

RESEARCH PAPER

# Mitochondrial and plastidial COG0354 proteins have folate-dependent functions in iron–sulphur cluster metabolism

Jeffrey C. Waller<sup>1</sup>, Kenneth W. Ellens<sup>1</sup>, Sophie Alvarez<sup>2</sup>, Karen Loizeau<sup>3</sup>, Stéphane Ravel<sup>3</sup> and Andrew D. Hanson<sup>1,\*</sup>

<sup>1</sup> Horticultural Sciences Department, University of Florida, Gainesville, FL 32611, USA

<sup>2</sup> Donald Danforth Plant Science Center, St Louis, MO 63132, USA

<sup>3</sup> Laboratoire de Physiologie Cellulaire Végétale, CNRS/CEA/INRA/Université Joseph Fourier, CEA-Grenoble, F-38054 Grenoble cedex 9, France

\* To whom correspondence should be addressed. E-mail: [adha@ufl.edu](mailto:adha@ufl.edu)

Received 21 June 2011; Revised 4 August 2011; Accepted 12 August 2011

## Abstract

COG0354 proteins have been implicated in synthesis or repair of iron/sulfur (Fe/S) clusters in all domains of life, and those of bacteria, animals, and protists have been shown to require a tetrahydrofolate to function. Two COG0354 proteins were identified in *Arabidopsis* and many other plants, one (At4g12130) related to those of  $\alpha$ -proteobacteria and predicted to be mitochondrial, the other (At1g60990) related to those of cyanobacteria and predicted to be plastidial. Grasses and poplar appear to lack the latter. The predicted subcellular locations of the *Arabidopsis* proteins were validated by *in vitro* import assays with purified pea organelles and by targeting assays in *Arabidopsis* and tobacco protoplasts using green fluorescent protein fusions. The At4g12130 protein was shown to be expressed mainly in flowers, siliques, and seeds, whereas the At1g60990 protein was expressed mainly in young leaves. The folate dependence of both *Arabidopsis* proteins was established by functional complementation of an *Escherichia coli* COG0354 (*ygfZ*) deletant; both plant genes restored *in vivo* activity of the Fe/S enzyme MiaB but restoration was abrogated when folates were eliminated by deleting *folP*. Insertional inactivation of At4g12130 was embryo lethal; this phenotype was reversed by genetic complementation of the mutant. These data establish that COG0354 proteins have a folate-dependent function in mitochondria and plastids, and that the mitochondrial protein is essential. That plants retain mitochondrial and plastidial COG0354 proteins with distinct phylogenetic origins emphasizes how deeply the extant Fe/S cluster assembly machinery still reflects the ancient endosymbioses that gave rise to plants.

**Key words:** COG0354, folate, iron–sulfur cluster, mitochondrion, plastid.

## Introduction

Iron/sulfur (Fe/S) clusters are highly versatile but labile cofactors of ancient origin that occur in all domains of life (Beinert, 2000). Fe/S cluster proteins play central roles in electron transfer, catalysis, and gene expression (Johnson *et al.*, 2005). Plants have over a hundred such proteins, located in mitochondria, plastids, cytosol, and nuclei (Balk and Pilon, 2011).

Although Fe/S clusters are simple structures, their biogenesis and maintenance require a complex machinery. The grand lines of this machinery are now clear and are similar

in bacteria and eukaryotes (Johnson *et al.*, 2005; Lill, 2009). Thus, Fe/S cluster assembly begins with a cysteine desulfurase that transfers sulfur from free cysteine to a scaffold protein, which also receives Fe from an Fe-binding protein. The scaffold protein acts as a platform on which clusters are assembled and from which they are mobilized to target apoproteins, often with the help of chaperones. Fe/S cluster biogenesis systems are of several types. Bacteria such as *Escherichia coli* have two independent systems, ISC and SUF. Plants also have ISC and SUF systems, ISC being in

mitochondria and SUF in plastids; these organelles thus have autonomous Fe/S assembly machinery that reflects their endosymbiotic origins (Xu and Møller, 2008; Balk and Pilon, 2011). Plants have in addition a cytosolic system, CIA, which depends in part on the mitochondrial ISC system (Balk and Pilon, 2011). Other eukaryotes have mitochondrial ISC and cytosolic CIA systems (Lill, 2009).

Despite the progress in dissecting Fe/S cluster biogenesis, much remains unclear, including the roles of proteins involved in the maturation of specific subsets of Fe/S enzymes (Sheftel *et al.*, 2010). One such protein is COG0354, which occurs in all domains of life and is known to be mitochondrial in yeast and animals (Gelling *et al.*, 2008; Pagliarini *et al.*, 2008). Yeast COG0354 (Iba57p) is required for Fe/S cluster formation on aconitase and activation of two radical SAM (*S*-adenosylmethionine) Fe/S enzymes (Gelling *et al.*, 2008). *E. coli* COG0354 (YgfZ) supports succinate dehydrogenase, fumarase, dimethylsulphoxide reductase, and the radical SAM enzyme MiaB, particularly during oxidative stress (Ote *et al.*, 2006; Waller *et al.*, 2010). These functions are crucial at the organism level: deleting *iba57* causes loss of mitochondrial function and a petite phenotype (Gelling *et al.*, 2008); deleting *ygfZ* impairs growth and increases oxidative stress sensitivity (Lin *et al.*, 2010; Waller *et al.*, 2010); and knockdown of zebrafish COG0354 (*C1orf69*) causes anaemia (Nilsson *et al.*, 2009).

