Two *Chlamydomonas* OPR proteins stabilize chloroplast mRNAs encoding small subunits of photosystem II and cytochrome b_6f.

Fei Wang[1,2], Xenie Johnson[1§], Marina Cavaiuolo[1], Alexandra-Viola Bohne[2], Joerg Nickelsen[2] and Olivier Vallon[1*]

1 UMR 7141, Centre National de la Recherche Scientifique/Université Pierre et Marie Curie, Institut de Biologie Physico-Chimique, Paris 75005, France
2 Biozentrum Ludwig-Maximilians-Universität München, D-82152 Planegg-Martinsried, Germany

* corresponding author: Olivier Vallon
UMR7141, Institut de Biologie Physico-Chimique
13 rue Pierre et Marie Curie
75005 Paris
Tel: +33 1 5841 5058 ; Fax: : +33 1 5841 5022
ovallon@ibpc.fr

Other E-mails:
Fei Wang <fwang831212@gmail.com>
Xenie Johnson <xenie.johnson@ibpc.fr>
Marina Cavaiuolo <cavaiuolo@ibpc.fr>
Alex Bohne <alexbohne@yahoo.de>
Joerg Nickelsen <joerg.nickelsen@lrz.uni-muenchen.de>

§ current address: CEA, IBEB, Lab Bioenerget Biotechnol Bacteries & Microalgues, CNRS, UMR Biol Veget & Microbiol Environ, Aix-Marseille Université, Saint-Paul-lez-Durance, F-13108, France.
running title : OPR proteins stabilizing chloroplast mRNAs

Keywords: OctotrichoPeptide Repeat; chloroplast gene expression; photosystem II; cytochrome b_6f; *Chlamydomonas reinhardtii*; *psbI*; *petG*; RNA-binding protein; RNA footprint.

Summary

In plants and algae, chloroplast gene expression is controlled by nucleus-encoded proteins that bind to mRNAs in a specific manner, stabilizing mRNAs or promoting their splicing, editing, or translation. Here, we present the characterization of two mRNA stabilization factors of the green alga *Chlamydomonas reinhardtii*, which both belong to the OctotricoPeptide Repeat (OPR) family. MCG1 is necessary to stabilize the *petG* mRNA, encoding a small subunit of the cytochrome b_6f complex, while MBI1 stabilizes the *psbI* mRNA, coding for a small subunit of Photosystem II. In the *mcg1* mutant, the small RNA footprint corresponding to the 5' end of the *petG* transcript is reduced in abundance. In both cases, the absence of the small subunit perturbs assembly of the cognate complex. Whereas PetG is essential for formation of a functional cytochrome b_6f dimer, PsbI appears partly dispensable as a low level of PSII activity can still be measured in its absence. Thus, nuclear control of chloroplast gene expression is not only exerted on the major core subunits of the complexes, but also on small subunits with a single transmembrane helix. While OPR proteins have thus far been involved in translation or trans-splicing of plastid mRNAs, our results expand the potential roles of this repeat family to their stabilization.

Introduction

Following the establishment of mitochondrial and plastid endosymbiosis, the genomes of both organelles have undergone extensive streamlining, resulting from combined gene loss and gene transfer to the nuclear genome. Yet, both organelles have retained their own gene expression machinery, with many ancestral bacterial features, for example in the working of the ribosomes, tRNAs, RNases and, in the case of the plastid, of the RNA polymerase. In turn, the necessary cooperation between the organellar and nuclear genomes has fundamentally reshaped the mechanisms of gene regulation. Control of gene expression is now exerted mostly at the post-transcriptional, rather than transcriptional level. And the key players of this control are now encoded in the nucleus, in the form of Regulators of Organelle Gene expression (ROGEs). These proteins are translated in the cytosol and imported into the organelles where they regulate the accumulation, processing, and translation of mRNAs in a gene-specific manner. In the model green alga *Chlamydomonas reinhardtii*, the term "M-

factors" has been coined for proteins involved in the stabilization and splicing of plastid mRNAs, while "T-factors" are necessary for translation initiation. Most of the M- and T-factors identified thus far have been found to carry tandem repeats of a short amino-acid motif, allowing sequence-specific protein-RNA interaction. The best-characterized of these repeat protein families is the PentatricoPeptide repeat (PPR family), with 14 members in *Chlamydomonas* (Tourasse *et al.* 2013) and 458 in Arabidopsis (O'Toole *et al.* 2008). While PPR proteins in land plants have been implicated in all aspects of mRNA metabolism, from stabilization to splicing to editing to translation (Barkan and Small 2014), the two characterized *Chlamydomonas* PPRs are both M-factors: MCA1 stabilizes the *petA* mRNA (Loiselay *et al.* 2008), while MRL1, which is conserved in land plants, stabilizes the *rbcL* transcript (Johnson *et al.* 2010). Other well-characterized M-factors are MBB1 (and its ortholog in land plants, Hcf107) and NAC2, which both belongs to the TetratricoPeptide repeat (TPR) family, or more precisely to the HAT (Half-A-Tetratricopeptide) repeat family (Boudreau *et al.* 2000, Loizeau *et al.* 2014, Vaistij *et al.* 2000).

However, in *Chlamydomonas*, most identified ROGEs belong to a less-well characterized family, the OctotricoPeptide Repeat (OPR) family (Eberhard *et al.* 2011, Rahire *et al.* 2012). The sequence of the repeat itself is completely unrelated to the PPR and TPR repeats, but its length is similar and it is also predicted to fold as a pair of α-helices. OPR proteins thus presumably fold as a super-helical solenoid, as do PPRs and TPRs, and they also probably bind RNA in their inner groove by virtue of specific interactions between certain residues of the protein and the nucleotide bases in the target RNA. The OPR family has a single representative in land plants, but is much more developed in *Chlamydomonas*, with 42 or 43 members depending on the analysis (Eberhard*, et al.* 2011, Rahire*, et al.* 2012). Recently, the sole OPR protein of Arabidopsis has been shown to be involved in RNA processing, acting on the 5' end of the 16S ribosomal RNA (Kleinknecht *et al.* 2014). However, most studies on OPRs have been performed in *C. reinhardtii*. Initially described when the *psbC* translation enhancer TBC2 was identified (Auchincloss *et al.* 2002), the family has seen its functional repertoire expand since, in addition to two other T-factors, TDA1 (Eberhard*, et al.* 2011) and TAB1 (Rahire*, et al.* 2012), it was shown to comprise two proteins involved in trans-splicing of the *psaA* mRNA: RAT2 is necessary for the 3' processing of the *tscA* RNA which forms part of intron 1 (Balczun *et al.* 2005), while RAA1 is necessary for the trans-splicing of both introns 1 and 2 (Merendino *et al.* 2006).

In this study, we describe two *C. reinhardtii* OPR proteins, MBI1 and MCG1, involved in the stabilization of the *psbI* and *petG* transcripts, respectively. This shows that OPR proteins can serve primarily in the stabilization of an mRNA, rather than in its translation or processing.

Results

Isolation of *Chlamydomonas* photosynthetic mutants and identification of two mutated loci.

Using insertional mutagenesis, we generated two collections of photosynthesis mutants of *C. reinhardtii*. Two antibiotics resistance markers, *ble* and *aphVIII* conferring resistance respectively to zeocin or to paromomycin, were introduced by electroporation into the genome of two distinct phototrophic recipient strains, either into the cell-wall less XS1 (Johnson *et al.* 2007) or the walled Jex4 (Houille-Vernes *et al.* 2011). Primary screening for a non-phototrophic (acetate-requiring) phenotype was followed by analysis of the fluorescence induction curves to identify the affected photosynthetic complex.

We here present four of the many mutants lacking cytochrome b_6f activity identified in the XS1 library. For reason that will be explained later, we will refer to these mutants as *mcg1-1*, *mcg1-2*, *mcg1-3* and *mcg1-4*. While *mcg1-1* had been generated using the *ble* cassette, *mcg1-2*, *mcg1-3* and *mcg1-4* had been obtained in different transformation experiments with *aphVIII*. When the mutants were backcrossed to the wild-type (WT) strain t222, the non-photosynthetic phenotype always segregated 2:2 (between 14 and 25 tetrads for each cross), indicating a single mutation in a nuclear gene. For *mcg1-1*, the phenotype was not linked to zeocin resistance in the cross, and at the time of crossing the three *aphVIII* transformants were found to have lost paromomycin resistance. To identify the mutated gene(s), we thus resorted to complementation by an indexed cosmid library (Depege *et al.* 2003, Kuras *et al.* 2007). Surprisingly, we found that the four cytochrome b_6f mutants could be complemented by the same cosmid. This cosmid carries a 31,541 bp portion of the genome (Chromosome_10:1577044-1608584, Figure 1A) containing five genes (www.phytozome.org). When the cosmid DNA was cut with restriction endonucleases *Bst*XI or *Nsi*I prior to transformation, complementation was still observed, but not with *Ear*I, *Sma*I, *Dra*III or *Nco*I. A single gene, which we call *MCG1*, was found to contain sites for the latter, but not the former enzymes: Cre10.g429400 (Figure 1A). Southern blot analysis (Figure S1) showed that *mcg1-3* carried a structural rearrangement in the gene, confirming that mutation of *MCG1* was the cause of the observed phenotype.

