Intracompartmental and Intercompartmental Transcriptional Networks Coordinate the Expression of Genes for Organellar Functions[1][W]

Dario Leister[2], Xi Wang[2], Georg Haberer[2], Klaus F.X. Mayer, and Tatjana Kleine*

Lehrstuhl für Molekularbiologie der Pflanzen (Botanik), Department Biologie I, Ludwig-Maximilians-Universität München, D–82152 Planegg-Martinsried, Germany (D.L., T.K.); and Munich Information Center for Protein Sequences/Institute of Bioinformatics and Systems Biology, Helmholtz Center Munich, D–85764 Neuherberg, Germany (X.W., G.H., K.F.X.M.)

Genes for mitochondrial and chloroplast proteins are distributed between the nuclear and organellar genomes. Organelle biogenesis and metabolism, therefore, require appropriate coordination of gene expression in the different compartments to ensure efficient synthesis of essential multiprotein complexes of mixed genetic origin. Whereas organelle-to-nucleus signaling influences nuclear gene expression at the transcriptional level, organellar gene expression (OGE) is thought to be primarily regulated posttranscriptionally. Here, we show that intracompartmental and intercompartmental transcriptional networks coordinate the expression of genes for organellar functions. Nearly 1,300 ATH1 microarray-based transcriptional profiles of nuclear and organellar genes for mitochondrial and chloroplast proteins in the model plant Arabidopsis (*Arabidopsis thaliana*) were analyzed. The activity of genes involved in organellar energy production (OEP) or OGE in each of the organelles and in the nucleus is highly coordinated. Intracompartmental networks that link the OEP and OGE gene sets serve to synchronize the expression of nucleus- and organelle-encoded proteins. At a higher regulatory level, coexpression of organellar and nuclear OEP/OGE genes typically modulates chloroplast functions but affects mitochondria only when chloroplast functions are perturbed. Under conditions that induce energy shortage, the intercompartmental coregulation of photosynthesis genes can even override intracompartmental networks. We conclude that dynamic intracompartmental and intercompartmental transcriptional networks for OEP and OGE genes adjust the activity of organelles in response to the cellular energy state and environmental stresses, and we identify candidate cis-elements involved in the transcriptional coregulation of nuclear genes. Regarding the transcriptional regulation of chloroplast genes, novel tentative target genes of σ factors are identified.

In eukaryotes, genetic information is stored in the nucleus and in the organellar genomes of mitochondria and chloroplasts. The organellar genomes are of ancient endosymbiotic origin but are now highly impoverished owing to either gene loss or gene transfer to the nucleus (Rand et al., 2004; Timmis et al., 2004; Kleine et al., 2009a). Therefore the majority of mitochondrial and chloroplast proteins are encoded in the nucleus, synthesized in the cytoplasm, and posttranslationally imported into the organelles (Jarvis, 2008). The residual organellar genomes code for proteins involved in organellar gene expression (OGE) or organellar energy production (OEP; i.e. in the light reactions of photosynthesis in chloroplasts and in the respiratory chain in mitochondria). Hence, organellar multiprotein complexes, such as 70S-type ribosomes, photosystems, and the respiratory chain complexes are actually mosaics of subunits encoded by nuclear and organellar genes. Their correct assembly obviously requires the coordination of OGE and nuclear gene expression (NGE) at different levels (Rodermel and Park, 2003; Beck, 2005; Nott et al., 2006; Pogson et al., 2008; Woodson and Chory, 2008) and is thought to include primarily posttranscriptional and translational mechanisms that provide for direct control of OGE by nuclear genes ("anterograde signaling"; Somanchi and Mayfield, 1999; Barkan and Goldschmidt-Clermont, 2000; Rochaix, 2001; Choquet and Wollman, 2002; Giegé et al., 2005; Stern et al., 2010). Conversely, "retrograde signaling" from the organelle to the nucleus is believed to enable NGE to be modified at the transcriptional level, in accordance with the developmental and metabolic state of the organelles (Nott et al., 2006; Pogson et al., 2008; Kleine et al., 2009b). However, evidence is accumulating that also OGE can be regulated at the transcriptional level. Thus, data derived from transcriptomic and proteomic studies of acclimation responses in photosynthetic eukaryotes showed that the photosynthetic apparatus can rapidly

[1] This work was supported by the Deutsche Forschungsgemeinschaft (grant no. FOR804 to D.L., T.K., and K.F.X.M.).

[2] These authors contributed equally to the article.

* Corresponding author; e-mail tatjana.kleine@lrz.uni-muenchen.de.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantphysiol.org) is: Tatjana Kleine (tatjana.kleine@lrz.uni-muenchen.de).

[W] The online version of this article contains Web-only data.

www.plantphysiol.org/cgi/doi/10.1104/pp.111.177691

adapt to metabolic and light fluctuations also at the transcriptional level (Eberhard et al., 2008). Moreover, cluster analyses of chloroplast genes conducted on transcriptomes from mutants with severe photosynthetic defects or from plants exposed to stresses suggested that the accumulation of plastid gene transcripts is regulated in response to altered states of the chloroplast (Cho et al., 2009). The importance of transcriptional control in the chloroplast is underpinned by work with σ (SIG) factors, which are nucleus encoded and required as chloroplast RNA polymerase transcription factors for chloroplast genes (Schweer et al., 2010; Lerbs-Mache, 2011). Accordingly, the loss of SIG2 and SIG6 delays chloroplast biogenesis (Hanaoka et al., 2003; Ishizaki et al., 2005; Loschelder et al., 2006).

Transcriptome analyses in the model plant Arabidopsis (*Arabidopsis thaliana*) have suggested the existence of different layers of retrograde control over the transcriptional expression of nuclear genes for chloroplast or mitochondrial proteins; here, they are named "nuclear chloroplast (or mitochondrial) genes," in contrast to "organellar chloroplast (or mitochondrial) genes." Retrograde signaling seems to involve two distinct types of mechanisms: (1) a "master switch" that acts in a binary mode to either induce or repress the same large set of genes (Richly et al., 2003; Biehl et al., 2005); and (2) a mechanism that supports the coregulation of nuclear and organellar genes for photosynthesis and chloroplast (cp)OGE components (Biehl et al., 2005; Maclean et al., 2008). The latter mechanism appears to be widely conserved in eukaryotes, because *Drosophila* nuclear genes for mitochondrial respiration and the mitochondrial (mt)OGE machinery share a common sequence element, the "nuclear respiratory gene" element (Sardiello et al., 2005). Thus, in animal mitochondria and plant chloroplasts, the transcriptional coordination of nuclear OEP and OGE gene expression seems to follow similar regulatory principles.

Here, we have analyzed in the model plant Arabidopsis the transcriptional regulation of nuclear and organelle genes for OEP and OGE proteins, together with proteins for chloroplast tetrapyrrole biosynthesis. By considering eight different categories of environmental and genetic conditions, focusing particularly on those that are known to influence organelle function or signaling, we identified sets of genes that responded in a coherent manner to various genetic and environmental conditions or perturbations. This enabled us to define coexpression networks within genetic compartments. Strikingly, transcriptional coregulation in nucleus and organelle was found to coordinate the expression of cpOEP and cpOGE genes and to be functionally dominant over intracompartmental coexpression networks under certain conditions. In addition, we identify candidate cis-elements involved in the transcriptional regulation of nuclear genes coding for organellar proteins as well as novel tentative target genes of σ factors.

RESULTS

Selection of Gene Sets

To identify instances of transcriptional coregulation at the intracompartmental and intercompartmental levels, appropriate sets of genes residing in the three genetic compartments in Arabidopsis were selected. The four major sets chosen contained all known genes for chloroplast and mitochondrial proteins (see "Materials and Methods"), located either in the nucleus (cp_n^{all} and mt_n^{all}) or the relevant organelle (cp_{org}^{all} and mt_{org}^{all}). Nine subsets were derived from these four main classes by extracting genes with functions in photosynthesis (cp_n^{photo} and cp_{org}^{photo}), chloroplast tetrapyrrole biosynthesis (cp_n^{tetpy}), mitochondrial respiration (mt_n^{resp} and mt_{org}^{resp}), and OGE (cp_n^{OGE}, cp_{org}^{OGE}, mt_n^{OGE}, and mt_{org}^{OGE}; Table I; Supplemental Table S1). In addition, three control sets of randomly selected nuclear genes were used, consisting of 20 ($ctrl_n^{20}$), 100 ($ctrl_n^{100}$), and 1,500 ($ctrl_n^{1500}$) members and equivalent in size to the corresponding nuclear gene sets (which range from 37 to 1,476 genes; Table I).