The biochemical action of COG0354 is not yet understood, but it appears to be universally conserved because the  $\Delta ygfZ$  growth defects are complemented by COG0354 genes from all domains of life, including plants (Waller *et al.*, 2010), and the  $\Delta iba57$  growth defects are complemented by human *C1orf69* (Gelling *et al.*, 2008). Major clues to the action of COG0354 have been the prediction of a folate-binding site on the *E. coli* protein (Tepljakov *et al.*, 2004), the finding that this protein indeed binds a model folate (5-formyltetrahydrofolate) *in vitro* (Waller *et al.*, 2010), and the genetic demonstration that *E. coli*, mammalian, and protistan COG0354 proteins depend on a tetrahydrofolate to function *in vivo* (Waller *et al.*, 2010).

Nothing was known until now about plant COG0354 proteins except that there are two types, and that both types from *Arabidopsis* complement the *E. coli*  $\Delta ygfZ$  mutation (Waller *et al.*, 2010). Here, it is shown that one plant COG0354 protein is mitochondrial and that the other is plastidial, that they are respectively predicted to derive from  $\alpha$ -proteobacteria and cyanobacteria, that they both require folate for *in vivo* activity, and that the mitochondrial protein is essential for embryo development.

## Materials and methods

### Bioinformatics

COG0354 protein sequences (identified by the KGC-[F/Y]-x-GQE-x(3)-[K/R] motif) were taken from NCBI (<http://www.ncbi.nlm.nih.gov/>) and SEED (<http://www.theseed.org/>) databases. Plant genomes and expressed sequence tags (ESTs) were searched for COG0354 sequences using BLAST algorithms at the NCBI and Joint Genome Institute (<http://genome.jgi-psf.org/>) databases.

Sequences were aligned using Multalin (<http://multalin.toulouse.inra.fr/multalin/>) or ClustalW; phylogenetic analyses were carried out with MEGA version 4.0 (Tamura *et al.*, 2007). Organellar targeting was predicted using TargetP (<http://www.cbs.dtu.dk/services/TargetP/>), WoLF PSORT (<http://wolfpsort.org/>), and Predotar (<http://urgi.versailles.inra.fr/predotar/predotar.html>). Ambiguous (mitochondrial/plastidial) targeting was evaluated using the COSMOSS Ambiguous Targeting Predictor (<http://www.cosmoss.org/bm/ATP>).

### Plant materials and growth conditions

*Arabidopsis thaliana* (ecotype Col-0) seeds were sterilized, imbibed at 4 °C for 5 d, and germinated on 0.5× MS salts containing 0.5% (w/v) sucrose. Seedlings were transferred after 14 d either into potting soil and supplemented biweekly with Peters Professional 20:20:20 fertilizer, or (for protein expression analysis) to hydroponic culture (Gibeaut *et al.*, 1997). Photosynthetic photon flux density was 150  $\mu\text{mol quanta m}^{-2} \text{s}^{-1}$  with a 16 h photoperiod (22 °C day, 18 °C night). Seeds of pea (*Pisum sativum* cv. Laxton's Improved Progress 9) were imbibed overnight in running tap water, sown in moist vermiculite, and grown at 16–18 °C under fluorescent light (75  $\mu\text{mol quanta m}^{-2} \text{s}^{-1}$ ) with a 12 h photoperiod for 11 d.

### In vitro organellar import assays

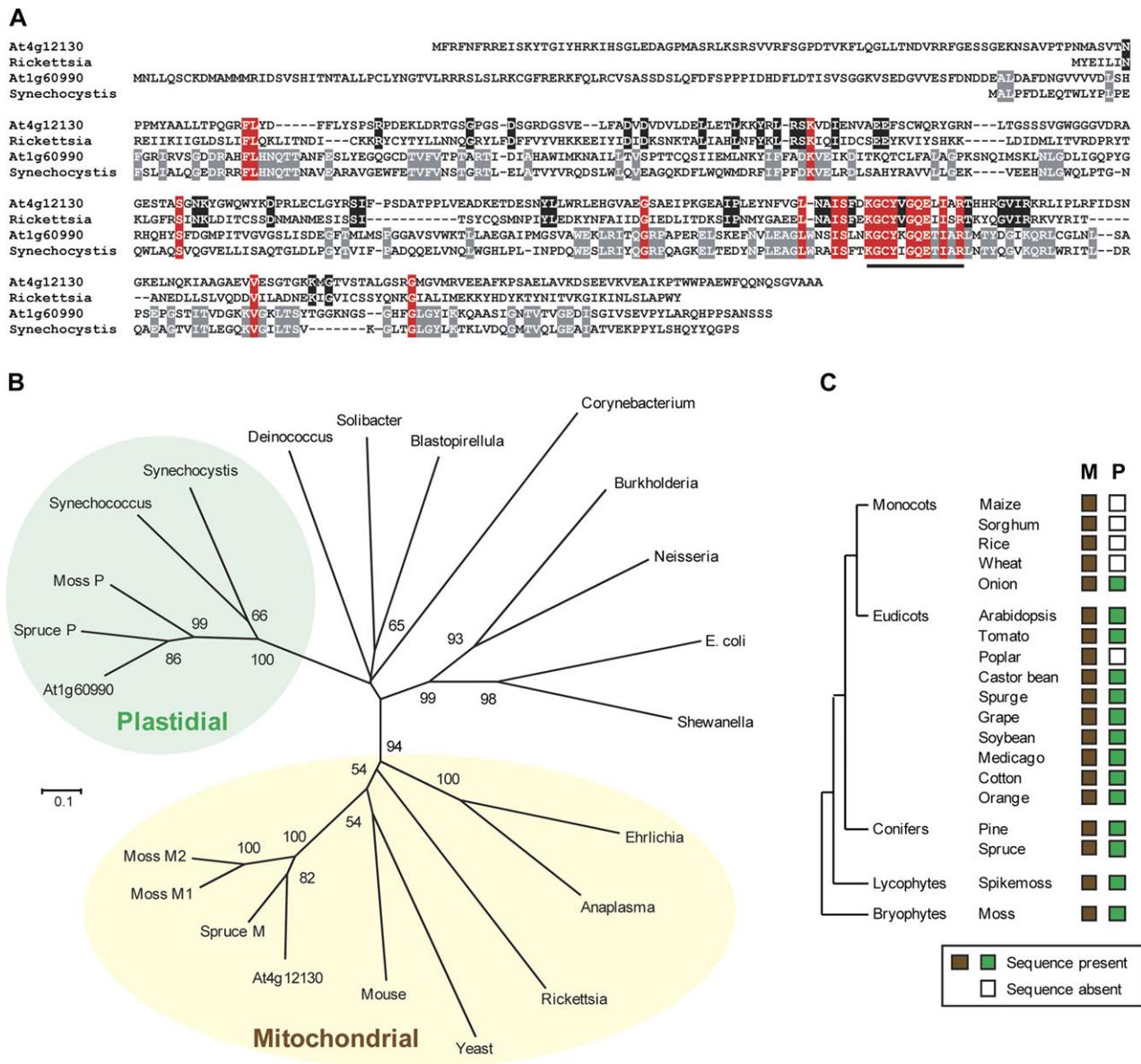
For dual import assays (Rudhe *et al.*, 2002), cDNAs encoding full-length At4g12130 and At1g60990 were PCR amplified using primers At4g12130-pGEM-fwd and -rev, and At1g60990-pGEM-fwd and -rev, respectively (Supplementary Table S1 available at *JXB* online); forward primers included a Kozak sequence. Amplicons were cloned into pGEM-4Z (Promega). Coupled *in vitro* transcription–translation, isolation of pea chloroplasts and mitochondria, dual import assays, thermolysin treatment, and repurification of organelles were performed as described (Pribat *et al.*, 2010). Soybean alternative oxidase (Rudhe *et al.*, 2002) and pine phenylalanine hydroxylase (Pribat *et al.*, 2010) served as positive controls for mitochondrial and chloroplast targeting, respectively.