In a separate screen of a library obtained by transforming Jex4 with the *aphVIII* cassette, we isolated strain *mbi1*, with a marked deficiency in photosystem II (PSII) activity. Because of the residual growth on Minimum medium (Figure 2), cloning by complementation was impossible and we sought to obtain a Flanking Sequence Tag (FST) that would reveal the cassette insertion locus. Using inverse PCR, we were able to amplify a FST from the 3' end of the cassette, and mapped it to the fourth and last intron of gene Cre06.g272450 (Figure 1B and Data S1). In a backcross to the wild-type strain WT-S24, the PSII-deficient phenotype segregated 2:2, indicating a single nuclear mutation. PCR genotyping of the progeny showed tight linkage with the insertion. Figure 2B shows the result for 7 of the 40 clones tested (20 phototrophic and 20 non-phototrophic). Final proof that the PSII phenotype was due to the insertion in Cre06.g272450 was provided by the successful complementation of the mutant with *Sca*I-digested DNA of a Bacterial Artificial Chromosome (28D19) covering this locus (*Sca*I cuts 7 of the 13 genes in the BAC, but not Cre06.g272450). Transformants, selected as fast-growing dark-green colonies, were checked by PCR with primers P1 and P2 for integration of Cre06.g272450, and showed fully-restored phototrophic growth (Figure 2). Our FST data also showed that a retrotransposition event had accompanied the insertion of the cassette (Data S1), preventing us from sequencing the junction at its 5' end, but PCR amplification upstream of the insertion point (dotted lines on Figure 1B) showed that the Cre06.g272450 gene had not undergone extensive deletion. This was confirmed by Southern blot analysis with a probe flanking the insertion point (Figure 1B), where two BamHI fragments are observed. FST sequencing also showed that the cassette had incurred a 260 bp deletion upstream of the *RBCS2* promoter, removing the *HSP70A* element known to prevent transcriptional silencing of transgenes (Schroda *et al.* 2002). Accordingly, genetic analysis showed that the insertion in Cre06.g272450 was not functional: paromomycin resistance segregated 2:2 in tetrads but independently from the photosynthesis phenotype, thus revealing the presence of another (functional) resistance cassette.

MCG1 and MBI1 belong to the OPR protein family

The *MCG1* and *MBI1* gene structures are fully supported by EST evidence (http://genomes.mcdb.ucla.edu/Cre454/). Note that *MCG1* presents an additional minor splice variant (JGI v4 gene model ID: 536506). MCG1 and MBI1 are large proteins (~178 kDa and ~138 kDa respectively) which both show the characteristic OPR of ~38 amino acids previously described (Eberhard*, et al.* 2011, Rahire*, et al.* 2012). Both are well conserved in the closely related colonial alga *Volvox carteri* (Figure S2), and a more remote ortholog of

MBI1 can be found in the transcriptome of *C. acidophila* (Data S2). But no ortholog of either gene was found in other algal genomes or transcriptomes (including more distantly related Volvocale such as *C. chlamydogama*, *C. euryale*, *C. leiostraca* and *C. sp* CCMP681). Based on the results of the repeat-searching algorithm FT-Rep (Rahire*, et al.* 2012) and careful comparison of the Chlamydomonas and Volvox sequences, MCG1 and MBI1 were found to present a total of 12 and 17 repeats, respectively, grouped in three or four blocks (Figure 1 and Figure S1). Prediction of intracellular targeting using TargetP (Emanuelsson *et al.* 2000) and Predalgo (Tardif *et al.* 2012) leaned towards a mitochondrial location, but our genetic analysis leaves no doubt that both proteins act in the plastid. Assuming chloroplast localization, the program ChloroP predicts targeting peptides of 22, 19, and 74 amino acids for the main and minor MCG1 variants and for MBI1, respectively.

PSII and cytochrome b_6f in the *mbi1* and *mcg1* mutants

To determine whether residual cytochrome b_6f and PSII electron transport was still occurring in the *mcg1* and *mbi1* mutants, we used fluorescence induction kinetics and flash absorption spectroscopy measurements (Table 1 and Figures S2, S3). Variable fluorescence (Fv/Fm), a measure of PSII activity, was reduced 4.2-fold in the *mbi1* mutant. The PSII/PSI ratio, measured from the amplitude of the initial electrochromic shift (ECS) after a flash (Joliot and Delosme 1974) was about seven times lower in *mbi1* than in wild-type. Altogether, these results confirm that PSII charge separation, even though severely reduced in the *mbi1* mutant, is not completely abolished. In contrast, no cytochrome b_6f activity could be detected in *mcg1* mutants: fluorescence rose rapidly almost to Fm, indicating absence of electron transfer beyond the PQ pool, and the slow phase of the ECS kinetics (Joliot and Joliot 1985) was completely abolished.

In the *mcg1* mutants, all the cytochrome b_6f core subunits (cytochrome *f*, cytochrome b_6, Rieske protein and subunit IV) were drastically reduced in abundance compared to WT or to the *pcy1-2* mutant lacking plastocyanin (Johnson*, et al.* 2007), but none was completely missing (Figure 3A): cytochrome *f* accumulated to about 10% of WT, the other subunits to levels comparable to those observed in mutant *mca1-7* lacking cytochrome *f* (Johnson*, et al.* 2007, Loiselay*, et al.* 2008). This suggested that the core subunits of the complex were still produced, but strongly destabilized due to the absence of another subunit. A similar conclusion was reached for the *mbi1* mutant, where we observed a substantial accumulation of the large PSII subunits, for example D2 (approximately 25% of WT, Figure 3B). Northern blot analysis (Figure S4) showed no defect in the RNA levels for the major PSII subunits..

Protein pulse-labelling experiments with radioactive ^{35}S-sulphate (Figure S5) showed normal translation of D1, D2 and CP43. Altogether, these data suggest that *mbi1* is deficient in the production of a small PSII subunit.

To probe assembly of the complexes, thylakoid membrane proteins were separated by 2D electrophoresis (Blue Native PAGE in first dimension, SDS-PAGE in second dimension) and analyzed by Western blotting. In *mbi1*, the PSII monomer was detected, as well as the RC47 complex lacking CP43 and free CP43 (Figure 4). However, the mutant showed almost no PSII dimer, and a complete absence of PSII supercomplexes. We examined the assembly of newly-synthesized proteins after ^{35}S pulse-labeling (Figure S6), and found that the mutant accumulated normal amounts of radioactivity in PSII monomers, but very little in the region of the dimer and almost none in the supercomplexes. We conclude that the *mbi1* mutation perturbs the assembly of the PSII dimer, without completely preventing it. As a consequence, higher order assemblies with antenna proteins are not formed. A similar analysis of the *mcg1-1* mutant (Figure S7) showed that the cytochrome b_6f dimer, which constitutes the functional unit (Stroebel *et al.* 2003), was entirely missing. Trace amounts of the cytochrome b_6f monomer lacking the Rieske protein were also observed, as well as free Rieske.

MBI1 stabilizes the *psbI* mRNA

Since all previous studies on OPR proteins have uncovered a role in post-transcriptional regulation of chloroplast gene expression, we set out to identify chloroplast *psb* and *pet* genes whose mRNA accumulation or translation would be affected. Since *mbi1* produced the four major PSII subunits, we concentrated on the low molecular weight subunits of PSII, excluding those like *psbE*, *psbF* or *psbH* that are necessary for formation of the PSII center (O'Connor *et al.* 1998, Pakrasi *et al.* 1989) or *psbJ* whose deletion allows formation of PSII-LHCII supercomplexes (Swiatek *et al.* 2003). Northern blot analysis showed normal accumulation of *psbK, psbL, psbM, psbT,* and *psbZ* mRNAs in the *mbi1* mutant (Figure 5)*,* but complete absence of the *psbI* transcript. This gene encodes a small subunit of 37 amino acids, which was shown in tobacco to be involved in PSII dimer formation (Schwenkert *et al.* 2006). The accepted gene nomenclature for ROGEs in *Chlamydomonas* uses **M** or **T** as the first letter for mRNA modification/binding and translation enhancement, respectively; **A**, **B**, **C**, **D** or **R** as the second letter for targets in PSI, PSII, cytochrome b6f, ATP synthase and RuBisCO respectively; and a third letter corresponding to the last one of the target gene. We therefore coined the name *MBI1* for the mutated gene (**M**-factor for ps**bI**).

In *C. reinhardtii*, the *psbI* gene is co-transcribed with *atpA*, *cemA* and *atpH*, respectively encoding the α-CF1 subunit of the ATP synthase, a chloroplast envelope protein and subunit III of the ATP synthase CFo. The mature 5' end of the transcript lies 92 nt upstream of the translation initiation codon (Drapier *et al.* 1998). In contrast to previous assumptions, the *psbI* mRNA is transcribed only from the upstream *atpA* promoter (S.-I. Ozawa, Y. Choquet, personal communication), generating a tetracistronic transcript that is processed to tri-, di- and mono-cistronic forms (Figure 6A). Northern blot with probes covering various regions of the *atpA-psbI-cemA-atpH* gene cluster (Figure 6B), revealed a complete absence of the major mono-cistronic form (#7), but traces of the di-, tri- and tetracistronic transcripts (#3, 2 and 1). Results with downstream probes indicated that the two transcripts with the same 5' end as the mono-cistronic *psbI* (#5 and 6) also failed to accumulate. Therefore, we conclude that the role of the MBI1 protein is to stabilize transcripts whose 5' end lies immediately upstream of *psbI*, presumably by protecting them from 5'->3' degradation. The precursors of the missing transcripts, i.e. transcripts #1, 2, and 3, do not overaccumulate in the mutant, suggesting that MBI is not involved in the processing itself but in transcript stabilization. Interestingly, the *atpA* monocistronic transcript (#4) overaccumulated in the mutant: we propose that the additional signal comes from 3'->5' exonucleolytic degradation of longer transcripts which in the absence of MBI1 can proceed up to the mature *atpA* 3' end.