Selection of Conditions

mRNA expression data derived from plants that carried mutations affecting functions in organelles or at the whole-cell level, or that had been subjected to environmentally induced perturbations, were retrieved from public Arabidopsis databases (see "Materials and Methods"). The perturbations considered and conditions imposed were assigned to eight major categories of environmental or genetic stimuli (Table II; Supplemental Table S2), which are known or assumed to be relevant for chloroplast and/or mitochondrial functions and retrograde signaling.

Chloroplast

A wealth of evidence shows that mutants affecting certain chloroplast functions give rise to changes in the transcription of nuclear and chloroplast genes (for review, see Pfannschmidt, 2003; Pogson et al., 2008). Therefore, we have considered microarray data from mutants for chloroplast proteins, for which effects on gene expression at the transcription level have been reported, including the thylakoid Ser/Thr protein kinase mutant *stn7* (Bonardi et al., 2005; Bräutigam et al., 2009), photosystem I subunits D and E (*psad1* and *psae1*) mutants that affect electron flow through PSI (Pesaresi et al., 2009), and a transposon-tagged albino mutant (Tian et al., 2007).

Genomes Uncoupled Signaling

This category includes microarray data that reveal the effects of treatments and mutations thought to specifically alter plastid-to-nucleus signaling. This includes the application of the herbicide norflurazon

Table I. *Overview of sets and subsets of genes*

The designation of gene sets refers to the subcellular location and function of the gene product and indicates the location of the gene; for example, cp_n^{photo} refers to nuclear (n) genes for chloroplast (cp) proteins with a function in photosynthesis (photo). Accordingly, mt stands for mitochondrion, resp for respiration, and org (organelle) indicates that the gene is located in chloroplasts or mitochondria. A complete list of the genes analyzed is provided in Supplemental Table S1. Note that 96 genes are represented in both cp_n^{all} and mt_n^{all} gene sets, because of dual targeting of their gene products. Also note that for each of the three control sets, $ctrl_n^{20}$, $ctrl_n^{100}$, and $ctrl_n^{1500}$, all nuclear Arabidopsis genes were considered and mean values of 100 repetitions were analyzed.

Designation	Description	No. of Genes
cp_n^{all}	All known nuclear chloroplast genes	1,476
cp_n^{photo}	Subset of cp_n^{all} involved in photosynthesis	129
cp_n^{OGE}	Subset of cp_n^{all} involved in OGE	90
cp_n^{tetpy}	Subset of cp_n^{all} involved in tetrapyrrole biosynthesis	39
cp_{org}^{all}	All organellar chloroplast genes	80
cp_{org}^{photo}	Subset of cp_{org}^{all} involved in photosynthesis	45
cp_{org}^{OGE}	Subset of cp_{org}^{all} involved in OGE	30
mt_n^{all}	All known nuclear mitochondrial genes	1,323
mt_n^{resp}	Subset of mt_n^{all} involved in respiration	110
mt_n^{OGE}	Subset of mt_n^{all} involved in OGE	37
mt_{org}^{all}	All organellar mitochondrial genes	98
mt_{org}^{resp}	Subset of mt_{org}^{all} involved in respiration	18
mt_{org}^{OGE}	Subset of mt_{org}^{all} involved in OGE	7
$ctrl_n^{20}$	20 randomly selected nuclear genes	20
$ctrl_n^{100}$	100 randomly selected nuclear genes	100
$ctrl_n^{1500}$	1,500 randomly selected nuclear genes	1,500

(which damages plastids and induces photobleaching) or lincomycin (an inhibitor of OGE); both agents prevent the expression of nuclear genes that encode photosynthesis-related proteins (Oelmüller and Mohr, 1986; Gray et al., 2003). In the *genomes uncoupled* (*gun*) mutants, the expression of several nuclear photosynthesis genes is maintained in the presence of norflurazon and lincomycin (Susek et al., 1993; Gray et al., 2003).

Reactive Oxygen Species

The pool of reactive oxygen species (ROS), such as singlet oxygen, superoxide, and hydrogen peroxide, is also thought to provide signals relevant for organelle-to-nucleus communication (Apel and Hirt, 2004; Pogson et al., 2008; Kleine et al., 2009b). Therefore, microarray data that reflect the consequences of increased ROS accumulation induced either by the application of an inhibitor (methyl viologen to steadily produce superoxide in the light; Mehler, 1951) or by down-regulation of peroxisomal catalase, a ROS scavenger, were considered. In the latter case, data were collected from plants grown under ambient growth conditions and after exposure to high light, which exacerbates the effects of defective ROS scavenging (Vanderauwera et al., 2005). Because the alternative oxidase (AOX) of plant mitochondria prevents ROS formation from an overreduced ubiquinone pool, microarray data from a study that showed that a decrease in AOX activity induces oxidative stress in chloroplasts (Umbach et al., 2005) were integrated into this category. Moreover, mRNA profiles from mutants for transcription factors that play a central role in ROS and abiotic stress signaling in Arabidopsis, like the zinc finger transcription factors ZAT10 (Rossel et al., 2007) and ZAT12 (Davletova et al., 2005), were included. Because exposure of the Arabidopsis *fluorescent* mutant to light generates singlet oxygen in plastids (Meskauskiene et al., 2001), an effect that provided the basis for analyses that led to the conclusion that hydrogen peroxide antagonizes singlet oxygen-mediated signaling and that the corresponding signaling pathways interact (op den Camp et al., 2003; Laloi et al., 2007), the corresponding microarray data were assigned to this category also.

Light Signaling

Light regulates chloroplast function and development by promoting the expression of nuclear chloroplast genes (Ohgishi et al., 2004). The expression of nuclear photosynthesis genes is induced by red/far-red and blue light photoreceptors: the phytochromes and cryptochromes, respectively (Ohgishi et al., 2004; Larkin and Ruckle, 2008). Light and plastid signaling are closely connected (Larkin and Ruckle, 2008). Thus, *long hypocotyl1* (*hy1*) and *hy2*, as well as *cryptochrome1* (*cry1*) mutant alleles, were isolated in *gun* mutant screens (Mochizuki et al., 2001; Ruckle et al., 2007), and plastid signals influencing photomorphogenesis are dependent on GUN1 and cry1 (Ruckle and Larkin, 2009). Consequently, microarray data from plants subjected to different light treatments, from mutants defective either in photoreceptors or their downstream transcription factors, like HY5 and HY5 homolog, and mutants displaying defects in photomorphogenesis or deregulation of nuclear photosynthesis genes in certain light conditions, were considered in this category.

Table II. *Overview of conditions*

"All" refers to the combination of all eight categories. A complete list of conditions is given in Supplemental Table S2. Data sets refer to different experimental series that contained up to 60 array experiments.

Category of Condition	Data Sets	Arrays
Chloroplast	3	25
gun signaling	5	38
Light signaling	18	247
Hormones	25	349
ROS	14	104
General stresses	15	312
Sugars	13	139
Nutrient supply	8	76
All	101	1,290

Hormones

Multiple interactions between the retrograde signaling pathway and hormone signaling, especially abscisic acid (ABA) signaling, have been described (Penfield et al., 2006; Shen et al., 2006; Koussevitzky et al., 2007; Kim et al., 2009). In contrast to brassinosteroids and gibberellins, which repress photomorphogenesis and negatively regulate chloroplast development (Li and Chory, 1999; Alabadí and Blázquez, 2009), cytokinins promote the etioplast-to-chloroplast transition and the formation of the electron transport chain (Cherniad'ev, 2000) and control the expression of nuclear chloroplast genes (Schmülling et al., 1997). In response to ethylene, nuclear photosynthesis genes are generally down-regulated. Moreover, ethylene signaling appears to interact with auxin-related signal transduction processes (Zhong and Burns, 2003). Given the important role of ABA, auxin, ethylene, gibberellins, brassinosteroids, and cytokinins in regulating the expression of nuclear chloroplast genes, a collection of microarray data from hormone-treated plants and hormone-related mutants were included in this category.