### Transient expression of GFP fusion proteins

The N-terminal 126 residues of At4g12130 or 122 residues At1g60990 were fused, via a tetraglycine linker, upstream and in-frame with green fluorescent protein (GFP) as follows. The first 378 bp of *At4g12130* or 366 bp of *At1g60990* were amplified from cDNAs (ABRC stock no. U17502 and Riken stock no. RAFL21-91-J21, respectively) using primers At4g12130-GFP-fwd and -rev, and At1g60990-GFP-fwd and -rev, respectively (Supplementary Table S1 at *JXB* online). Amplicons were digested with appropriate restriction enzymes and ligated into pTH2 (Niwa, 2003). Protoplasts were prepared from suspension culture cells of *Arabidopsis* or tobacco (*Nicotiana tabacum*) cv. BY-2, transformed, and stained with MitoTracker Orange (Molecular Probes) as described (Ravanel *et al.*, 2001). Controls for GFP targeting and epifluorescence microscopy were as described (Pinon *et al.*, 2005). BY-2 protoplasts lack chlorophyll, which enables co-localization by epifluorescence microscopy of signals from GFP and the MitoTracker mitochondrial marker. Mitochondrial localization was therefore tested in BY-2 cells, and chloroplastic localization in *Arabidopsis* cells.

### Protein expression analyses

Roots and stems were harvested from 6-week-old hydroponically cultured plants; other organs (except seedlings) were from soil-grown plants. Tissues were ground in liquid N<sub>2</sub>. For immunoblot analysis, 100 mg tissue samples were extracted in 0.1 M TRIS-HCl (pH 7.5), 0.2 M KCl, 3 mM MgCl<sub>2</sub>, 1 mM EDTA, 1 mM EGTA, 0.05% (v/v) Nonidet P-40, 0.1% (w/v) polyvinylpyrrolidone, 2% (w/v) polyvinylpolypyrrolidone, 1 mM dithiothreitol, 1 mM benzamidine, 1 mM phenylmethylsulphonyl fluoride, 5 mM  $\epsilon$ -aminocaproic acid, and centrifuged to clear. Protein was estimated by dye binding (Bradford, 1976) with bovine serum albumin as standard.