MCG1 stabilizes the *petG* mRNA by protecting its 5' end

For MCG1, the only candidate targets were *petL* (Takahashi *et al.* 1996) and *petG* (Fong and Surzycki 1992). As shown in Figure 7, the two major *petL* transcripts accumulated normally, while the 0.7 kb *petG* mRNA (Berthold *et al.* 1995) was undetectable. The *petG* gene is located dowsntream of *psbF* and *psbL* (Fong and Surzycki 1992) which both showed normal accumulation of their transcripts (Figure S8). We conclude that *MCG1* is necessary for the production or stabilization of the *petG* mRNA, and thence derive its name. Using RLM-RACE (RNA-Ligation Mediated Rapid Amplification of cDNA ends, (Liu and Gorovsky 1993), we determined the position of the *petG* 5' end in the WT (Figure S9). It lies 277 bp upstream of the translation start, with no candidate "-10" promoter sequence immediately upstream. Accordingly, the corresponding band at ~530 bp was observed even without pyrophosphatase treatment, indicating that it is a cleavage product rather than a primary transcript. The 530 bp band was completely absent from the mutant, in line with its

total lack of *petG* transcript. A minor shorter band was also detected in both the WT and the mutant. In both strains, its 5' end lies 156 nt upstream of the translation start, again with no promoter sequence upstream. It does not correspond to a *petG* mRNA detectable by Northern blotting, and we suspect that it has little biological significance.

RNA "footprints" correspond to short regions of the chloroplast transcripts protected from degradation by the specific binding of a ROGE (Ruwe and Schmitz-Linneweber 2012). Recently, data mining of *Chlamydomonas* sRNA data has unraveled a series of putative footprints in this organism (Loizeau*, et al.* 2014). One of them corresponds exactly to the 5' end of *petG*, as uncovered in our RLM-RACE experiments, and we thought that MCG1 could be responsible for its stabilization. Using a primer corresponding to this footprint, we measured its abundance by qRT-PCR and found a severe reduction of the footprint in the mutant, with a *mcg1-1*/WT ratio of 0.10-0.13 ± 0.04.

Discussion

As commonly found in the maintenance of biological systems, control of organelle genes by nucleus-encoded ROGEs carries a cost, in terms of genetic burden (the need to maintain and regulate the ROGE genes in the genome) and energy investment (the need to synthesize and import the regulatory proteins). From this point of view, it could be expected that this type of control would be limited to the large subunits of the photosynthetic enzymes whose synthesis is costly or which bind potentially toxic reactive co-factors. In *Chlamydomonas*, up to now only the large subunits of the photosynthetic complexes have been shown to be under such control, be it for cytochrome b_6f (Girard-Bascou *et al.* 1995, Loiselay*, et al.* 2008, Murakami *et al.* 2005, Wostrikoff *et al.* 2001), PSII (Auchincloss*, et al.* 2002, Boudreau*, et al.* 2000, Somanchi *et al.* 2005, Vaistij*, et al.* 2000), PSI (Rahire*, et al.* 2012, Rochaix *et al.* 2004) or the ATP synthetase (Drapier *et al.* 1992, Eberhard*, et al.* 2011, Lemaire and Wollman 1989). In land plants, PPR10 stabilizes the monocistronic mRNA for *psaJ* but it also acts on *rpl33*, *atpI* and *atpH* (Pfalz *et al.* 2009), while PGR3 stabilizes the *petL* transcripts in addition to its effect on *ndhA*, and CRP1 activates *psaC* translation (Schmitz-Linneweber *et al.* 2005) in addition to acting on *petB*, *petD* and *petA* (Barkan *et al.* 1994). In this study, we show that the genes for *petG* and *psbI*, two small subunits (37 residues) of the cytochrome b_6f and PSII complexes are specifically controlled by nucleus-encoded ROGEs belonging to the OPR family of RNA-binding proteins. In both cases, the mutant phenotypes suggest that the small subunit is the main, if not sole, target. In terms of economy of resources, the benefit seems at best minimal.

Two categories of explanations can be presented for the existence of ROGE genes necessary for the stabilization of organelle mRNAs. Constructive Neutral Evolution (CNE) (Covello and Gray 1993, Stoltzfus 1999) has been invoked to explain such apprently uselessly complex traits of organelle gene expression as RNA editing and splicing (Lukes *et al.* 2011) or trans-splicing of *psaA* (Lefebvre-Legendre *et al.* 2014). In the case of the M-factors we present here, CNE would posit that the repeat proteins that gave rise to MCG1 or MBI1 entered in a selectively neutral interaction with the 5' regions of *petG* and *psbI* mRNAs, respectively. This did not initially present the cell with a selective advantage. But when later on the plastome mutated or rearranged so that the mRNAs became sensitive to 5'->3' exonucleases, MCG1 and MBI1 became essential as they allowed some stabilization of the RNAs they bound to. This dependence was fixed by a ratchet mechanism whereby additional plastome mutations prevented the targets to regain their original intrinsic stability. The mRNA and the protein then co-evolved, leading to enhanced fit.

CNE is very appealing for RNA editing and trans-splicing, where presuppression appears absolutely necessary in the light of the very severe phenotype of the initial chloroplast mutations. But for M-factors, more classical selectionist views are also reasonable because the initial mutation probably has a milder phenotype (polycistronic mRNAs are usually only partially cleaved). A new RNA-binding specificity, arising by duplication and neofunctionalization of an OPR or PPR gene, can then be selected for because by increasing the stability of the transcript it reduces the energetic cost of gene expression. Additional mutations may then allow a better stabilization of the cleaved transcript, which in time may become the only stable form (*petG*), or not (*psbI*).

Whatever the scenario, even a CNE origin of M-factors is not incompatible with their later acquiring adaptive functions. In this category, regulation stands out as an attractive candidate. If the accumulation of MCG1 or MBI1 is limiting for the accumulation of the *petG* and *psbI* mRNAs, changes in their expression could be used by the cell to change the level of accumulation of cytochrome b_6f and PSII. Such a mechanism has been proposed for MCA1 (Raynaud *et al.* 2007), but it is yet to be generalized to other ROGEs. Based on available data, neither *MCG1* nor *MBI1* appeared to change their transcript level during nitrogen or sulfur starvation, two conditions that are known to affect cytochrome *b6f* and PSII accumulation, respectively (Boyle *et al.* 2012, Miller *et al.* 2010, Wei *et al.* 2014). However, cursory examination of the gene expression data on the Phytozome browser revealed two RNAseq experiments showing interesting variations of *MCG1* and *MBI1* expression levels (Table S2 lists the conditions for all the experiments depicted in the Phytozome browser, as collected

from the SRA database). Figure 8 (top panel) shows the effect of light and biliverdin treatment in WT and a mutant lacking plastidial heme oxygenase (*hmox1*) and therefore deficient in the biliverdin-mediated retrograde signaling pathway (Duanmu *et al.* 2013). *MBI1* appears to be downregulated by light in the WT, more than in the mutant background. Unexpectedly, this light effect is abolished by biliverdin, or even reversed in the mutant. The response of *MCG1* is completely different: it shows a massive induction by the 0.5 h white light treatment, a response diminished by the *hmox1* mutation or biliverdin treatment. This study noted that genes induced by light and suppressed by biliverdin were enriched for protein families involved in high light stress and this effect was attributed to the induction by biliverdin of the response to oxidative stress (Duanmu et al., 2013).

In this respect, it is interesting to note that another experiment analyzing oxidative stress also shows significant variations, this time in the same direction for *MCG1* and *MBI1* (Figure 8, bottom panel): both genes appeared to be repressed by H_2O_2 treatment (Urzica *et al.* 2012). It is tempting to speculate that this differential response of ROGEs controlling the biogenesis of PSII and cytochrome b_6f reflects different needs for building up or repairing the complexes, and that this could be regulated by ROS signaling. But grasping the exact physiological significance of these changes would require more in-depth study, in particular measuring protein level for MCG1 and MBI1 and mRNA levels for *petG* and *psbI* in the conditions tested. This nevertheless lends support to the notion that ROGEs offer opportunities for the nuclear genome to regulate, based on environmental conditions, the expression of chloroplast genes coding for photosynthetic function. Interestingly, *MCA1*, *TCA1* and *MCD1*, other ROGEs involved in cytochrome b_6f biogenesis, appear to react to light and biliverdin signaling in a similar manner to *MCG1*, while the response of *MBB1* which controls *psbB* mRNA levels resembles that of *MBI1* (see Phytozome website). Still, other PSII ROGEs (*NAC2*, *TBC2*) behave differently from *MBI1*, so it seems that different regulons can be involved in the biogenesis of a given complex. What is clear is that many ROGE genes respond to environmental signals. The coordination of changes in gene expression may require that all photosynthetic genes, including those for the smaller subunits, be under control of nucleus-encoded regulators, fine-tuning their activity to optimize acclimation of the photosynthetic apparatus to changing environmental conditions. As mentioned above, the acquisition of such regulatory properties is not in contradiction with the genes originating by CNE: genes involved in the trans-splicing of *psaA* also show light induction (*RAA1, RAA2*) or repression (*RAA3*), repression by H_2O_2 (*RAA1*) or induction by anoxia (*RAA2*).

Based on our result, we anticipate that other M-factors will be found to belong to the OPR family, as MCG1 and MBI1 both do. Up to now, *Chlamydomonas* OPR proteins have been found to be involved in mRNA processing (more precisely, *psaA* trans-splicing) (Balczun, *et al.* 2005, Merendino, *et al.* 2006) and in translation control (Auchincloss, *et al.* 2002, Eberhard, *et al.* 2011, Rahire, *et al.* 2012). It is not clear which feature of an OPR (or more generally, of any type of ROGE) defines the nature of the process it controls. Some ROGEs have been found to act in different ways on different mRNAs they bind to, for example CRP1 (Barkan, *et al.* 1994) or PPR10 (Prikryl *et al.* 2011) in maize. More generally, it can be considered that the primary molecular function of a ROGE is to bind to its target, and the biological function (enhancement of splicing, translation initiation, editing, prevention of exonucleolytic degradation, or any other effect) depends either on the properties of additional domains of the ROGE and partner proteins, or on the changes induced in the target by the binding of the ROGE (unfolding, folding, unmasking or masking of specific regions). In this way, it is not surprising that OPRs, like PPRs and TPRs, also prove able to bind permanently to specific mRNAs, and thus stabilize their 5' or 3' ends. Here again, we expect that many OPRs, especially in Chlorophycean algae where they are so numerous, will be found to function as M- or T-factors.