Sugars

In this category, microarray data for plants treated with Suc or Glc, for instance, and of mutants affected in carbohydrate metabolism or subcellular partitioning, like *hexokinase1* (*hxk1*) and *triosephosphate translocator* (*tpt*), were investigated. Sugars act as signaling molecules to control the expression of nuclear genes involved in photosynthesis, glyoxylate metabolism, respiration, starch, and Suc synthesis. Increased levels of Glc or Suc, the end products of photosynthesis, repress photosynthetic gene expression (Rolland et al., 2006). HXK1 is crucial for sensing and responding to intracellular Glc signals (Cho et al., 2006). The TPT functions primarily as a dihydroxyacetone phosphate/phosphate exchanger to maintain Suc synthesis in the cytosol (Flügge, 1999). Interestingly, the *tpt* and *gun* mutations provoke similar alterations in the nuclear chloroplast transcriptome (Biehl et al., 2005).

General Stresses and Nutrient Supply

The remaining two conditions include mutations or treatments that have major effects on general cell functions. We included, for instance, microarray data from drought- and cold-stressed plants in the "general stresses" category and data for plants grown under limiting nitrogen and sulfate conditions in the category "nutrient supply," because these latter conditions impinge on photosynthesis and the expression of photosynthesis genes (Jung et al., 2003; Zhang et al., 2004; Peng et al., 2007; Pinheiro and Chaves, 2011).

All

Because the specific stimuli for plastid signaling considered in each of the eight categories are often interconnected, the effects of individual stimuli are sometimes difficult to separate. Thus, interactions between light and hormone signaling pathways exist. For instance, HY5, which was included in the category "light signaling," also acts as a signal integrator that connects the light and hormone pathways (Lau and Deng, 2010). Moreover, singlet oxygen-mediated plastid signaling affects plastid development in seedlings and depends on the recruitment of ABA during seedling development (Kim et al., 2009). As a third example, AOX has been suggested to link several processes in mitochondria and chloroplasts to optimize photosynthetic performance (Strodtkötter et al., 2009). Therefore, to test whether transcriptional responses are specific for certain conditions (or categories), we also performed our analyses with the whole set of microarray data considered for the eight categories.

Overall, data from 101 genome-wide Affymetrix ATH1 oligonucleotide array analyses, encompassing 1,290 independent hybridization experiments (Table II; Supplemental Table S2), were analyzed.

Transcripts of Genes for OEP and OGE Are Generally Highly Abundant

For each microarray experiment, normalized absolute gene expression values were extracted from the downloaded files. To compensate for potential quantitative differences between microarray experiments, for instance differences in absolute hybridization levels, gene expression data were subjected to z-score transformation (see "Materials and Methods"). Positive scores indicate transcript levels above average and negative scores indicate transcript abundances below average for the relevant control ensemble.

Comparison of the mean z-scores, as a measure for absolute gene expression levels, revealed clear differences between the eight different conditions, with "*gun* signaling" and "chloroplast" conditions eliciting the lowest and highest transcript abundances, respectively (Supplemental Fig. S1A). Even more pronounced differences were observed between the different gene sets (Fig. 1A; Supplemental Fig. S1A). Under all conditions considered, organellar mitochondrial genes represented

the least expressed group of genes, always giving slightly negative z-scores, indicative of transcript abundances below average with respect to the total Arabidopsis transcriptome (Fig. 1A; Supplemental Fig. S1A). Organellar chloroplast genes, together with nuclear photosynthetic genes, represented the most highly expressed gene classes. Mean expression levels for the cp_n^{photo} and mt_n^{resp} subsets were 4-fold higher than the means for the corresponding master sets cp_n^{all} and mt_n^{all}, respectively. Similarly, mt_n^{OGE} genes were expressed at markedly higher levels than mt_n^{all}, in contrast to cp_n^{OGE} and cp_n^{tetpy} genes, which behaved similarly to cp_n^{all}.

Taken together, these analyses revealed that transcripts encoding chloroplast proteins were generally present in very high abundances, irrespective of whether the corresponding gene is located in the nucleus or the organelle. Moreover, nuclear OEP genes, as well as nuclear mtOGE genes, were also expressed at significantly higher levels than the average for the corresponding master set (all nuclear chloroplast or mitochondrial genes).

Similarities in Expression Responses within Gene Sets

A transcriptional signature common to most nuclear chloroplast genes (master switch) was previously proposed based on the characterization of the nuclear chloroplast transcriptome under 35 environmental and genetic conditions (Richly et al., 2003). To critically evaluate this hypothesis on the basis of a sufficiently comprehensive data set and to test whether also others of our predefined gene sets of nuclear and organellar genes for chloroplast and mitochondrial proteins are specifically coregulated, expression responses within gene sets were investigated. To determine the level of expression similarity within a given gene set, the mean Pearson correlation coefficient (PCC) was calculated from the PCCs (see "Materials and Methods") of all possible gene pairs within the gene set (Fig. 1B; Supplemental Fig. S1B). Once again, the three sets of randomly selected nuclear genes $ctrl_n^{20}$, $ctrl_n^{100}$, and $ctrl_n^{1500}$ and the two master sets of all organellar chloroplast (cp_{org}^{all}) and mitochondrial (mt_{org}^{all}) genes served as controls. As expected, the three nuclear control groups exhibited PCCs close to zero (Fig. 1B), although the conditions "sugars" and "ROS" also elicited some degree of coexpression in randomly selected nuclear genes (PCCs from 0.15 to 0.20; Supplemental Fig. S1B). The organellar genes exhibited the highest degrees of coregulation (cp_{org}^{all}, 0.52; mt_{org}^{all}, 0.53); interestingly, "chloroplast" conditions were associated with a lesser degree of organellar gene coexpression. With the exception of mt_n^{all}, which for most conditions showed only marginally higher PCCs than randomly selected

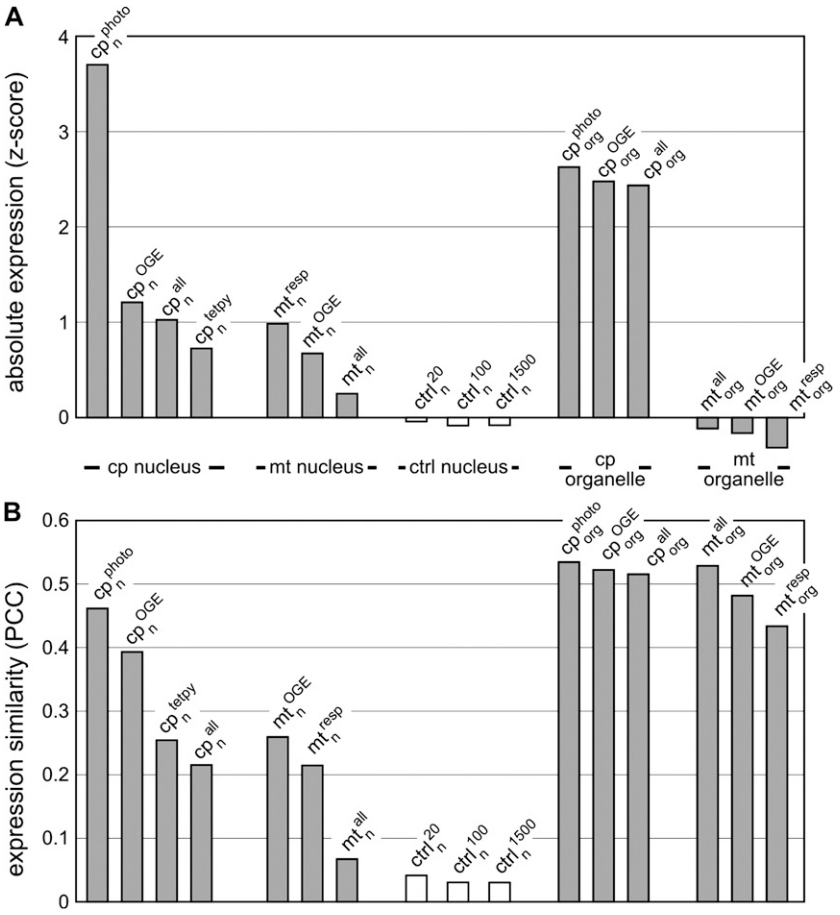


Figure 1. Absolute mRNA expression levels of gene sets and expression similarities. A, Mean z-scores are given over all categories of conditions. SE values were always less than 3%. The designation of gene sets follows Table I. Data for nuclear control groups are given as white bars. B, Mean PCCs are given for each gene set over all conditions. Values between 0 and 1 indicate increasing positive correlation of expression profiles; 0 to –1 stand for increasing negative correlations. All nuclear gene sets show significantly higher expression similarities than do random control sets in Monte Carlo simulations (Supplemental Table S3). For a detailed analysis of the individual categories, see Supplemental Figure S1.

nuclear genes, all sets of nuclear chloroplast or mitochondrial genes showed increased coregulation to varying degrees. Thus, cp_n^{photo} (PCC = 0.46) and cp_n^{OGE} (PCC = 0.39) were the most highly coregulated nuclear gene sets (Fig. 1B), and again "chloroplast" conditions decreased the level of coregulation, as in all other nuclear chloroplast gene sets (Supplemental Fig. S1B). Nuclear mtOGE genes were the most highly coexpressed gene subset among the nuclear mitochondrial genes; here, "chloroplast" conditions provoked the highest level of coexpression (PCC = 0.53), as indeed for nuclear mitochondrial genes as a whole. Monte Carlo simulations confirmed that almost all nuclear gene sets were significantly more highly coregulated than corresponding random control groups (Supplemental Table S3).