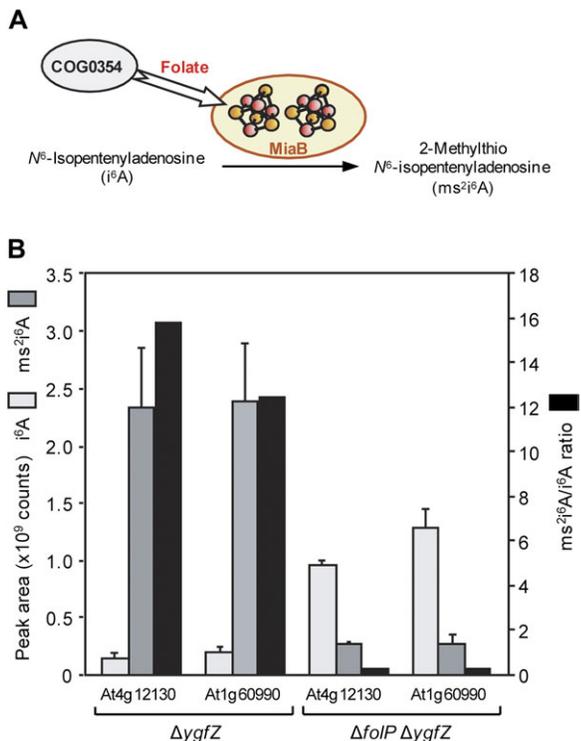
**Fig. 1.** Primary structure, phylogeny, and distribution of plant COG0354 proteins. (A) Amino acid sequence alignment of At4g12130, At1g60990, and a close bacterial homologue of each (*Rickettsia prowazekii* and *Synechocystis* sp. PCC 6803, respectively). Residues conserved in all four sequences are in red; those conserved only in At4g12130 and *Rickettsia* are in black, and those conserved only in At1g60990 and *Synechocystis* are in grey. Dashes are gaps to maximize alignment. The COG0354 signature motif (KGC-[YF]-x-GQE-x(3)-[KR]) is red underlined; note that this motif is slightly longer than that defined by Teplyakov *et al.* (2004). (B) Neighbor-joining phylogenetic tree of COG0354 proteins from diverse plants and bacteria; yeast lba57p and mouse C1orf69 are also included. Bootstrap values (1000 replicates) are next to branches; values <50% are not shown. Branch length indicates the number of inferred amino acid changes per position. At4g12130, other plant proteins predicted to be mitochondrial (M), and yeast and mouse sequences fall into the same clade ('Mitochondrial') as *Rickettsia* and other Rickettsiales (*Anaplasma*, *Ehrlichia*). At1g60990 and other plant proteins predicted to be plastidial (P) fall into the same clade ('Plastidial') as cyanobacteria. (C) Occurrence of orthologues of At4g12130 (M) and At1g60990 (P) in genomes and/or EST collections of land plants. Full species names: maize, *Zea mays*; sorghum, *Sorghum bicolor*; rice, *Oryza sativa*; wheat, *Triticum aestivum*; onion, *Allium cepa*; tomato, *Solanum lycopersicum*; poplar, *Populus trichocarpa*; castor bean, *Ricinus communis*; spurge, *Euphorbia esula*; grape, *Vitis vinifera*; soybean, *Glycine max*; medicago, *Medicago truncatula*; cotton, *Gossypium hirsutum*; orange, *Citrus sinensis*; pine, *Pinus taeda*; spruce, *Picea glauca*; spikemoss, *Selaginella moellendorffii*; moss, *Physcomitrella patens*.

Denaturing gel electrophoresis and immunoblotting were as described (Turner *et al.*, 2005); antisera were diluted 1:1000. Antisera were raised in rabbits (Cocalico Biologicals Inc.) against hexahistidine-tagged, denatured At4g12130 or At1g60990 prepared as described (Waller *et al.*, 2010).

*Folate dependence tests*

The  $\Delta ygfZ$  and  $\Delta folP::kan^r \Delta ygfZ$  *E. coli* strains (K12 MG1655 background) were as described (Waller *et al.*, 2010). Both harboured pACYC-RP (Stratagene), which carries genes for rare arginine and proline tRNAs. Each of these strains was transformed with pBAD24 containing the truncated *At4g12130* or





**Fig. 3.** Evidence that At4g12130 and At1g60990 require folate for activity *in vivo*. (A) The reaction catalysed by the tRNA modification enzyme MiaB, and the relationship between MiaB, the COG0354 protein, and folate. MiaB has two [4Fe–4S] clusters (Hernández *et al.*, 2007), which are depicted schematically. (B) LC-MS/MS quantification of i<sup>6</sup>A and ms<sup>2</sup>i<sup>6</sup>A, and the ms<sup>2</sup>i<sup>6</sup>A/i<sup>6</sup>A ratio, in tRNA of the ΔygfZ or ΔfolP ΔygfZ strains expressing At4g12130 or At1g60990, grown in Antibiotic Medium 3 plus 300 μM thymidine, 0.02% L-arabinose, and appropriate antibiotics. Data are means and SE for three independent samples.

its plant orthologues fall into the same clade as *Rickettsia* and other α-proteobacteria, which are the closest living relatives of the mitochondrial progenitor (Williams *et al.*, 2007). This clade also contained the yeast and mouse COG0354 proteins, which have both been shown to be mitochondrial (Gelling *et al.*, 2008; Pagliarini *et al.*, 2008) (Fig. 1B). Conversely, At1g60990 and its counterparts in other plants are in the same clade as proteins from cyanobacteria, the group that gave rise to plastids (Gould *et al.*, 2008) (Fig. 1B). These clades are referred to here as ‘mitochondrial’ and ‘plastidial’, respectively (Fig. 1B). Alignment of the sequences of the *Arabidopsis* proteins with their closest bacterial homologues (Fig. 1A) underscores the marked divergence between mitochondrial- and plastidial-type proteins. While there is substantial overall sequence identity within each type, there is little identity between the types outside the COG0354 signature motif KGC-[YF]-x-GQE-x(3)-R.

Although many plant genomes encode both COG0354 proteins, some seem to lack the plastidial type (Fig. 1C). Those without this type include poplar, and all grasses for which genome sequences and/or deep EST collections exist.

EST evidence indicates that a plastidial-type protein occurs in other monocots, for example onion (Fig. 1C).

