentes, M. Meneses, A. Nunes-Nesi, W. L. Araújo, R. Tapia, I. Gómez, L. Holuigue, R. A. Gutiérrez, A. R. Fernie, X. Jordana, A deficiency in the flavoprotein of Arabidopsis mitochondrial complex II results in elevated photosynthesis and better growth in nitrogen-limiting conditions. *Plant Physiol.* **157**, 1114–1127 (2011).

44. E. Soubeyrand, M. Kelly, S. A. Keene, A. C. Bernert, S. Latimer, T. S. Johnson, C. Elowsky, T. A. Colquhoun, A. K. Block, G. J. Basset, Arabidopsis 4-COUMAROYL-COA LIGASE 8 contributes to the biosynthesis of the benzenoid ring of coenzyme Q in peroxisomes. *Biochem. J.* **476**, 3521–3532 (2019).

45. I. B. Verdaguer, M. Crispim, C. A. Zafra, R. A. C. Sussmann, N. L. Buritica, H. R. Melo, M. F. Azevedo, F. G. Almeida, E. A. Kimura, A. M. Katzin, Exploring ubiquinone biosynthesis inhibition as a strategy for improving atovaquone efficacy in malaria. *Antimicrob. Agents Chemother.* **65**, e01516-20 (2021).

46. Y. Jiang, B. Chen, C. Duan, B. Sun, J. Yang, S. Yang, R. M. Kelly, Multigene editing in the *Escherichia coli* genome via the CRISPR-Cas9 system. *Appl. Environ. Microbiol.* **81**, 2506–2514 (2015).

47. S. J. Clough, A. F. Bent, Floral dip: A simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. *Plant J.* **16**, 735–743 (1998).

48. B. K. Nelson, X. Cai, A. Nebenführ, A multicolored set of in vivo organelle markers for co-localization studies in *Arabidopsis* and other plants. *Plant J.* **51**, 1126–1136 (2007).

49. R. P. Hellens, E. A. Edwards, N. R. Leyland, S. Bean, P. M. Mullineaux, pGreen: A versatile and flexible binary Ti vector for *Agrobacterium*-mediated plant transformation. *Plant Mol. Biol.* **42**, 819–832 (2000).

50. T. Czechowski, M. Stitt, T. Altmann, M. K. Udvardi, W. R. Scheible, Genome-wide identification and testing of superior reference genes for transcript normalization in *Arabidopsis*. *Plant Physiol.* **139**, 5–17 (2005).

51. Q. He, J. Peng, F. Yan, L. Lin, Y. Lu, H. Zheng, H. Chen, J. Chen, Intron retention and 3'-UTR analysis of Arabidopsis Dicer-like 2 transcripts. *Mol. Biol. Rep.* **39**, 3271–3280 (2012).

52. M. Padmanabhan, S. P. Dinesh-Kumar, Virus-induced gene silencing as a tool for delivery of dsRNA into plants. *Cold Spring Harb. Protoc.* **2009**, pdb.prot5139 (2009).

53. K. Katoh, D. M. Standley, MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).

54. D. T. Hoang, O. Chernomor, A. von Haeseler, B. Q. Minh, L. S. Vinh, UFBoot2: Improving the ultrafast bootstrap approximation. *Mol. Biol. Evol.* **35**, 518–522 (2018).

55. S. R. Eddy, Accelerated profile HMM searches. *PLoS Comput. Biol.* **7**, e1002195 (2011).

56. S. Capella-Gutiérrez, J. M. Silla-Martínez, T. Gabaldón, trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972–1973 (2009).

57. G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell, A. J. Olson, AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* **30**, 2785–2791 (2009).

58. J. F. H. Strassert, I. Irisarri, T. A. Williams, F. Burki, A molecular timescale for eukaryote evolution with implications for the origin of red algal-derived plastids. *Nat. Commun.* **12**, 1879 (2021).

59. F. Burki, A. J. Roger, M. W. Brown, A. G. B. Simpson, The new tree of eukaryotes. *Trends Ecol. Evol.* **35**, 43–55 (2020).

60. F. Leliaert, D. R. Smith, H. Moreau, M. D. Herron, H. Verbruggen, C. F. Delwiche, O. De Clerck, Phylogeny and molecular evolution of the green Algae. *Crit. Rev. Plant Sci.* **31**, 1–46 (2012).