One of the key features expected for an M-factor is that its binding should be tight enough to efficiently protect the RNA from exonucleolytic degradation by stopping the progress of an exonuclease (Prikryl, *et al.* 2011, Zhelyazkova *et al.* 2012). This has been shown to result in the accumulation of short RNAs in the chloroplast, presumably resulting from the protection of the RNA against nucleases (Ruwe and Schmitz-Linneweber 2012). In *Chlamydomonas*, a recent survey of published small RNA data has unraveled a few of those footprints, in particular those of the MBB1 protein at the 5' end of *psbB* and *psbH* transcripts (Loizeau, *et al.* 2014). One such footprint was found between *atpA* and *psbI*, at position 126,304. However, it lies in the 3' UTR of *atpA*, 272 nt upstream of the *psbI* 5' end determined by (Drapier, *et al.* 1998), and probably does not represent the footprint of MBI1. Our attempts to amplify a footprint starting at the *psbI* 5' end were unsuccessful. Interestingly, however, (Loizeau, *et al.* 2014) also uncovered a 21 nt small RNA that maps exactly at the 5' end of the *petG* mRNA determined in the present study. Our finding that this sRNA is severely reduced in *mcg1* strongly suggests that it is indeed the footprint of the MCG1 protein. It is not completely absent, though, which suggests that our assay can also capture unprotected mRNA fragments on their way to complete degradation. Methods need to be developed where unprotected RNA is efficiently degraded prior to the footprint assay.

While *MCG1* inactivation results in a complete loss of cytochrome b_6f assembly and function, that of *MBI1* has a much milder phenotype: PSII is still partly assembled and active, and can still support some level of photo-autotrophic growth. This is in line with previous studies on transplastomic deletion mutants: deletion of *petG* completely prevents cytochrome b_6f assembly and function (Berthold*, et al.* 1995, Schneider *et al.* 2007, Schwenkert *et al.* 2007), while *psbI* deletants still assemble PSII and are photo-autotrophic (Ikeuchi *et al.* 1995, Künstner *et al.* 1995, Schwenkert*, et al.* 2006). The reason for this difference can be found in the different roles that PetG and PsbI play in the assembly and stability of the cytochrome b_6f and PSII proteins. In the 3D-structures of the two dimeric complexes (Figure S10), PetG appears sandwiched between PetD and the other small subunits PetL, PetM and PetN, with hardly any surface exposed to the bilayer, while PsbI is found at the periphery of the protein, interacting exclusively with D1 and CP43. It is therefore not surprising that the absence of PetG would prevent proper interaction between the other subunits of the cytochrome b_6f complex, while that of PsbI would only perturb the latest steps of PSII assembly. Note that the exact consequence of the absence of PsbI seems to vary between organisms. All studies found accumulation of monomeric PSII and CP43-less core at the expense of dimers, but this was ascribed to a lesser stability of the higher order assemblies in tobacco (Schwenkert*, et al.* 2006) while studies in *Synechocystis* rather indicated a lesser stability of the interaction with CP43 preventing further assembly steps (Dobakova *et al.* 2007). An *in vitro* study of a thermophilic cyanobacterial PSII even showed that the stability of the dimer, once formed, did not depend on PsbI (Kawakami *et al.* 2011). It was not the aim of our study to resolve this dispute, but our results, in particular our finding that the assembly of *de novo* synthesized D1/D2 is altered in the mutant (Figure S6), rather points to an assembly defect.

But is PsbI completely missing in the *mbi1* mutant? In the absence of an antibody to the protein, we can only rely on mRNA accumulation and mutant phenotypes. The monocistronic *psbI* transcript is undetectable, but other transcripts containing the *psbI* CDS are still accumulated: the rather abundant transcript #3, as well as the minor transcripts #1 and 2 (Figure 11). However, if PsbI could be translated from these messengers, we would expect to see the accumulation of at least a low level of stable PSII dimers and supercomplexes. This, however, was not observed: even PSII dimers were hardly detectable, including in pulse-labelling experiments. In addition, the phenotype of *mbi1* is very similar to that described for the various *psbI* deletion strains characterized in *Chlamydomonas* (Drapier*, et al.* 1998, Künstner*, et al.* 1995). We thus favor the hypothesis that in the absence of the MBI1 protein, stable transcripts that do not start at the 5' end of *psbI* are not translatable. This

could be due to an intrinsic inability of the poly-cistronic transcripts to undergo translation initiation at inner cistrons. However, we rather like to envision a more specific effect, namely that MBI1, in addition to stabilizing the transcript, helps its translation. Such a dual action has been shown for PPR10 (Prikryl, *et al.* 2011) and HCF107 (Hammani *et al.* 2012) in land plants and for MCA1 in Chlamydomonas, in this case via its interaction with the TCA1 translation enhancer (Boulouis *et al.* 2011). The AtpH protein also is not produced in mutant *mdh1-ac46* that lacks the monocistronic mRNA, even if poly-cistronic transcripts still accumulate (Drapier, *et al.* 1998). While a direct effect on the structure of the translation initiation region is possible, interaction with other proteins need in some cases be considered. Multiprotein complexes involving various types of ROGEs (including OPRs) have been found to mediate essential steps of chloroplast mRNA maturation or expression (Jacobs *et al.* 2013).

Is the production of CemA also affected by the *mbi1* mutation? In other words, can CemA production be supported by the two polycistronic transcripts #1 and #2, which in the absence of MBI1 are the only transcripts that carry the *cemA* CDS? In the absence of an antibody and in view of the mild phenotype of the *cemA* deletion (Rolland *et al.* 1997), we cannot conclude, but we note that no mono-cistronic transcript has ever been detected for this gene, so that a coupling between mRNA stabilization and translation seems less likely.

Experimental Procedures

Chlamydomonas **strains**

Chlamydomonas strains were grown in Tris-acetate-phosphate (TAP) or Minimum medium, under continuous light (Harris, 1989). For the description of recipient strain XS1 (*cw15 arg7 mt+*) and the production and complementation of mutants, refer to Johnson et al. (2007). For cell-walled strain Jex4 (Houille-Vernes, *et al.* 2011), transformation by electroporation was performed (Raynaud, *et al.* 2007) using plasmid pBC1 cut with *SacI* and *KpnI*.

Spectroscopy and Fluorescence Measurements

The JTS-10 spectrophotometer (Biologic, Grenoble, France) was used for *in vivo* spectroscopy measurement. Cells grown in 100 mL TAP were harvested in the mid-exponential phase and resuspended in 2 mL Hepes (20 mM, pH 7.2) containing 20% (w/w) Ficoll. They were kept in the dark with vigorous shaking for 20 min before measurements. Fluorescence was excited with a green LED (520 nm) and measured in the near far red. Fv/Fm was calculated as the average over 6 experiments at various light intensities. PSII and PSI ratios were evaluated from the amplitude of the ECS signal measured in the presence or

absence of the PSII inhibitors DCMU (20 µM) and hydroxylamine (1 mM) (Joliot and Delosme 1974).

Inverse PCR

100 ng genomic DNA, digested with *Pst*I or *Nhe*I and purified by the phenol/chloroform and NaAc-ETOH precipitation, was self-ligated by T4 DNA Ligase in 100 μL reactions overnight at 16℃. Ligation products were purified and used as templates for nested PCR reactions with primers binding on the pBC1 vector (Table S1). The obtained products were sequenced using primer P3.

Quantitative real time PCR (qRT-PCR) of small RNAs

Small RNAs (20- to 30-nt) were isolated from total RNA on a 15% polyacrylamide gel, polyadenylated and retro-transcribed using the miRNA 1st-Strand cDNA Synthesis Kit (Agilent). The assay involves addition of a poly-A tail using Poly-A Polymerase, reverse transcription with a primer annealing to this tail and qPCR between a generic and a gene-specific primer (here petG_5'_F, corresponding to the footprint to which a CGCAC 5' tail was added to reach the desired melting temperature). qRT-PCR was performed on the Biorad CFX96 instrument using the SsoAdvanced™ universal SYBR® Green supermix (Biorad). All reactions were run in duplicate in three independent assays. To standardize input, we amplified in parallel an abundant miRNA, cre-miR1157-3p (Molnar *et al.* 2007), whose sequence was retrieved from the miRBase database v21 (http://www.mirbase.org/index.shtml). Relative expression levels were calculated using the delta-delta Cq (ΔΔCq) method based on PCR efficiency (E) described by (Pfaffl 2001). The efficiency for each set of primers used was determined by generating a standard curve with 1:5 serial dilution of cDNA. The Cq values of the diluted samples were plotted against the log of each sample dilution and E was calculated from the slope of the regression line according to the formula $E = 10^{[-1/\text{slope}]}$.

Nucleic acid analysis

Total RNA from *Chlamydomonas* was extracted by using the TRI reagent (Sigma). For Northern blot, total RNA was separated on denaturing formaldehyde agarose gels (1% or 2%), then transferred to Roti Nylon$^+$ membrane (Roth), followed by UV light cross-linking (UV Crosslinker, UVC 500, Hoefer). Dig-labeled probes were synthesized by PCR using dig-dNTP. Hybridizations and detection of dig-labeled probes were as described by (Sambrook and Russel 2001). RLM-RACE was performed using the Generacer kit (Invitrogen) according to the manufacturer's instruction, with primer Cr petG RLM rev.