Coregulation of Different Gene Sets in the Same Genetic Compartment

Pairwise comparisons of the mRNA expression of genes in different gene sets allowed us to identify expression similarities between gene sets. Whereas already all nuclear chloroplast genes exhibited a certain level of coregulation (see above; Supplemental Fig. S1B), the combinations cp_n^{photo}/cp_n^{OGE}, $cp_n^{photo}/cp_n^{tetpy}$, and cp_n^{OGE}/cp_n^{tetpy} displayed even higher degrees of coregulation over all conditions (Fig. 2A). This implies that coordination of the expression of nucleus-encoded photosynthetic proteins, the expression of chloroplast-encoded photosynthetic proteins via the OGE machinery, and the synthesis of the tetrapyrrole chlorophyll, three processes that are all necessary for efficient assembly of the photosynthetic machinery, is controlled, at least in part, at the level of mRNA expression in the nucleus. For mitochondria, the combination mt_n^{resp}/mt_n^{OGE} also exhibited a certain level of coexpression (Fig. 2A), implying that the coregulation of the transcription of nuclear mtOGE and respiration discovered in animals (Sardiello et al., 2005) also extends to photosynthetic eukaryotes. Strikingly, under "chloroplast" conditions, the coregulation generally observed in the combinations cp_n^{photo}/cp_n^{OGE} and $cp_n^{photo}/cp_n^{tetpy}$ breaks down, whereas the same conditions enhance coregulation of the mt_n^{resp}/mt_n^{OGE} sets (Supplemental Fig. S2A). At the level of organellar genes, $cp_{org}^{OGE}/cp_{org}^{photo}$ or $mt_{org}^{OGE}/mt_{org}^{resp}$ were no more highly coexpressed than the corresponding control sets of all organellar genes (Fig. 2B; Supplemental Fig. S2B). However, their levels of coregulation were still higher than those of any gene class or pair of nuclear gene sets.

We also tested whether coregulation between nuclear chloroplast and mitochondrial genes occurs. Indeed, nuclear OGE genes (cp_n^{OGE}/mt_n^{OGE}) were moderately coregulated under most conditions, while the different OEP genes (cp_n^{photo}/mt_n^{resp}) could even display a moderate negative correlation coefficient (i.e. be regulated in opposite senses; Fig. 2C). For organellar OGE ($cp_{org}^{OGE}/mt_{org}^{OGE}$) and OEP ($mt_{org}^{resp}/cp_{org}^{photo}$) genes, moderate coregulation was evident only under two conditions ("light signaling" and "sugars"; Supplemental Fig. S3A).

Coregulation between Different Compartments

We then compared the expression profiles of gene sets for proteins with similar function that are distributed to different genetic compartments. For mitochondrial proteins, genes for respiration ($mt_n^{resp}/mt_{org}^{resp}$) or OGE ($mt_n^{OGE}/mt_{org}^{OGE}$) were not coexpressed at all, except under "chloroplast" and, less obviously, "ROS" conditions (Fig. 2D; Supplemental Fig. S3B). At the chloroplast level, a modest degree of coregulation was already noted at the level of all nuclear and organellar chloroplast genes ($cp_n^{all}/cp_{org}^{all}$). More pronounced levels of nucleus-organelle coregulation were observed for photosynthesis ($cp_n^{photo}/cp_{org}^{photo}$) and cpOGE ($cp_n^{OGE}/cp_{org}^{OGE}$) genes (Fig. 2D). Here, "general stresses" and "nutrient supply" conditions, which should affect the general energy status of the cell, resulted in maximal intercompartmental coregulation, whereas coregulation was markedly reduced under "chloroplast" conditions (Supplemental Fig. S3B), implying a differentiated transcriptional response to adjust the photosynthetic process.

Therefore, coregulation of nuclear and organellar genes is especially characteristic for chloroplast functions. Only under conditions that perturb chloroplast function or ROS homeostasis does a certain level of intercompartmental coexpression ensue between nuclei and mitochondria.

Dynamics of Transcriptional Networks

The results of our pairwise comparisons indicated that coregulation between different gene sets can vary in different conditions. Therefore, to obtain an overall picture of expression similarities, we performed hierarchical clustering of gene set expression for all eight individual categories of conditions as well as over all conditions (Fig. 3). The coregulation in the combinations cp_n^{tetpy}/cp_n^{OGE} and $mt_{org}^{OGE}/mt_{org}^{resp}$ was very robust and was found in all eight categories. Coexpression in the combinations $cp_{org}^{photo}/cp_{org}^{OGE}$ and mt_n^{resp}/mt_n^{OGE} was found in seven of the eight categories. Nuclear photosynthesis genes were coexpressed with cp_n^{tetpy}/cp_n^{OGE} in six categories (Fig. 3).

The three categories "*gun* signaling," "light signaling," and "sugars" generally increased coexpression among gene sets, as indicated by the shorter branch lengths. Nucleus-organelle coregulation was more robust for chloroplasts (six categories) than for mitochondria (four categories) and was particularly perturbed by "light signaling" and "sugars" conditions, which resulted in organelle-organelle coregulation. In contrast, "general stresses" and "chloroplast" provoked the strongest coregulation of nuclear and chloroplast photosynthesis genes (Fig. 3).

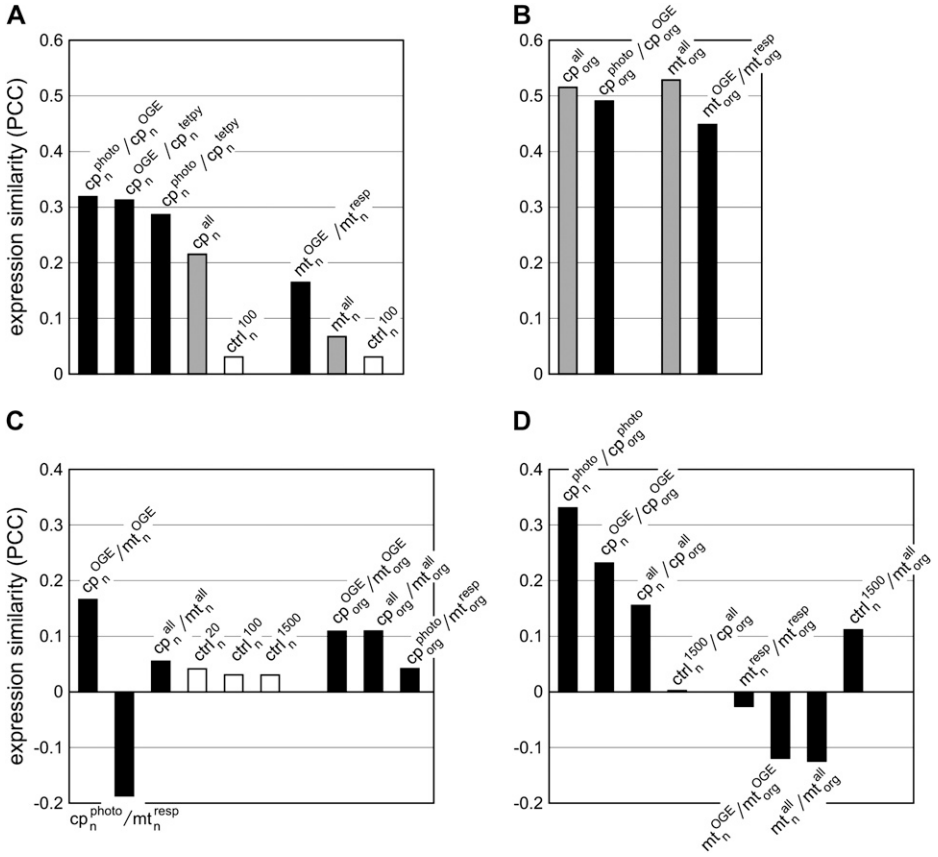


Figure 2. Expression similarities between different gene sets. A and B, Genes in the same genetic compartment and products in the same organelle. C, Genes in the same genetic compartment and products in different organelles. D, Genes in different genetic compartments. For comparison, some expression similarities within gene sets (gray bars) or nuclear control sets (white bars) are provided. A detailed analysis for the eight conditions is provided in Supplemental Figures S2 and S3, and results of Monte Carlo simulations are provided in Supplemental Table S4.