### Subcellular localization of Arabidopsis COG0354 proteins

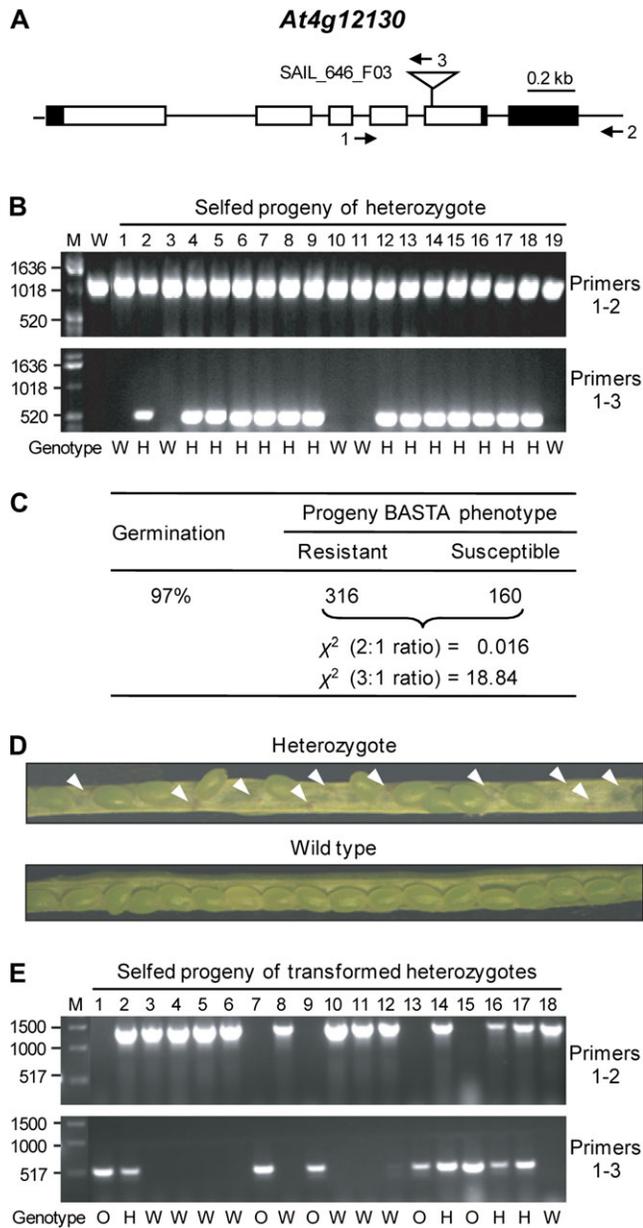
The subcellular location of the two *Arabidopsis* COG0354 proteins was investigated first using dual import assays, in which radiolabelled full-length translation products are incubated with a mixture of purified chloroplasts and mitochondria (Rudhe *et al.*, 2002). In reactions with At4g12130, mitochondria contained a labelled product that was smaller than the full-length precursor and resistant to attack by thermolysin, as expected for a translocated protein, but no thermolysin-resistant protein was present in chloroplasts (Fig. 2A). Conversely, in reactions with At1g60990, chloroplasts contained a smaller, thermolysin-resistant product but mitochondria did not (Fig. 2A). To complement this approach, subcellular localization was analysed in tobacco BY-2 protoplasts and *Arabidopsis* protoplasts using fusions between the N-terminus of each COG0354 protein and GFP. Expression of the At4g12130::GFP fusion in BY-2 cells resulted in a punctate pattern of fluorescence matching that of the MitoTracker mitochondrial marker (Fig. 2B). In contrast, the At1g60990::GFP fusion expressed in *Arabidopsis* cells gave larger punctate fluorescent areas that coincided with chlorophyll autofluorescence (Fig. 2B). Thus both *in vitro* and *in vivo* approaches indicate that, as predicted, the At4g12130 protein is targeted to mitochondria and the At1g60990 protein is targeted to chloroplasts. High throughput proteomic data also support a chloroplastic location for At1g60990 (Olinares *et al.*, 2010).

### Expression patterns of At4g12130 and At1g60990

At4g12130 and At1g60990 protein levels were analysed by immunoblotting using rabbit antisera (Fig. 2C). At4g12130 protein was most abundant in flowers, siliques, and dry seeds, with cleavage products detectable in the latter samples. At1g60990 protein was most abundant in young leaves. These results agree broadly with transcriptome data, which show that *At4g12130* mRNA is most strongly expressed in flowers, pollen, and seeds, whereas *At1g60990* mRNA is expressed mainly in green tissues (Steinhauser *et al.*, 2004; Xiang *et al.*, 2011).

### Evidence for folate dependence

The folate requirement of COG0354 can be established using a heterologous *E. coli* system in which the Fe/S enzyme MiaB serves as an *in vivo* reporter of COG0354 activity (Waller *et al.*, 2010). MiaB mediates the conversion of N<sup>6</sup>-isopentenyladenosine (i<sup>6</sup>A) in tRNA to 2-methylthio-N<sup>6</sup>-isopentenyladenosine (ms<sup>2</sup>i<sup>6</sup>A); its *in vivo* activity can thus be gauged from the ms<sup>2</sup>i<sup>6</sup>A/i<sup>6</sup>A ratio in tRNA. As *E. coli* MiaB activity depends strongly on the action of COG0354, the ms<sup>2</sup>i<sup>6</sup>A/i<sup>6</sup>A ratio is a sensitive measure of COG0354 function (Fig. 3A). This ratio is >10 in wild-type *E. coli* cells, but can



**Fig. 4.** Evidence that *At4g12130* is essential for embryo development. (A) Scheme showing the position of the T-DNA insertion in the *At4g12130* gene. Introns are represented by lines and exons by boxes; 5' - and 3' -untranslated regions are in black. The positions of primers used for genomic PCR are shown by numbered arrows. Key to primers: 1, *At4g12130*-GS-fwd; 2, *At4g12130*-GS-rev; 3, TDNA-LB. (B) PCR analysis of genomic DNA of 19 representative selfed progeny of heterozygotes, showing the absence of mutant homozygotes; a wild-type control (W) is included. Lane M shows molecular size markers (bp). The progeny genotypes [i.e. wild type (W) or heterozygous (H)] are indicated beneath the gel images. (C) Segregation of BASTA resistance in the selfed progeny of heterozygous plants. The hypothesis of a 2:1 ratio (resistant:susceptible) is accepted by the  $\chi^2$  test; a 3:1 ratio is strongly rejected ( $P=0.001$ ). (D) Dissection of representative siliques from self-pollinated heterozygous and wild-type plants. Arrowheads mark aborted seeds. (E) PCR analysis of genomic DNA of 18 representative selfed progeny of heterozy-

drop to  $<1$  when the gene encoding COG0354 (*ygfZ*) is deleted (Ote et al., 2006; Waller et al., 2010).