Immunoblot Analysis

Exponentially growing *Chlamydomonas* cells were pelleted, resuspended in lysis buffer which was 20 mM KCl, 20 mM tricine pH 7.8, 0.4 mM EDTA, 5 mM β-mercaptoethanol, 0.2% -1% Triton×100 for cell-wall less strains and 200 mM Tris-Cl, pH 8.0, 150 mM NaCl, 50 mM $MgCl_2$, 20 mM EDTA, 0.02 g/mL Nonidet P40, 5 mM β-mercaptoethanol for cell-walled strains. Cells were centrifuged again for 2 min in a microcentrifuge and the supernatant was used. The protein concentration was measured in the supernatant by Bradford Protein Assay (Bradford 1976). SDS-PAGE, protein gel blotting and immunodetection were performed as described by (Sambrook and Russel 2001)).

2D Blue Native-PAGE

Exponentially growing *Chlamydomonas* cells were pelleted from 20 mL culture by centrifugation at 6,000 g, 10 min, 4°C, and resuspended in 1 mL TMK buffer (10 mM Tris/HCl, pH 6.8; 10 mM $MgCl_2$; 20 mM KCl, with protease inhibitors). To prepare crude thylakoid fractions, cells were broken with glass beads by vortexing 2×60 seconds with an intermediate cooling on ice for at least 2 min, and centrifuged at 3,000 g for 1 min to remove glass beads. The supernatant was collected and centrifuged for 10 min, 20,000 g at 4°C. The pellet was washed twice with 500 μL TMK buffer, and finally resuspended in 500 μL TMK buffer. Membranes corresponding to 25 μg of chlorophyll were resuspended in 60 μL ACA Buffer (750 mM ε-aminocaproic acid, 50 mM Bis-Tris-Cl pH 7.0, 5 mM pH 7.0 EDTA, 50 mM NaCl) and solubilized on ice for 10 min by applying n-dodecyl-β-D-maltoside (β-DM) to a final concentration of 1.5% (w/v). After centrifugation at 16,000 g, 4°C for 20 min, the supernatant was mixed with 1/20 volume of BN loading dye (750 mM ε-aminocaproic acid and 5% Coomassie G 250 (w/v)) and fractionated in a non-denaturing 4.5%-12% BN-PAGE gel (polyacrylamide gel containing 0.5 M ε-aminocaproic acid, 20% glycerol and 50 mM Bis-Tris-Cl pH 7.0). BN-PAGE was carried out overnight with cathode buffer (50 mM Tricine, 15 mM Bis-Tris-Cl pH 7.0 and 0.02 g/ mL Coomassie G 250) and anode buffer (50 mM Bis-Tris-HCl pH 7.0) at a constant voltage of 70 V at 4°C. Gel lanes were denatured in 0.125 M Tris-HCl pH 6.8, 4 g/mL SDS and 1% β-mercaptoethanol for 30 min at room temperature, and resolved on the second dimension gel (15% acrylamide, 5M urea).

In vivo pulse-labelling experiments

Exponentially growing cells ($2 \cdot 10^6$ cells/mL) were pulse-labelled with 35S-sulphate (10 mCi/mL, Hartmann analytic) in the presence of cycloheximide, essentially as described by (Klinkert *et al.* 2006). Membrane-enriched fractions were loaded on an equal-chlorophyll

basis either on denaturing SDS gels (16% acrylamide with 60 g/mL urea) or on native gels as described above.

Acknowledgements

Supporting Information

Figure S1: Southern blot analysis of *mcg1* mutants.

Figure S2: Alignment of OPR repeats in MCG1 and MBI1 of *Chlamydomonas* and *Volvox.*

Figure S3: Fluorescence induction kinetics.

Figure S4: ECS measurement.

Figure S5: Accumulation of transcripts encoding PSII core subunits.

Figure S6: Radioactive *in vivo* labeling of newly synthesized PSII proteins.

Figure S7: Assembly of PSII complexes detected by *in vivo* labelling.

Figure S8: Assembly of cytochrome b_6f complexes in the wild-type and *mcg1-1* mutant.

Figure S9: *psbF* and *psbL* transcripts in *mcg1* mutants

Figure S10: RLM-RACE analysis of *petG* transcript 5' ends in the wild-type and the *mcg1-1* mutant.

Figure S11: Position of PetG (top) and PsbI (bottom) within the cytochrome b_6f and PSII complexes.

Table S1: Primers used in this study

Table S2: Summary of RNAseq experiments displayed on the Phytozome browser

Data S1: Flanking Sequence Tags obtained by inverse PCR

Data S2: Sequence of the MBI1 ortholog in *Chlamydomonas acidophila*

References

Auchincloss, A.H., Zerges, W., Perron, K., Girard-Bascou, J. and Rochaix, J.D. (2002) Characterization of Tbc2, a nucleus-encoded factor specifically required for translation of the chloroplast *psbC* mRNA in *Chlamydomonas reinhardtii*. *The Journal of cell biology*, **157**, 953-962.

Balczun, C., Bunse, A., Hahn, D., Bennoun, P., Nickelsen, J. and Kuck, U. (2005) Two adjacent nuclear genes are required for functional complementation of a chloroplast trans-splicing mutant from *Chlamydomonas reinhardtii*. *The Plant journal : for cell and molecular biology*, **43**, 636-648.

Barkan, A. and Small, I. (2014) Pentatricopeptide repeat proteins in plants. *Annual review of plant biology*, **65**, 415-442.

Barkan, A., Walker, M., Nolasco, M. and Johnson, D. (1994) A nuclear mutation in maize blocks the processing and translation of several chloroplast mRNAs and provides evidence for the differential translation of alternative mRNA forms. *EMBO Journal*, **13**, 3170-3181.

Berthold, D.A., Schmidt, C.L. and Malkin, R. (1995) The deletion of *petG* in *Chlamydomonas reinhardtii* disrupts the cytochrome *bf* complex. *Journal of Biological Chemistry*, **270**, 29293-29298.

Boudreau, E., Nickelsen, J., Lemaire, S.D., Ossenbuhl, F. and Rochaix, J.D. (2000) The Nac2 gene of *Chlamydomonas* encodes a chloroplast TPR-like protein involved in *psbD* mRNA stability. *EMBO J.*, **19**, 3366-3376.

Boulouis, A., Raynaud, C., Bujaldon, S., Aznar, A., Wollman, F.A. and Choquet, Y. (2011) The nucleus-encoded trans-acting factor MCA1 plays a critical role in the regulation of cytochrome *f* synthesis in *Chlamydomonas* chloroplasts. *The Plant cell*, **23**, 333-349.

Boyle, N.R., Page, M.D., Liu, B., Blaby, I.K., Casero, D., Kropat, J., Cokus, S.J., Hong-Hermesdorf, A., Shaw, J., Karpowicz, S.J., Gallaher, S.D., Johnson, S., Benning, C., Pellegrini, M., Grossman, A. and Merchant, S.S. (2012) Three acyltransferases and nitrogen-responsive regulator are implicated in nitrogen starvation-induced triacylglycerol accumulation in Chlamydomonas. *The Journal of biological chemistry*, **287**, 15811-15825.

Bradford, M.M. (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.

Covello, P.S. and Gray, M.W. (1993) On the evolution of RNA editing. *Trends in genetics : TIG*, **9**, 265-268.

Depege, N., Bellafiore, S. and Rochaix, J.D. (2003) Role of chloroplast protein kinase Stt7 in LHCII phosphorylation and state transition in *Chlamydomonas*. *Science*, **299**, 1572-1575.

Dobakova, M., Tichy, M. and Komenda, J. (2007) Role of the PsbI protein in photosystem II assembly and repair in the cyanobacterium *Synechocystis* sp. PCC 6803. *Plant physiology*, **145**, 1681-1691.

Drapier, D., Girard-Bascou, J. and Wollman, F.A. (1992) Evidence for Nuclear Control of the Expression of the *atpA* and *atpB* Chloroplast Genes in *Chlamydomonas*. *The Plant cell*, **4**, 283-295.

Drapier, D., Suzuki, H., Levy, H., Rimbault, B., Kindle, K.L., Stern, D.B. and Wollman, F.A. (1998) The chloroplast *atpA* gene cluster in *Chlamydomonas reinhardtii*. Functional analysis of a polycistronic transcription unit. *Plant physiology*, **117**, 629-641.

Duanmu, D., Casero, D., Dent, R.M., Gallaher, S., Yang, W., Rockwell, N.C., Martin, S.S., Pellegrini, M., Niyogi, K.K., Merchant, S.S., Grossman, A.R. and Lagarias, J.C. (2013) Retrograde bilin signaling enables *Chlamydomonas* greening and phototrophic survival. *Proceedings of the National Academy of Sciences of the United States of America*, **110**, 3621-3626.

Eberhard, S., Loiselay, C., Drapier, D., Bujaldon, S., Girard-Bascou, J., Kuras, R., Choquet, Y. and Wollman, F.A. (2011) Dual functions of the nucleus-encoded factor TDA1 in trapping and translation activation of *atpA* transcripts in *Chlamydomonas reinhardtii* chloroplasts. *The Plant journal : for cell and molecular biology*, **67**, 1055-1066.

Emanuelsson, O., Nielsen, H., Brunak, S. and von Heijne, G. (2000) Predicting subcellular localization of proteins based on their N- terminal amino acid sequence. *Journal of molecular biology*, **300**, 1005-1016.

Fong, S.E. and Surzycki, S.J. (1992) Organization and structure of plastome *psbF, psbL, petG* and *ORF712* genes in *Chlamydomonas reinhardtii. Curr.Genet.*, **21**, 527-530.