Coregulation between Different Genetic Compartments at the Gene Pair Level

The PCC analysis was further exploited to identify coregulated gene pairs from among gene sets encoded in different compartments that specified proteins with similar functions. This includes (1) as an instance of organelle-organelle coregulation, OGE ($mt^{OGE}_{org}/cp^{OGE}_{org}$) and OEP ($mt^{resp}_{org}/cp^{photo}_{org}$) genes from mitochondria and chloroplasts, which were moderately coregulated under "light signaling" and "sugars" conditions (Supplemental Fig. S3A); (2) as an example for nucleus-organelle coregulation at the mitochondrial level, genes for OGE ($mt^{OGE}_{n}/mt^{OGE}_{org}$) and respiration ($mt^{resp}_{n}/mt^{resp}_{org}$), which were coexpressed under "chloroplast" conditions (Supplemental Fig. S3B); and (3) at the chloroplast level, the pronounced nucleus-organelle coregulation observed for cpOGE ($cp^{OGE}_{n}/cp^{OGE}_{org}$) and photosynthesis ($cp^{photo}_{n}/cp^{photo}_{org}$) genes already under "all" conditions (Fig. 2D).

Therefore, the PCCs of gene pairs of the combinations mentioned above under the respective conditions were calculated (Supplemental Table S5). As a control, all-against-all Pearson correlation matrices from 23,913 Arabidopsis genes were computed for each condition. From these correlation matrices, the background distributions of pairwise expression similarities were determined. Gene pairs exceeding the 90% and 95% quantile of the distribution (corresponding to $P = 0.1$ and $P = 0.05$, respectively) were then considered as "coregulated" and "significantly coregulated," respectively, under the respective conditions. Whereas no instance of negative coregulation was detected among the gene pairs and conditions investigated, 12 cases of significant positive coregulation and 151 cases of positive coregulation were detected (Supplemental Table S5). Representative cases are discussed in the following.

$mt^{OGE}_{org}/cp^{OGE}_{org}$ ("Light Signaling" and "Sugars")

For this combination, the genes encoding mtRPL16 and mtRPS3 are represented by the same Affymetrix array element and are coregulated with two ("light signaling") and four ("sugars") chloroplast genes, respectively, including the gene for cpRPL20 under both conditions (Supplemental Table S5, A and B).

$mt^{resp}_{org}/cp^{photo}_{org}$ ("Light Signaling" and "Sugars")

In this case, the mitochondrial genes encoding ATPase subunit 6, NADH dehydrogenase subunits 3 and 4, and apocytochrome *b* in "light signaling" conditions and the mitochondrial genes encoding the NADH dehydrogenase subunit 4 and ATP synthase subunit 8 in "sugars" conditions were coregulated with chloroplast genes, in particular genes encoding the PSII reaction center proteins D1 and CP43 and the

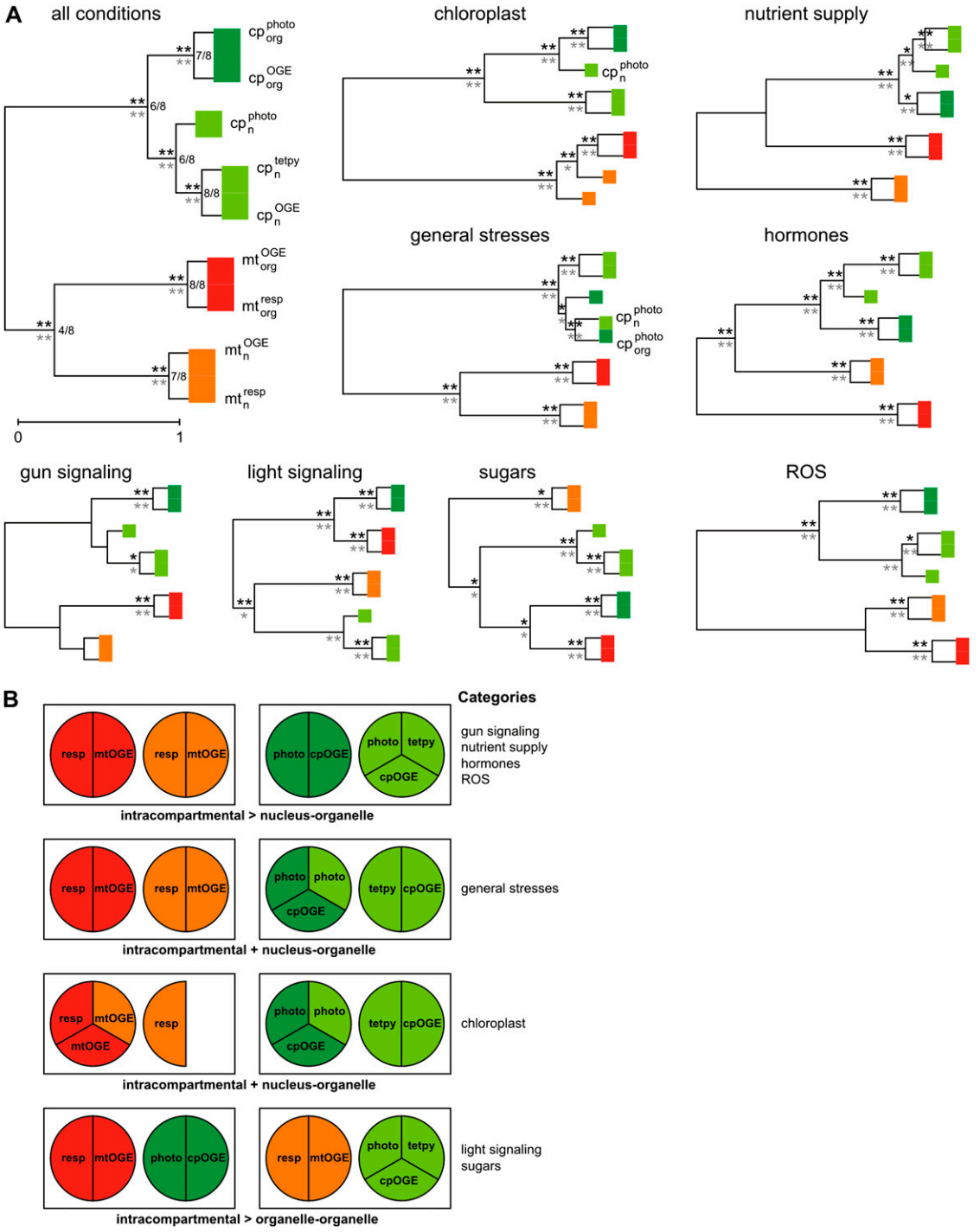


Figure 3. Different layers of transcriptional control regulate the expression of genes for organellar proteins. A, Hierarchical clustering of the expression profiles of gene sets. Results over all conditions and for each condition are provided. The distance between profiles has been calculated as 1 – PCC. The bar indicates a distance unit of 1. Bootstrap values of $P \geq 80\%$ and $P \geq 95\%$ are indicated as one and two asterisks, respectively. Black asterisks stand for approximately unbiased P values and gray asterisks for P values from standard bootstrap resampling. B, Scheme summarizing the results from A. Highly coregulated gene sets (intracompartmental or nucleus-organelle) are symbolized by sectors, and moderately coregulated gene sets (nucleus-organelle or organelle-organelle) are enclosed by boxes. The color code is as in A. x > y indicates that mode x of transcriptional coregulation prevails over mode y.