To test folate dependence of the *Arabidopsis* COG0354 proteins, each was expressed from a plasmid in an *E. coli* *ygfZ* deletant strain ( $\Delta ygfZ$ ). As expected from the observation that *At4g12130* and *At1g60990* both complement the growth phenotypes of this strain (Waller et al., 2010), they both gave high  $ms^2i^6A/i^6A$  ratios in the wild-type range (Fig. 3B). However, when the *folP* gene (encoding the folate synthesis enzyme dihydropteroate synthase) was also deleted, the ratios were  $<1$  (Fig. 3B). Since  $\Delta folP$  strains completely lack folates (Waller et al., 2010), this result indicates that the action of both *Arabidopsis* COG0354 proteins requires folate, as has been shown for animal and protist COG0354 proteins (Waller et al., 2010).

#### An insertional mutant of *At4g12130* and its complementation

A T-DNA mutant carrying a single insertion in exon 5 of *At4g12130* was identified in the GABI-Kat collection (Fig. 4A); the insertion interrupts the *At4g12130* open reading frame 69 residues upstream of the C-terminus. The selfed progeny of heterozygotes contained only wild-type and heterozygous individuals (Fig. 4B), and segregated for the BASTA resistance marker in a 2:1 (resistant:susceptible) ratio (Fig. 4C). These observations both point to embryo lethality. Consistent with embryo lethality, immature siliques of heterozygous plants contained aborted seeds at about one-quarter of the positions; such defective seeds were absent from wild-type siliques (Fig. 4D). To confirm that the lethal phenotype is due to ablation of *At4g12130*, heterozygous mutants were transformed with a genomic clone of *At4g12130* that included 2000 bp of upstream promoter sequence. The selfed progeny of transformants included individuals homozygous for the mutant *At4g12130* locus, as determined by PCR (Fig. 4E). Of 68 progeny screened, eight were complemented homozygotes, 38 were wild type, and 22 were heterozygotes. The complemented plants were normal in appearance. These complementation data validate the conclusion that *At4g12130* is essential for embryo development.

## Discussion

Bioinformatic analysis and targeting studies showed that plants, as a group, have two COG0354 proteins with distinct subcellular locations and origins. One is mitochondrial, is related to the mitochondrial COG0354 proteins of other eukaryotes, and is most probably derived from the  $\alpha$ -proteobacterial progenitor of mitochondria. The other

gotes transformed with a genomic clone of *At4g12130* that included the native promoter. Note the presence of mutant homozygotes (O) as well as heterozygous (H) and wild-type (W) individuals.

COG0354 protein is unique to plants, is located in plastids, and probably originated from cyanobacteria. This dual distribution is consistent with the presence of independent Fe/S cluster assembly machinery in mitochondria and plastids, and underscores how these types of machinery mirror the endosymbiotic history of the plant cell (Balk and Pilon, 2011). Also, the presence of COG0354 proteins in both organelles in which Fe/S clusters are made is consistent with the evidence that COG0354 family proteins interact physically with other Fe/S assembly components (Gelling *et al.*, 2008; Hu *et al.*, 2009), namely that they form part of the assembly complex.

It is at first sight surprising that certain plant taxa—grasses and poplar—have no plastidial-type COG0354, presumably as a result of relatively recent losses in independent angiosperm lineages. However, inasmuch as either plant protein can complement the *E. coli* COG0354 deletion (Waller *et al.*, 2010), it seems likely that they could replace each other in plants, in which case dual targeting of the mitochondrial-type protein to mitochondria and plastids would provide a simple explanation. There are many examples of such dual targeting in plants (Mackenzie *et al.*, 2005; Millar *et al.*, 2006). In this connection it is noteworthy that the COSMOSS targeting algorithm (Mitschke *et al.*, 2009) predicts that the maize, rice, and poplar mitochondrial-type COG0354 proteins are dual targeted and that their *Arabidopsis*, grape, and castor bean counterparts are not (scores for the former trio being 0.70–0.76 and the latter 0.47–0.51). As the genomes of grasses and poplar encode other components of the plastidial Fe/S assembly machinery, the lack of a plastidial-type COG0354 in these plants is not attributable to the absence of the whole cluster assembly system.

The expression pattern of the *Arabidopsis* plastidial-type COG0354 protein (At1g60990) is consistent with a chloroplast location, being highest in young leaves. In contrast, the mitochondrial-type protein (At4g12130) is expressed principally in seeds and flowers. Within flowers, microarray data (Steinhauser *et al.*, 2004) and proteome analysis (Sheoran *et al.*, 2006) point to pollen as a main site of expression. The high expression of *At4g12130* mRNA and its translation product in pollen and seeds—tissues with low water contents—shows that both message and protein are stable to drying and so are available upon hydration to support repair processes (Bewley, 1997). Since desiccation entails oxidative stress (Smirnov, 1993), expression in drying and dry tissues is consistent with the evidence from *E. coli* that COG0354 helps maintain Fe/S protein function during oxidative stress (Waller *et al.*, 2010). High-level *At4g12130* expression in seeds may also contribute to the embryo lethality observed when this gene is ablated, although such lethality is characteristic of most components of the Fe/S cluster assembly machinery (Balk and Pilon, 2011).