Girard-Bascou, J., Choquet, Y., Gumpel, N.J., Culler, D., Purton, S., Merchant, S., Laquerriere, F. and Wollman, F.-A. (1995) Nuclear control of the expression of the chloroplast *pet* genes in *Chlamydomonas reinhardtii*. In *Photosynthesis: from light to biosphere* (Mathis, P. ed. Dordrecht: Kluwer Academic publishers, pp. 683-686.

Hammani, K., Cook, W.B. and Barkan, A. (2012) RNA binding and RNA remodeling activities of the half-a-tetratricopeptide (HAT) protein HCF107 underlie its effects on gene expression. *Proceedings of the National Academy of Sciences of the United States of America*, **109**, 5651-5656.

Houille-Vernes, L., Rappaport, F., Wollman, F.A., Alric, J. and Johnson, X. (2011) Plastid terminal oxidase 2 (PTOX2) is the major oxidase involved in chlororespiration in *Chlamydomonas. Proceedings of the National Academy of Sciences of the United States of America*, **08**, 20820-20825.

Ikeuchi, M., Shukla, V.K., Pakrasi, H.B. and Inoue, Y. (1995) Directed inactivation of the *psbI* gene does not affect photosystem II in the cyanobacterium *Synechocystis* sp. PCC 6803. *Mol.Gen.Genet.*, **249**, 622-628.

Jacobs, J., Marx, C., Kock, V., Reifschneider, O., Franzel, B., Krisp, C., Wolters, D. and Kuck, U. (2013) Identification of a chloroplast ribonucleoprotein complex containing trans-splicing factors, intron RNA, and novel components. *Molecular & cellular proteomics : MCP*, **12**, 1912-1925.

Johnson, X., Kuras, R., Wollman, F.A. and Vallon, O. (2007) Gene hunting by complementation of pooled *Chlamydomonas* mutants. In *14th International Congress of Photosynthesis* (Allen, J. ed. Glasgow, UK: Springer, pp. 1093-1097.

Johnson, X., Wostrikoff, K., Finazzi, G., Kuras, R., Schwarz, C., Bujaldon, S., Nickelsen, J., Stern, D.B., Wollman, F.A. and Vallon, O. (2010) MRL1, a Conserved Pentatricopeptide Repeat Protein, Is Required for Stabilization of rbcL mRNA in *Chlamydomonas* and *Arabidopsis. The Plant cell*, **22**, 234-248.

Joliot, P. and Delosme, R. (1974) Flash-induced 519 nm absorption change in green algae. *Biochimica et biophysica acta*, **357**, 267-284.

Joliot, P. and Joliot, A. (1985) Slow electronic phase and intersystem electron transfer in algae. *Biochimica et biophysica acta*, **806**, 398-409.

Kawakami, K., Umena, Y., Iwai, M., Kawabata, Y., Ikeuchi, M., Kamiya, N. and Shen, J.R. (2011) Roles of PsbI and PsbM in photosystem II dimer formation and stability studied by deletion mutagenesis and X-ray crystallography. *Biochimica et biophysica acta*, **1807**, 319-325.

Kleinknecht, L., Wang, F., Stube, R., Philippar, K., Nickelsen, J. and Bohne, A.V. (2014) RAP, the Sole Octotricopeptide Repeat Protein in *Arabidopsis*, Is Required for Chloroplast 16S rRNA Maturation. *The Plant cell*, **26**, 777-787.

Klinkert, B., Elles, I. and Nickelsen, J. (2006) Translation of chloroplast *psbD* mRNA in *Chlamydomonas* is controlled by a secondary RNA structure blocking the AUG start codon. *Nucleic acids research*, **34**, 386-394.

Künstner, P., Guardiola, A., Takahashi, Y. and Rochaix, J.D. (1995) A mutant strain of *Chlamydomonas reinhardtii* lacking the chloroplast photosystem II *psbI* gene grows photoautotrophically. *Journal of Biological Chemistry*, **270**, 9651-9654.

Kuras, R., Saint-Marcoux, D., Wollman, F.A. and de Vitry, C. (2007) A specific c-type cytochrome maturation system is required for oxygenic photosynthesis. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 9906-9910.

Lefebvre-Legendre, L., Merendino, L., Rivier, C. and Goldschmidt-Clermont, M. (2014) On the complexity of chloroplast RNA metabolism: psaA trans-splicing can be bypassed in chlamydomonas. *Mol Biol Evol*, **31**, 2697-2707.

Lemaire, C. and Wollman, F.A. (1989) The chloroplast ATP synthase in *Chlamydomonas reinhardtii*. II. Biochemical studies on its biogenesis using mutants defective in photophosphorylation. *Journal of Biological Chemistry*, **264**, 10235-10242.

Liu, X. and Gorovsky, M.A. (1993) Mapping the 5' and 3' ends of *Tetrahymena thermophila* mRNAs using RNA ligase mediated amplification of cDNA ends (RLM-RACE). *Nucleic acids research*, **21**, 4954-4960.

Loiselay, C., Gumpel, N.J., Girard-Bascou, J., Watson, A.T., Purton, S., Wollman, F.A. and Choquet, Y. (2008) Molecular identification and function of cis- and trans-acting determinants for *petA* transcript stability in *Chlamydomonas reinhardtii* chloroplasts. *Molecular and cellular biology*, **28**, 5529-5542.

Loizeau, K., Qu, Y., Depp, S., Fiechter, V., Ruwe, H., Lefebvre-Legendre, L., Schmitz-Linneweber, C. and Goldschmidt-Clermont, M. (2014) Small RNAs reveal two target sites of the RNA-maturation factor Mbb1 in the chloroplast of *Chlamydomonas*. *Nucleic acids research*, **42**, 3286-3297.

Lukes, J., Archibald, J.M., Keeling, P.J., Doolittle, W.F. and Gray, M.W. (2011) How a neutral evolutionary ratchet can build cellular complexity. *IUBMB life*, **63**, 528-537.

Merendino, L., Perron, K., Rahire, M., Howald, I., Rochaix, J.D. and Goldschmidt-Clermont, M. (2006) A novel multifunctional factor involved in trans-splicing of chloroplast introns in *Chlamydomonas*. *Nucleic acids research*, **34**, 262-274.

Miller, R., Wu, G., Deshpande, R.R., Vieler, A., Gartner, K., Li, X., Moellering, E.R., Zauner, S., Cornish, A.J., Liu, B., Bullard, B., Sears, B.B., Kuo, M.H., Hegg, E.L., Shachar-Hill, Y., Shiu, S.H. and Benning, C. (2010) Changes in transcript abundance in *Chlamydomonas reinhardtii* following nitrogen deprivation predict diversion of metabolism. *Plant physiology*, **154**, 1737-1752.

Molnar, A., Schwach, F., Studholme, D.J., Thuenemann, E.C. and Baulcombe, D.C. (2007) miRNAs control gene expression in the single-cell alga *Chlamydomonas reinhardtii*. *Nature*.

Murakami, S., Kuehnle, K. and Stern, D.B. (2005) A spontaneous tRNA suppressor of a mutation in the *Chlamydomonas reinhardtii* nuclear *MCD1* gene required for stability of the chloroplast *petD* mRNA. *Nucleic acids research*, **33**, 3372-3380.

O'Connor, H.E., Ruffle, S.V., Cain, A.J., Deak, Z., Vass, I., Nugent, J.H. and Purton, S. (1998) The 9-kDa phosphoprotein of photosystem II. Generation and characterisation of *Chlamydomonas* mutants lacking PSII-H and a site- directed mutant lacking the phosphorylation site. *Biochimica et biophysica acta*, **1364**, 63-72.

O'Toole, N., Hattori, M., Andres, C., Iida, K., Lurin, C., Schmitz-Linneweber, C., Sugita, M. and Small, I. (2008) On the expansion of the pentatricopeptide repeat gene family in plants. *Mol Biol Evol*, **25**, 1120-1128.

Pakrasi, H.B., Diner, B.A., Williams, J. and Arntzen, C.J. (1989) Deletion Mutagenesis of the Cytochrome b_{559} Protein Inactivates the Reaction Center of Photosystem II. *The Plant cell*, **1**, 591-597.

Pfaffl, M.W. (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic acids research*, **29**, e45.

Pfalz, J., Bayraktar, O.A., Prikryl, J. and Barkan, A. (2009) Site-specific binding of a PPR protein defines and stabilizes 5' and 3' mRNA termini in chloroplasts. *The EMBO journal*, **28**, 2042-2052.

Prikryl, J., Rojas, M., Schuster, G. and Barkan, A. (2011) Mechanism of RNA stabilization and translational activation by a pentatricopeptide repeat protein. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 415-420.

Rahire, M., Laroche, F., Cerutti, L. and Rochaix, J.D. (2012) Identification of an OPR protein involved in the translation initiation of the PsaB subunit of photosystem I. *The Plant journal : for cell and molecular biology*, **72**, 652-661.

Raynaud, C., Loiselay, C., Wostrikoff, K., Kuras, R., Girard-Bascou, J., Wollman, F.A. and Choquet, Y. (2007) Evidence for regulatory function of nucleus-encoded factors on mRNA stabilization and translation in the chloroplast. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 9093-9098.

Rochaix, J.D., Perron, K., Dauvillee, D., Laroche, F., Takahashi, Y. and Goldschmidt-Clermont, M. (2004) Post-transcriptional steps involved in the assembly of photosystem I in *Chlamydomonas*. *Biochemical Society transactions*, **32**, 567-570.

Rolland, N., Dorne, A.J., Amoroso, G., Sultemeyer, D.F., Joyard, J. and Rochaix, J.D. (1997) Disruption of the plastid *ycf10* open reading frame affects uptake of inorganic carbon in the chloroplast of *Chlamydomonas*. *EMBO Journal*, **16**, 6713-6726.

Ruwe, H. and Schmitz-Linneweber, C. (2012) Short non-coding RNA fragments accumulating in chloroplasts: footprints of RNA binding proteins? *Nucleic acids research*, **40**, 3106-3116.