NAD(P)H dehydrogenase subunits J and K (Supplemental Table S5, C and D).

$mt_n^{OGE}/mt_{org}^{OGE}$ ("Chloroplast")

In this combination, the highest degree of coregulation (PCC = 1; $P = 0$) was found for the gene pairs *AT2G07675*/*ATMG00980* and *AT2G07715*/*ATMG00980* (Supplemental Table S5E). *AT2G07675* encodes a ribosomal protein of the S12/S23 family, *AT2G07715* a nucleic acid-binding protein with an oligonucleotide/oligosaccharide binding-like fold, and *ATMG00980* the ribosomal protein L2 (mtRPL2). In addition, the mitochondrial genes for mtRPS3, mtRPS4, mtRPL5, and mtRPL16 constituted a subset of mt_{org}^{OGE}, which was significantly coregulated (four cases) and coregulated (13 cases) with various genes from mt_n^{OGE}.

$mt_n^{resp}/mt_{org}^{resp}$ ("Chloroplast")

Here, six cases of significant coregulation and 27 cases of coregulation were found (Supplemental Table S5F). In particular, mitochondrial genes encoding ATPase subunits (*atp1* [*ATMG01190*], *atp6-1* [*ATMG00410*], *atp6-2* [*ATMG01170*], and *atp9* [*ATMG01080*]), cytochrome oxidase 1 (*cox1* [*ATMG01360*]), and NADH dehydrogenase subunit 9 (*nad9* [*ATMG00070*]) were coregulated with nuclear genes for ATPase subunits like *atpC* (*AT2G07671*) and *atpG* (*AT4G26210*).

$cp_n^{OGE}/cp_{org}^{OGE}$ ("All")

Only one clear example of nucleus-organelle coregulation of cpOGE genes was found: this involved the nuclear gene *EMBRYO DEFECTIVE2369* (*AT4G04350*) and the chloroplast gene for cpPRPS16 (*ATCG00050*; Supplemental Table S5G).

$cp_n^{photo}/cp_{org}^{photo}$ ("All")

In contrast to the paucity of coregulation between nuclear and chloroplast genes for OGE, 77 pairs of photosynthesis genes were coregulated. For instance, chloroplast genes encoding subunits of the ATP synthase (*atpB* [*ATCG00480*] and *atpE* [*ATCG00470*]) or the cytochrome b_6/f complex (*petB* [*ATCG00720*]) were coregulated with nuclear genes encoding diverse Rubisco small chain subunits (*AT1G67090*, *AT5G38410*, *AT5G38420*, and *AT5G38430*), Rubisco activase (*RCA* [*AT2G39730*]), and several proteins of the light-harvesting complexes of PSI and PSII (Supplemental Table S5H).

A Case Study: Absolute mRNA Expression, Expression Similarities, and Coregulation of Class I, II, and III Chloroplast Genes

In higher plants, chloroplast transcription is performed by three different plastid RNA polymerases: a multimeric plastid-encoded plastid (PEP) and two monomeric nucleus-encoded plastid (NEP) RNA polymerases (Lerbs-Mache, 2011; Liere et al., 2011). In principle, most plastid genes can be transcribed by both types of RNA polymerases. However, a subset of plastid genes, including PSI and PSII genes, are transcribed from PEP promoters only (class I genes), whereas some genes that are related to plastid transcription are transcribed exclusively by NEP (class III genes; Hajdukiewicz et al., 1997). Nonphotosynthetic and housekeeping genes are mostly transcribed by both PEP and NEP (class II genes; Swiatecka-Hagenbruch et al., 2007). In plastids, the specificity of promoter recognition by PEP is achieved by the nucleus-encoded σ factors SIG1 to SIG6.

To obtain insights into the regulation of class I, II, and III genes, the absolute mRNA expression and the expression similarities of chloroplast genes were analyzed (Supplemental Fig. S4). Comparison of the absolute gene expression levels revealed that the expression of class III genes was in general two times lower compared with class I and II and the rest of the chloroplast genes not yet classified (Supplemental Fig. S4A). Generally, "*gun* signaling" and "ROS" provoked the lowest transcript abundances in class I and II and in class III genes, respectively. Whereas the expression similarities within the three different gene classes was high in "all" conditions (PCCs from 0.46 to 0.55; Supplemental Fig. S4B), under "ROS" conditions, class II and III genes displayed low expression similarities, with PCCs around 0.3 and 0.2, respectively. To obtain a more comprehensive picture of the coregulation of chloroplast genes, we performed hierarchical clustering of chloroplast gene expression over "all" conditions and included also the nuclear *SIG1* to *SIG6* in this analysis (Supplemental Fig. S4C). *SIG1*, *SIG2*, *SIG3*, *SIG4*, and *SIG6* clustered together, whereas *SIG5* did not cluster with any of the investigated genes. Class I and II genes were distributed all over the clustering tree but formed some subclusters, for instance one containing the class I genes *psaA*, *psbC*, *psbD*, *psbH*, and *ndhA*. The class III genes *rpoC1*, *rpoC2*, and *accD* clustered together, whereas *rpoA* and *rpoB* were found in a different cluster. Generally, genes encoding ribosomal proteins clustered together.

Classification of the Genes According to Their Extent of Differential Expression

Of the microarray experiments considered in our analysis, 413 experiments represented comparisons either between treated and nontreated plants or between mutant and wild-type samples. All the genes coding for chloroplast or mitochondrial proteins that we investigated displayed at least a 2-fold difference in expression in at least two of the 413 experiments. A set of 127 genes showed such a differential response in at least 103 (or one-fourth) of the 413 experiments and were classified as "very highly responsive" (Supplemental Table S6). Except for the chloroplast gene *ATCG01010* encoding the chloroplast NAD(P)H dehydrogenase subunit F (NdhF), all very highly respon-

sive genes were found to be located in the nucleus or mitochondrion. Numerous genes coding for auxin-responsive proteins and P-glycoproteins were found among the very highly responsive genes. In addition, many genes associated with stress responses were identified, including nuclear genes for chloroplast (AT2G29500 and AT4G27670) and mitochondrial (AT1G52560, AT4G25200, and AT5G51440) heat shock proteins, alternative oxidases 1C and D (AOX1C and -D [AT3G27620 and AT1G32350]), early light-inducible proteins 1 and 2 (ELIP1 and -2 [AT3G22840 and AT4G14690]), cold-regulated 15B protein (COR15B [AT2G42530]), and the 9-cis-epoxycarotenoid dioxygenase (NCED) genes *NCED2*, *-3*, and *-5* (*AT4G18350*, *AT3G14440*, and *AT1G30100*) encoding subunits of the key enzyme in ABA biosynthesis. Related genes like *NCED4* and *-9*, or *AOX1A* and *-B*, which were not in the very highly responsive category, were found in the next category ("highly responsive" [differentially regulated in 71 to 102 experiments]; Supplemental Table S6). Genes for photosynthesis and tetrapyrrole biosynthesis were mainly found among the "moderately responsive" genes (2-fold differential in 36 to 70 experiments) and the "weakly responsive" genes (2-fold differential in six to 35 experiments).

One hundred twenty (or 4%) of the investigated genes were classified as "very weakly" responsive genes, with differential expression in only two to five experiments (Supplemental Table S6). All of these were nuclear genes, and no known biological function could be assigned to approximately 25% of them. Examples from the collection of very weakly responsive genes are those for subunits of the mitochondrial ATP synthase (*AT2G33040*, *AT5G08670*, and *AT5G08690*), the genes *TRANSLOCON AT THE OUTER ENVELOPE MEMBRANE OF CHLOROPLASTS34* (*AT5G0500*) and *TRANSLOCON AT THE INNER ENVELOPE MEMBRANE OF CHLOROPLASTS40* (*AT5G16620*), three *EMBRYO-DEFECTIVE* genes (*AT4G26300*, *AT5G02250*, and *AT5G24400*), and *GLOBULAR ARREST1* (*AT5G41480*).