The tests of recombinant plant COG0354 proteins in the surrogate *E. coli* system established that both of them require folate to function, as do COG0354 proteins from other kingdoms (Waller *et al.*, 2010). This conclusion rests on the

failure of either plant protein to restore activity to the MiaB Fe/S ‘reporter’ protein in a strain rendered completely folate deficient by deletion of the biosynthetic gene *folP*. FolP mediates the coupling of the pterin and *p*-aminobenzoate moieties of folate (de Crécy-Lagard *et al.*, 2007), and deletion of *folP* leaves pterin synthesis unaffected but greatly reduces MiaB activity (Waller *et al.*, 2010). The results thus make it very improbable that pterins can substitute for folates, and place the folate requirement for MiaB activity upstream of MiaB along with the requirement for the COG0354 protein itself (Fig. 3A). While the data for the *folP* deletion do not identify the type of folate that plant COG0354 proteins require, results with *E. coli* COG0354 implicate tetrahydrofolate as the active form (Waller *et al.*, 2010). The ability of the plant proteins to substitute for the *E. coli* protein implies that they, too, require tetrahydrofolate. As both mitochondria and chloroplasts contain tetrahydrofolates (Hanson and Gregory, 2011) the necessary folate cofactor would be available to both COG0354 proteins *in planta*.

In summary, the present data establish COG0354 proteins as components of both the mitochondrial and the plastidial Fe/S assembly machinery, and show that these types of machinery have a hitherto unsuspected folate requirement. The data also show that the mitochondrial COG0354 protein is required for embryo development. Finally, the loss of the plastidial-type protein from certain lineages raise interesting questions about the evolutionary forces that drove the loss of this protein and about its function in the species that retain it.

## Supplementary data

Supplementary data are available at *JXB* online.

**Table S1.** Oligonucleotide primers used in this study.

## Acknowledgements

This work was supported in part by National Science Foundation grant no. MCB-0839926 and by an endowment from the C. V. Griffin Sr Foundation. We thank J. Whelan for the soybean alternative oxidase clone, H. J. Klee and K. M. Folta for advice, and O. Frelin for assistance with transgenic plants.

## References

- Bailly M, Blaise M, Roy H, Denziak M, Lorber B, Birck C, Becker HD, Kern D. 2008. tRNA-dependent asparagine formation in prokaryotes: characterization, isolation and structural and functional analysis of a ribonucleoprotein particle generating Asn-tRNA<sup>Asn</sup>. *Methods* **44**, 146–163.
- Balk J, Pilon M. 2011. Ancient and essential: the assembly of iron–sulfur clusters in plants. *Trends in Plant Science* **16**, 218–226.
- Beinert H. 2000. Iron–sulfur proteins: ancient structures, still full of surprises. *Journal of Biological and Inorganic Chemistry* **5**, 2–15.