Sambrook, J. and Russel, D.W. (2001) *Molecular Cloning: A Laboratory Manual* Third edition edn.: Cold Spring Harbor Laboratory Press.

Schmitz-Linneweber, C., Williams-Carrier, R. and Barkan, A. (2005) RNA immunoprecipitation and microarray analysis show a chloroplast Pentatricopeptide repeat protein to be associated with the 5' region of mRNAs whose translation it activates. *The Plant cell*, **17**, 2791-2804.

Schneider, D., Volkmer, T. and Rogner, M. (2007) PetG and PetN, but not PetL, are essential subunits of the cytochrome b_6f complex from *Synechocystis* PCC 6803. *Research in microbiology*, **158**, 45-50.

Schroda, M., Beck, C.F. and Vallon, O. (2002) Sequence elements within an HSP70 promoter counteract transcriptional transgene silencing in *Chlamydomonas*. *The Plant journal : for cell and molecular biology*, **31**, 445-455.

Schwenkert, S., Legen, J., Takami, T., Shikanai, T., Herrmann, R.G. and Meurer, J. (2007) Role of the low-molecular-weight subunits PetL, PetG, and PetN in assembly, stability, and dimerization of the cytochrome b_6f complex in tobacco. *Plant physiology*, **144**, 1924-1935.

Schwenkert, S., Umate, P., Dal Bosco, C., Volz, S., Mlcochova, L., Zoryan, M., Eichacker, L.A., Ohad, I., Herrmann, R.G. and Meurer, J. (2006) PsbI affects the stability,

function, and phosphorylation patterns of photosystem II assemblies in tobacco. *The Journal of biological chemistry*, **281**, 34227-34238.

Somanchi, A., Barnes, D. and Mayfield, S.P. (2005) A nuclear gene of *Chlamydomonas reinhardtii*, Tba1, encodes a putative oxidoreductase required for translation of the chloroplast psbA mRNA. *The Plant journal : for cell and molecular biology*, **42**, 341-352.

Stoltzfus, A. (1999) On the possibility of constructive neutral evolution. *Journal of molecular evolution*, **49**, 169-181.

Stroebel, D., Choquet, Y., Popot, J.L. and Picot, D. (2003) An atypical haem in the cytochrome b(6)f complex. *Nature*, **426**, 413-418.

Swiatek, M., Regel, R.E., Meurer, J., Wanner, G., Pakrasi, H.B., Ohad, I. and Herrmann, R.G. (2003) Effects of selective inactivation of individual genes for low-molecular-mass subunits on the assembly of photosystem II, as revealed by chloroplast transformation: the *psbEFLJ* operon in *Nicotiana tabacum. Molecular genetics and genomics : MGG*, **268**, 699-710.

Takahashi, Y., Rahire, M., Breyton, C., Popot, J.L., Joliot, P. and Rochaix, J.D. (1996) The chloroplast *ycf7* (*petL*) open reading frame of *Chlamydomonas reinhardtii* encodes a small functionally important subunit of the cytochrome b_6f complex. *The EMBO journal*, **15**, 3498-3506.

Tardif, M., Atteia, A., Specht, M., Cogne, G., Rolland, N., Brugiere, S., Hippler, M., Ferro, M., Bruley, C., Peltier, G., Vallon, O. and Cournac, L. (2012) PredAlgo: a new subcellular localization prediction tool dedicated to green algae. *Mol Biol Evol*, **29**, 3625-3639.

Tourasse, N.J., Choquet, Y. and Vallon, O. (2013) PPR proteins of green algae. *RNA biology*, **10**, 1526-1542.

Urzica, E.I., Adler, L.N., Page, M.D., Linster, C.L., Arbing, M.A., Casero, D., Pellegrini, M., Merchant, S.S. and Clarke, S.G. (2012) Impact of oxidative stress on ascorbate biosynthesis in *Chlamydomonas* via regulation of the VTC2 gene encoding a GDP-L-galactose phosphorylase. *The Journal of biological chemistry*, **287**, 14234-14245.

Vaistij, F.E., Boudreau, E., Lemaire, S.D., Goldschmidt-Clermont, M. and Rochaix, J.D. (2000) Characterization of Mbb1, a nucleus-encoded tetratricopeptide-like repeat protein required for expression of the chloroplast *psbB*/*psbT*/*psbH* gene cluster in *Chlamydomonas reinhardtii. Proceedings of the National Academy of Sciences of the United States of America*, **97**, 14813-14818.

Wei, L., Derrien, B., Gautier, A., Houille-Vernes, L., Boulouis, A., Saint-Marcoux, D., Malnoe, A., Rappaport, F., de Vitry, C., Vallon, O., Choquet, Y. and Wollman, F.A. (2014) Nitric Oxide-Triggered Remodeling of Chloroplast Bioenergetics and Thylakoid Proteins upon Nitrogen Starvation in *Chlamydomonas reinhardtii. The Plant cell*, **26**, 353-372.

Wostrikoff, K., Choquet, Y., Wollman, F.A. and Girard-Bascou, J. (2001) TCA1, a single nuclear-encoded translational activator specific for *petA* mRNA in *Chlamydomonas reinhardtii* chloroplast. *Genetics*, **159**, 119-132.

Zhelyazkova, P., Hammani, K., Rojas, M., Voelker, R., Vargas-Suarez, M., Borner, T. and Barkan, A. (2012) Protein-mediated protection as the predominant mechanism for defining processed mRNA termini in land plant chloroplasts. *Nucleic acids research*, **40**, 3092-3105.

Table 1. Photosynthetic Parameters of *Chlamydomonas* wild-type, *mbi1* and *mcg1-1* cells grown in low light

strain	Fv/Fm	PSII/PSI	max of slow phase
WT	0.71 +/- 0.03	1.4	1.23
mbi1	0.17 +/- 0.03	0.2	1.27
mcg1-1	0.72 +/- 0.02	1.8	<0.05

Figure legends

Figure 1: The *MCG1* and *MBI1* genes. A) Cosmid 87H6 used for complementation of *mcg1* mutants, with the restriction enzymes used (*Nsi*I upward arrows, *Bst*XI round-end arrows, *Ear*I, *Sma*I, *Dra*III or *Nco*I shown with their initial). Gene structure of *MCG1* (exons as black boxes, UTRs as black arrows; the alternative exon 1 is shown in light grey). The probe used in Fig S1 is shown in black, the *Pci*I sites with downward diamonds. **B)** Insertion site of the *aphVIII* cassette in *MBI1* (R: *RBCS2* promoter; TE: transposable element of unknown length). Right: Southern blot with probe and restriction sites indicated on map (B for *Bam*HI, X for *Xma*I). Bottom: PCR genotyping of phototrophic (ac+) and non-phototrophic (ac-) progeny. **C)** FT-Rep analysis of the proteins, showing the Fourier-transformed pattern match signal and predicted repeats.

Figure 2: Growth tests. WT-Jex4, *mbi1* and complemented strain on Minimum or TAP plates, in high light (200 µE m-2 s-1, HL) or low light (30 µE m$^-$2 s^{-1}, LL).

Figure 3: Western blot analysis. A) *mcg1* mutants (underlined) probed with antibodies against cytochrome *f*, cytochrome b_6, Rieske Fe-S protein and subunit IV, with PsaA and β-ATPase as loading controls. Control strains are WT (including a 1/10 load, both from the same blot and exposure as the mutants), *mca1-7* lacking petA mRNA and *pcy1-2* lacking plastocyanin **B)** Immunodetection of D2 in *mbi1* compared to a dilution series of WT-Jex4

and *nac2-26* which completely lacks the *psbD* mRNA. PsaA, RbcL and cytochrome b_6 serve as loading controls.

Figure 4. Assembly of PSII complexes. Detection of PSII assembly complexes by immunoblot analysis of 2D BN/SDS gels with antibodies against D2 and CP43. The position of PSII supercomplexes (SC), dimer, monomer, and the CP43-less RC47 complex is indicated.

Figure 5. Transcript accumulation for chloroplast-encoded small PSII subunits. Five µg of total RNA from WT-Jex4 and *mbi1* were hybridized with probes specific for *psbI, psbK, psbL, psbM, psbT, psbZ,* respectively. The *rbcL* transcript and the ethidium bromide-stained gel (EtBr) served as a loading control. The bands shown are the mature monocistronic transcripts.

Figure 6. Expression of the *atpA* gene cluster is affected in *mbi1*.
A) Map of the cluster. Bent arrows indicate promoters. Transcript numbering and sizes are as in (Drapier*, et al.* 1998). **B)** Northern blot analysis. Five µg RNA, hybridized with the probes A-E shown in A. The EtBr-stained gels showing rRNA serve as loading controls.