Common cis-Elements in Coregulated Gene Sets

Because all nuclear chloroplast genes (cp_n^{all}) exhibited a certain level of coregulation (see above; Supplemental Fig. S1B), in particular genes from the cp_n^{photo}, cp_n^{OGE}, and cp_n^{tetpy} sets, and their combinations (Figs. 1B, 2A, and 3), the 500-bp promoter sequences of these coregulated gene sets were assessed for the occurrence of 21 known plant transcription factor binding sites by PScan (Zambelli et al., 2009; Fig. 4A), and scanned for novel cis-elements by the Amadeus program (Linhart et al., 2008; Fig. 4C). Employing PScan, binding sites for the basic leucine zipper (bZIP) proteins EmBP-1, bZIP910, bZIP911, and TGA1A, the Myb transcription factor ARABIDOPSIS RESPONSE REGULATOR10 (ARR10), and the AP2 MBD-like transcription factor ABA INSENSITIVE4 (ABI4) were found to be overrepresented in cp_n^{all} and cp_n^{photo} and, with the exception of ARR10, also in cp_n^{OGE}. However, in cp_n^{tetpy} promoters only, the binding motifs for EmBP-1 and ARR10 were enriched, suggesting that a distinct mode of transcriptional regulation exists for this gene set.

The use of Amadeus allowed us to identify three 8-bp elements containing the ACGT element in the three gene sets, which were 2-fold enriched compared with the genome-wide frequency in Arabidopsis 500-bp promoters and which were already identified by PScan as being associated with binding of EMPB-1, TGA1A, bZIP910, and bZIP911. In addition, two novel motifs, NGAAYRYY and BSKTATCY, designated cpCoReg1 and cpCoReg2, were identified, which were even more highly overrepresented than the ACGT-containing motifs (average enrichment factors in all three gene sets of 4.1 and 3.2, respectively; Fig. 4C). Thus, within the promoter sequences of the genes present in cp_n^{photo}, cp_n^{OGE}, and cp_n^{tetpy}, cpCoReg1 was found in 17 (or 12.8%), 14 (or 15.1%), and 11 (or 26.2%) promoters, respectively, whereas the genome-wide average was 4.6%. In the case of cpCoReg2, the motif was found in 46 (or 34.6%), 26 (or 28%), and 12 (or 28.6%) of the promoter sequences of the genes contained in cp_n^{photo}, cp_n^{OGE}, and cp_n^{tetpy}, respectively, but only in 2,881 (or 10.1%) of all 28,446 Arabidopsis promoter sequences. When the matrix library collected in TRANSFAC (Wingender et al., 1996) was searched by Amadeus for known transcription factors binding to sequences similar to cpCoReg1 and -2, the mammalian mitochondrial transcription factor A (mtTFA; Litonin et al., 2010) and the *Drosophila* heat shock factor (HSF; Jensen et al., 2008) were identified as best hits.

For mitochondria, mt_n^{resp} and mt_n^{OGE} and their combination exhibited a certain level of coexpression (Figs. 1B, 2A, and 3). Examination by PScan for known binding sites of transcription factors, using as a control the gene set mt_n^{all} (which displayed very little coregulation), resulted in a much more complex picture than the one obtained for nuclear chloroplast gene sets (Fig. 4B). Binding sites for seven transcription factors were found to be overrepresented in mt_n^{all}. Of these, only ABI4 and MNB1A (DNA-binding with one finger1 [Dof1]) displayed significant *P* values for the mt_n^{resp} gene set. For mt_n^{OGE}, no overrepresented transcription factor-binding site could be identified. The Amadeus program was then used to identify novel 8-bp consensus sequences overrepresented in the 500-bp promoter regions of the mt_n^{resp} and mt_n^{OGE} gene sets (Fig. 4D). Two hits with enrichment factors of 2.3 and 2.8, respectively, were obtained. Both resembled the AAAG-binding element common to MNB1A, Dof2, and Dof3 and already detected by PScan. In addition, GMNAANMY, designated in the following as mtCoReg1, was found as the most significant common element within mt_n^{resp} and mt_n^{OGE}, with frequencies of 44.5% and 60%, respectively, in the respective promoter sequences. These frequencies represent a marked enrichment when compared with the genome-wide average for mtCoReg1 (6,959 of

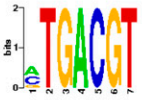
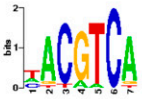
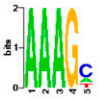
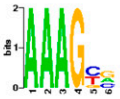
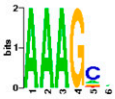
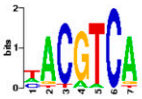
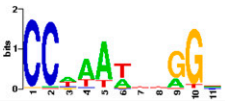
Transcription Factor	Motif Logo	*P* value			
A. Known cis-regulons identified in nuclear chloroplast gene sets					
		cp_n^{photo}	cp_n^{OGE}	cp_n^{tetpy}	cp_n^{all}
EmBP-1		8.14 E-13	0.447	0.043	1.15 E-8
bZIP911		8.13 E-7	0.023	>0.050	4.14 E-7
bZIP910		3.7 E-3	0.029	>0.050	4.39 E-5
ARR10		0.014	>0.050	0.04	0.050
TGA1A		0.027	0.027	>0.050	2.6 E-3
ABI4		0.034	0.016	>0.050	2.06 E-8
B. Known cis-regulons identified in nuclear mitochondrial gene sets					
		mt_n^{resp}	mt_n^{OGE}	mt_n^{all}	
ABI4		5.59 E-5	>0.050	5.08 E-3	
MNB1A		4.92 E-3	>0.050	2.78 E-3	
Dof3		0.018	>0.050	1.22 E-3	
Dof2		0.021	>0.050	9.29 E-3	
TGA1A		>0.050	8.00 E-3	>0.050	
EMBP-1		>0.050	>0.050	8.30 E-3	
AG		>0.050	>0.050	3.90 E-3	

Figure 4. (*Figure continues on following page.*)

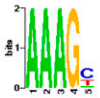
Transcription Factor	Motif Logo	*P* value		
Gamyb		>0.050	>0.050	3.90 E-3
PBF		>0.050	>0.050	0.017
C. Novel cis-regulons identified in nuclear chloroplast gene sets				
		cp_n^{photo}	cp_n^{OGE}	cp_n^{tetpy}
cpCoReg1		4.3 E-5	2 E-5	1.6 E-6
cpCoReg2		5.0 E-16	7.1 E-8	5.4 E-8
D. Novel cis-regulons identified in nuclear mitochondrial gene sets				
		mt_n^{resp}	mt_n^{OGE}	
mtCoReg1		6.2 E-11	1.0 E-7	
mtCoReg2		5.7 E-12	1.0 E-4	

Figure 4. Known (A and B) and novel (C and D) cis-regulons identified in coregulated gene sets. *P* values were calculated based on the whole-genome set of Arabidopsis. For the designation of the gene sets, see Table I.

28,446 promoters; 23.2%). Another cis-element identified in mt_n^{resp} and mt_n^{OGE} was HCGGRYHD (mtCoReg2; Fig. 4D), detected in 37.8% and 40% of the respective promoters. Again, the element occurs much more frequently than the genome-wide frequency would lead one to expect (3,618 of 28,446 promoters; 12.7%).

DISCUSSION

In this study, the absolute and relative expression of different sets of nuclear and organellar genes under eight different categories of conditions was analyzed. The comprehensive data collections and databases available (e.g. http://www.ncbi.nlm.nih.gov/geo, http://arabidopsis.org, and http://mapman.gabipd.org/web/guest/mapman) made the classification of the gene sets with respect to their functions in organellar gene expression, energy-transducing processes, or tetrapyrrole biosynthesis straightforward and unambiguous. The definition of the eight categories of environmental and genetic conditions relevant for organelle functions, however, is of necessity ambiguous, because of the complications introduced by the possible pleiotropy of some treatments or genetic lesions, which may trigger multiple signaling pathways and/or affect both organelles. Further impurities were added to the analysis by considering microarrays performed on plant material of different developmental stages. Nevertheless, we obtained clear evidence for specific reactions of the gene sets under certain categories of conditions, implying that within our predefined categories certain stimuli were indeed enriched, allowing us to recognize distinct transcriptional trends and responses.