- Bewley JD.** 1997. Seed germination and dormancy. *The Plant Cell* **9**, 1055–1066.
- Bradford MM.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein–dye binding. *Analytical Biochemistry* **72**, 248–254.
- Clough SJ, Bent AF.** 1998. Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *The Plant Journal* **16**, 735–743.
- de Crécy-Lagard V, El Yacoubi B, de la Garza RD, Noiriel A, Hanson AD.** 2007. Comparative genomics of bacterial and plant folate synthesis and salvage: predictions and validations. *BMC Genomics* **8**, 245.
- Gelling C, Dawes IW, Richhardt N, Lill R, Mühlhoff U.** 2008. Mitochondrial Iba57p is required for Fe/S cluster formation on aconitase and activation of radical SAM enzymes. *Molecular and Cellular Biology* **28**, 1851–1861.
- Gibeaut DM, Hulett J, Cramer GR, Seemann JR.** 1997. Maximal biomass of *Arabidopsis thaliana* using a simple, low-maintenance hydroponic method and favorable environmental conditions. *Plant Physiology* **115**, 317–319.
- Gleave AP.** 1992. A versatile binary vector system with a T-DNA organizational structure conducive to efficient integration of cloned DNA into the plant genome. *Plant Molecular Biology* **20**, 1203–1207.
- Gould SB, Waller RR, McFadden GI.** 2008. Plastid evolution. *Annual Review of Plant Biology* **59**, 491–517.
- Hanson AD, Gregory JF 3<sup>rd</sup>.** 2011. Folate biosynthesis, turnover, and transport in plants. *Annual Review of Plant Biology* **62**, 4.1–4.21.
- Hernández HL, Pierrel F, Elleingand E, García-Serres R, Huynh BH, Johnson MK, Fontecave M, Atta M.** 2007. MiaB, a bifunctional radical- *S*-adenosylmethionine enzyme involved in the thiolation and methylation of tRNA, contains two essential [4Fe–4S] clusters. *Biochemistry* **46**, 5140–5147.
- Hu P, Janga SC, Babu M, et al.** 2009. Global functional atlas of *Escherichia coli* encompassing previously uncharacterized proteins. *PLoS Biology* **7**, e96.
- Johnson DC, Dean DR, Smith AD, Johnson MK.** 2005. Structure, function, and formation of biological iron–sulfur clusters. *Annual Review of Biochemistry* **74**, 247–281.
- Lill R.** 2009. Function and biogenesis of iron–sulfur proteins. *Nature* **460**, 831–838.
- Lin CN, Syu WJ, Sun WS, Chen JW, Chen TH, Don MJ, Wang SH.** 2010. A role of *ygfZ* in the *Escherichia coli* response to plumbagin challenge. *Journal of Biomedical Science* **17**, 84.
- Mackenzie SA.** 2005. Plant organellar protein targeting: a traffic plan still under construction. *Trends in Cell Biology* **15**, 548–554.
- Millar AH, Whelan J, Small I.** 2006. Recent surprises in protein targeting to mitochondria and plastids. *Current Opinion in Plant Biology* **9**, 610–615.
- Mitschke J, Fuss J, Blum T, Höglund A, Reski R, Kohlbacher O, Rensing SA.** 2009. Prediction of dual protein targeting to plant organelles. *New Phytologist* **183**, 224–235.
- Nilsson R, Schultz IJ, Pierce EL, et al.** 2009. Discovery of genes essential for heme biosynthesis through large-scale gene expression analysis. *Cell Metabolism* **10**, 119–130.
- Niwa Y.** 2003. A synthetic green fluorescent protein gene for plant biotechnology. *Plant Biotechnology* **20**, 1–11.
- Olinares PD, Ponnala L, van Wijk KJ.** 2010. Megadalton complexes in the chloroplast stroma of *Arabidopsis thaliana* characterized by size exclusion chromatography, mass spectrometry, and hierarchical clustering. *Molecular and Cellular Proteomics* **9**, 1594–1615.
- Ote T, Hashimoto M, Ikeuchi Y, Su’etsugu M, Suzuki T, Katayama T, Kato J.** 2006. Involvement of the *Escherichia coli* folate-binding protein YgfZ in RNA modification and regulation of chromosomal replication initiation. *Molecular Microbiology* **59**, 265–275.
- Pagliarini DJ, Calvo SE, Chang B, et al.** 2008. A mitochondrial protein compendium elucidates complex I disease biology. *Cell* **134**, 112–123.
- Phillips G, El Yacoubi B, Lyons B, Alvarez S, Iwata-Reuyl D, de Crécy-Lagard V.** 2008. The biosynthesis of the 7-deazaguanosine modified tRNA nucleosides: a new role for GTP cyclohydrolase I. *Journal of Bacteriology* **190**, 7876–7884.
- Pinon V, Ravanel S, Douce R, Alban C.** 2005. Biotin synthesis in plants. The first committed step of the pathway is catalyzed by a cytosolic 7-keto-8-aminopelargonic acid synthase. *Plant Physiology* **139**, 1666–1676.
- Pribat A, Noiriel A, Morse AM, et al.** 2010. Nonflowering plants possess a unique folate-dependent phenylalanine hydroxylase that is localized in chloroplasts. *The Plant Cell* **22**, 3410–3422.
- Ravanel S, Cherest H, Jabrin S, Grunwald D, Surdin-Kerjan Y, Douce R, Rébeillé F.** 2001. Tetrahydrofolate biosynthesis in plants: molecular and functional characterization of dihydrofolate synthetase and three isoforms of polyglutamate synthetase in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences, USA* **98**, 15360–15365.
- Rosso MG, Li Y, Strizhov N, Reiss B, Dekker K, Weisshaar B.** 2003. An *Arabidopsis thaliana* T-DNA mutagenized population (GABI-Kat) for flanking sequence tag-based reverse genetics. *Plant Molecular Biology* **53**, 247–259.
- Rudhe C, Chew O, Whelan J, Glaser E.** 2002. A novel *in vitro* system for simultaneous import of precursor proteins into mitochondria and chloroplasts. *The Plant Journal* **30**, 213–220.
- Sheftel A, Stehling O, Lill R.** 2010. Iron–sulfur proteins in health and disease. *Trends in Endocrinology and Metabolism* **21**, 302–314.
- Sheoran IS, Sproule KA, Olson DJ, Ross AR, Sawhney VK.** 2006. Proteome profile and functional classification of proteins in *Arabidopsis thaliana* (*Landsberg erecta*) mature pollen. *Sexual Plant Reproduction* **19**, 185–196.
- Smirnov N.** 1993. The role of active oxygen in the response of plants to water deficit and desiccation. *New Phytologist* **125**, 27–58.
- Steinhauser D, Usadel B, Luedemann A, Thimm O, Kopka J.** 2004. CSB.DB: a comprehensive systems-biology database. *Bioinformatics* **20**, 3647–3651.

- Tamura K, Dudley J, Nei M, Kumar S.** 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution* **24**, 1596–1599.
- Teplyakov A, Obmolova G, Sarikaya E, Pullalarevu S, Krajewski W, Galkin A, Howard AJ, Herzberg O, Gilliland GL.** 2004. Crystal structure of the YgfZ protein from *Escherichia coli* suggests a folate-dependent regulatory role in one-carbon metabolism. *Journal of Bacteriology* **186**, 7134–7140.
- Turner WL, Waller JC, Snedden WA.** 2005. Identification, molecular cloning and functional characterization of a novel NADH kinase from *Arabidopsis thaliana* (thale cress). *Biochemical Journal* **385**, 217–223.
- Waller JC, Alvarez S, Naponelli V, et al.** 2010. A role for tetrahydrofolates in the metabolism of iron–sulfur clusters in all domains of life. *Proceedings of the National Academy of Sciences, USA* **107**, 10412–10417.
- Williams KP, Sobral BW, Dickerman AW.** 2007. A robust species tree for the alphaproteobacteria. *Journal of Bacteriology* **189**, 4578–4586.
- Xiang D, Venglat P, Tibiche C, et al.** 2011. Genome-wide analysis reveals gene expression and metabolic network dynamics during embryo development in *Arabidopsis*. *Plant Physiology* **156**, 346–356.
- Xu XM, Møller SG.** 2008. Iron–sulfur cluster biogenesis systems and their crosstalk. *ChemBiochem* **9**, 2355–2362.