Figure 7. Accumulation of transcripts encoding small subunits of cytochrome *b6f*.
Three µg RNA from wild-type and *mcg1-1*, *mcg1-2* and *mcg1-4*, hybridized with probes specific for *petG* and *petL*. The EtBr-stained gel served as a loading control.
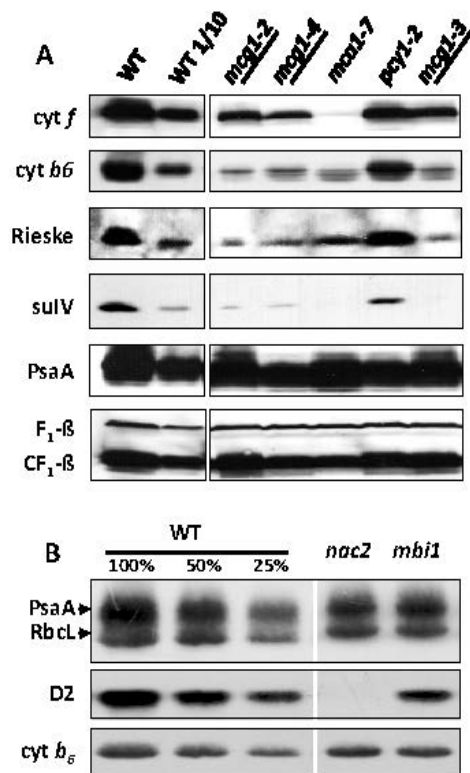
Figure 8: plot of *MCG1* (left) and *MBI1* (right) expression level. Extracted from the Phytozome website. For experiment SRP014795 (Duanmu*, et al.* 2013), WT and a *hmox1* mutant were subjected to a shift from dark (D) to light (L) for 30 min at 150 $\mu E \cdot m^{-2} \cdot s^{-1}$ with (+) or without (-) 0.1 mM biliverdin. For experiment SRP010084 (Urzica*, et al.* 2012), the time of treatment with H_2O_2 (h) is indicated. Expression levels are indicated as fragments per kilobase per million mapped reads (FPKM).
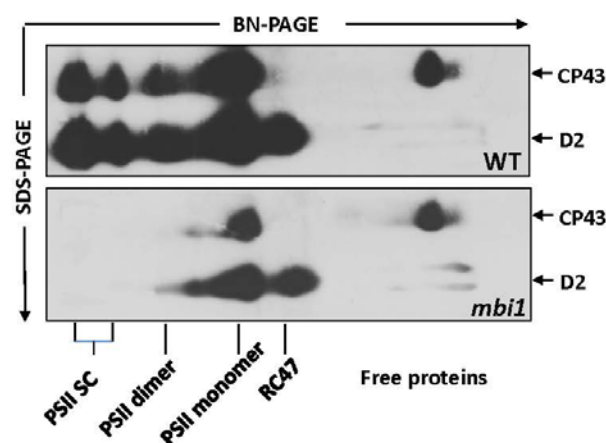
Figure 1: The *MCG1* and *MBI1* genes. A) Cosmid 87H6 used for complementation of *mcg1* mutants, with the restriction enzymes used (*Nsi*I upward arrows, *Bst*XI round-end arrows, *Ear*I, *Sma*I, *Dra*III or *Nco*I shown with their initial). Gene structure of *MCG1*(exons as black boxes, UTRs as black arrows; the alternative exon 1 is shown in light grey). The probe used in Fig S1 is shown in black, the *Pci*I sites with downward diamonds. **B)** Insertion site of the *aphVIII* cassette in *MBI1* (R: *RBCS2* promoter; TE: transposable element of unknown length). Right: Southern blot with probe and restriction sites indicated on map (B for *Bam*HI, X for *Xma*I). Bottom: PCR genotyping of phototrophic (ac+) and non-phototrophic (ac-) progeny. **C)** FT-Rep analysis of the proteins, showing the Fourier-transformed pattern match signal and predicted repeats.

Figure 2: Growth tests. WT-Jex4, *mbi1* and complemented strain on Minimum or TAP plates, in high light (200 µE m-2 s-1, HL) or low light (30 µE m-2 s-1, LL).

Figure 3: Western blot analysis. A) *mcg1* mutants (underlined) probed with antibodies against cytochrome *f*, cytochrome *b₆*, Rieske Fe-S protein and subunit IV, with PsaA and β-ATPase as loading controls. Control strains are WT (including a 1/10 load, both from the same blot and exposure as the mutants), *mca1-7* lacking petA mRNA and *pcy1-2* lacking plastocyanin **B)** Immunodetection of D2 in *mbi1* compared to a dilution series of WT-Jex4 and *nac2-26* which completely lacks the *psbD* mRNA. PsaA, RbcL and cytochrome *b₆* serve as loading controls.

Figure 4. Assembly of PSII complexes. Detection of PSII assembly complexes by immunoblot analysis of 2D BN/SDS gels with antibodies against D2 and CP43. The position of PSII supercomplexes (SC), dimer, monomer , and the CP43-less RC47 complex is indicated.
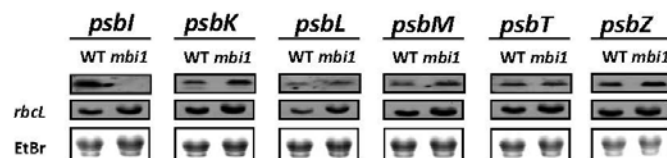
Figure 5. Transcript accumulation for chloroplast-encoded small PSII subunits. 5 µg of total RNA from WT-Jex4 and *mbi1* were hybridized with probes specific for *psbl, psbK, psbL, psbM, psbT, psbZ,* respectively. *rbcL* transcripts and the ethidium bromide-stained gel (EtBr) served as a loading control. The bands shown are the mature monocistronic transcripts.
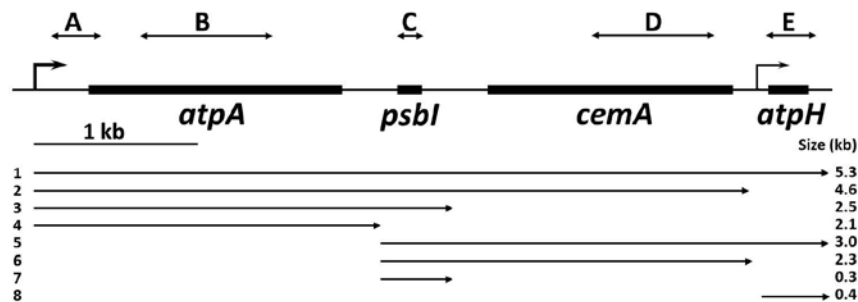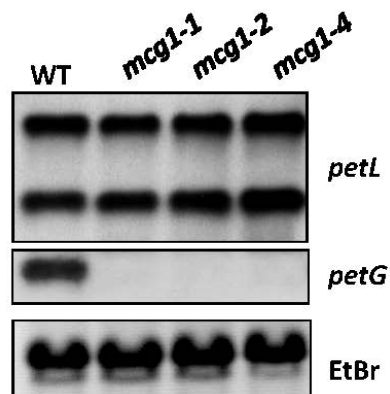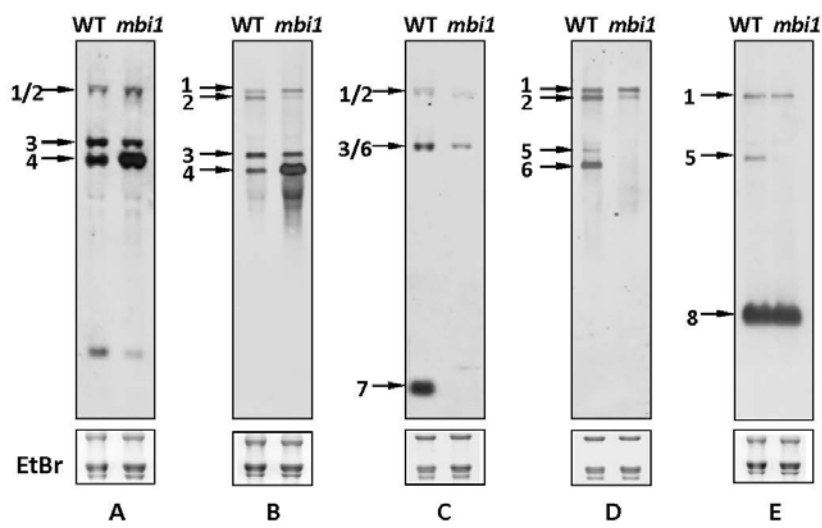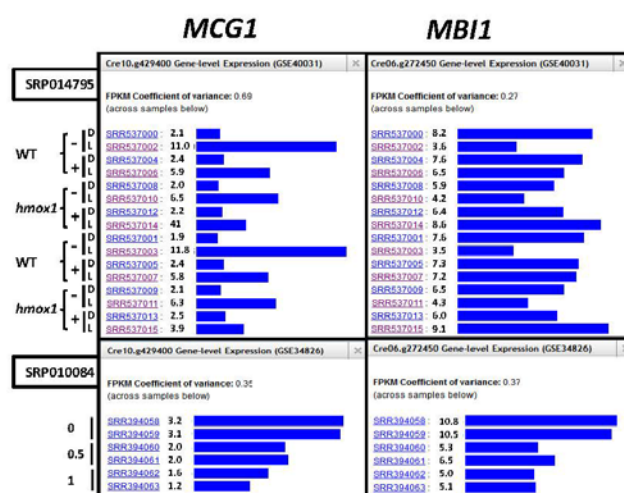
Figure 6A

Figure 6. Expression of the *atpA* gene cluster is affected in *mbi1*. A) Map of the cluster. Bent arrows indicate promoters. Transcript numbering and sizes are as in (Drapier et al., 1998). **B)** Northern blot analysis: 5 µg RNA were hybridized with the probes A-E shown in A. The EtBr-stained gels showing rRNA serve as loading controls.

Figure 6B

Figure 7. Accumulation of transcripts encoding small subunits of cytochrome *b6f*.
Three µg RNA from wild-type and *mcg1-1*, *mcg1-2* and *mcg1-4*, hybridized with probes specific for *petG* and *petL*. The EtBr-stained gel served as a loading control.

Figure 8: plot of *MCG1* (left) and *MBI1* (right) expression level. Extracted from the Phytozome website. For experiment SRP014795 (Duanmu, *et al.* 2013), WT and a *hmox1* mutant were subjected to a shift from dark (D) to light (L) for 30 min at 150 μE·m^{-2}·s^{-1} with (+) or without (-) 0.1 mM biliverdin. For experiment SRP010084 (Urzica, *et al.* 2012), the time of treatment with H_2O_2 (h) is indicated. Expression levels are indicated as fragments per kilobase per million mapped reads (FPKM).