High Levels of Transcripts for Abundant Proteins Might Reflect the Transcriptional Regulation of Protein Abundance

Transcript abundance can reflect absolute protein abundance, even in the presence of posttranslational regulation (Kleffmann et al., 2004) and thus give first insights into the actual level of activity of biological processes. Our analyses revealed that transcripts encoding chloroplast proteins, particularly ones involved in photosynthesis, were generally highly abundant, irrespective of whether the corresponding gene is located in the nucleus or the organelle. Photosynthesis proteins are also highly abundant; therefore, the

strong correlation between transcript and protein levels for photosynthesis-related proteins might reflect the highly dynamic nature of the photosynthetic apparatus, insofar as it can rapidly adapt to metabolic and light fluctuations even at the transcriptional level (Eberhard et al., 2008). Only in the "*gun* signaling" category, in which seedlings were treated with certain inhibitors, was the absolute expression of the cp_n^{photo} and cp_{org}^{photo} gene sets comparatively low (Supplemental Fig. S1A). This might be explained by the generally lower expression of photosynthesis genes in cotyledons (which were used for the analyses) compared with leaves (data not shown; https://www.genevestigator.com) and by the fact that the application of these inhibitors prevents the expression of many nuclear genes for photosynthesis not only in wild-type plants but also, to a certain degree, in *gun* mutants (Oelmüller and Mohr, 1986; Gray et al., 2003).

For proteins of the mitochondrial respiratory chain, the corresponding nuclear genes were also expressed at significantly higher levels than the genomic average, although four times lower than the values for those encoding photosynthesis proteins (Fig. 1A). As with photosynthesis genes, this might indicate that a close correlation between transcript and protein levels could contribute to dynamic alterations of the activity of the respiratory chain by transcriptional regulation. Such alterations might be required to cope with increased electron flux in the light without substantially increasing mitochondrial ROS production. Indeed, it was shown that light directly influences the transcription of nuclear genes coding for components of the respiratory electron transport chain, probably to support photosynthetic metabolism (Escobar et al., 2004; Yoshida and Noguchi, 2009). Strikingly, genes for mitochondrial proteins located in the organelle, especially those coding for respiration, displayed very low abundance, even lower than the genomic average (Fig. 1A; Supplemental Fig. S1A). It was shown previously that, although transcript synthesis in Arabidopsis mitochondria cycled in a diurnal rhythm, steady-state transcript levels were stable, implying that the available steady-state transcript levels in plant mitochondria are sufficient to provide the required translation capacity also at times of peak respiratory and physiological demands (Okada and Brennicke, 2006). This hypothesis based on the analysis of diurnal rhythms can now be generalized to the many conditions analyzed in our study and can explain the negative correlation between mRNAs coding for mitochondrial proteins responsible for the same process that are transcribed from nuclear and organellar genes (Fig. 2D).

Coregulation within Gene Sets and between Gene Sets in the Same Genetic Compartment: Metabolic Implications

With the exception of mt_n^{all}, all sets of nuclear chloroplast or mitochondrial genes showed markedly higher degrees of coregulation than did random controls (Fig. 1B), corroborating, for nuclear chloroplast genes, the previous proposal of a transcriptional signature common to most nuclear chloroplast genes (master switch; Richly et al., 2003). Moreover, both nuclear photosynthesis and cpOGE genes displayed even higher levels of coregulation, a finding that is in agreement with the fact that they cluster in only two of the 23 regulons identified on the basis of the behavior patterns of some 3,000 different nuclear Arabidopsis gene transcripts under 101 conditions (Biehl et al., 2005). In addition, the cp_n^{tetpy}, mt_n^{resp}, and mt_n^{OGE} gene sets each displayed levels of coregulation comparable to that in the cp_n^{all} set (Fig. 1B), implying that coregulation at the transcript level also encompasses nuclear genes for tetrapyrrole synthesis (the pathway that provides the photosynthetic pigments), for the respiratory chain, and for mitochondrial gene expression. The negative effect of "photosynthesis" conditions on the level of coexpression of genes for chloroplast proteins (Supplemental Fig. S1B) can be interpreted in the context of the complex metabolic reprogramming that must occur in chloroplasts in general (cp_n^{all}) to compensate for changes in photosynthetic activity and readjust the activity of photosynthetic electron flow. In contrast, the more highly synchronized expression of nuclear mitochondrial genes under "photosynthesis" conditions might reflect a general "emergency response" at the level of mitochondria to safeguard ATP supplies.

At the level of organellar genes, cluster analyses of chloroplast genes conducted on transcriptomes from mutants with severe effects on photosynthesis or from plants exposed to stresses suggested that the accumulation of plastid gene transcripts is regulated in response to altered states of the chloroplasts (Cho et al., 2009). To our knowledge, no comparable analysis has been conducted on mitochondrial genes. Intriguingly, we show here that all chloroplast and mitochondrial organellar gene sets are highly coregulated, even more so than their nuclear counterparts (Fig. 1B). Presumably, the high degree of coregulation of mitochondrial gene sets reflects a lack of transcriptional regulation, as steady-state levels of their transcripts generally remain stable (see above; Okada and Brennicke, 2006), whereas coregulation of the chloroplast gene sets has to be interpreted in the context of their coregulation with their nuclear pendants (see below: retrograde signaling).

Coregulation of nuclear gene sets impinging on different functions in the same genetic compartment has so far been shown only for cpOGE and photosynthesis in Arabidopsis and was interpreted as an instance of nuclear transcriptional control of plastid ribosome abundance, contributing to the coordinated expression of plastome- and nucleus-encoded proteins of the photosynthetic machinery (Biehl et al., 2005). Here, we have shown that the two sets of nuclear genes for cpOGE and tetrapyrrole biosynthesis, in particular, are coregulated and that, in turn, the cp_n^{tetpy}/cp_n^{OGE} module is coregulated with nuclear photosynthesis genes. This implies that coordination of the expression of nucleus-encoded photosynthetic proteins, the expression of chloroplast-encoded photosynthetic proteins via the OGE machin-

ery, and the synthesis of chlorophyll, three processes that are all necessary for the efficient assembly of the photosynthetic machinery, are controlled, at least in part, at the level of mRNA expression in the nucleus. The coexpression of mtOGE and respiration was first noted at the level of conserved nuclear promoter elements in *Drosophila* (Sardiello et al., 2005) but has not previously been investigated at the transcript level in any species. We now show that also in Arabidopsis, this phenomenon serves to coordinate the accumulation of nuclear transcripts coding for respiratory proteins and mtOGE proteins (Fig. 2A). Moreover, also nuclear genes for tetrapyrrole biosynthesis are coregulated with cpOGE genes and photosynthesis genes, making perfect sense in a physiological context.

At the level of organellar genes, the extent of coordination of OGE and photosynthesis genes in the chloroplast and of OGE and respiratory chain genes in mitochondria is higher than that of any gene class or pair of nuclear gene sets but not higher than the corresponding master sets of all chloroplast or mitochondrial genes (Fig. 2B; Supplemental Fig. S2B). Therefore, it remains unclear whether a specific mechanism for coordinating the expression of genes for OGE and energy-transducing elements also operates in the organelles.

Negative Coregulation between Gene Sets from Chloroplasts and Mitochondria: Implications for Energy-Dissipating Processes?

The expression of genes for proteins that function in chloroplasts and mitochondria needs to be coordinated under certain conditions because of their well-documented metabolic interdependence (Raghavendra and Padmasree, 2003). For instance, photosynthetic processes depend on a range of compounds synthesized by mitochondria, but they also provide substrates for mitochondrial respiration. Furthermore, mitochondrial respiration protects photosynthesis against photoinhibition by dissipating redox equivalents exported from the chloroplasts (for review, see Leister, 2005). The different OEP genes (cp_n^{photo}/mt_n^{resp}) displayed a moderate negative correlation coefficient (Fig. 2C). This can be tentatively explained by opposite regulation under stress conditions, a hypothesis that is supported by the observation that levels of transcripts specifying energy-dissipating respiratory components are increased under high light (Svensson and Rasmusson, 2001; Yoshida et al., 2008) while nuclear genes encoding photosynthesis proteins are mostly down-regulated (Kimura et al., 2003), as are the levels of their products, such as PSII antenna proteins (Melis, 1991).

Coregulation of Gene Sets in Different Compartments: Implications for Retrograde and Anterograde Signaling

Organellar multiprotein complexes, such as 70S-type ribosomes, photosystems, and the respiratory chain complexes, are actually mosaics of subunits encoded by nuclear and organellar genes. Hence, their correct assembly obviously requires the coordination of OGE and NGE at different levels, and the coordinate expression of nuclear genes for photosynthesis and cpOGE (see above) represents one level of control. Regulation directly at the transcript level would provide another mode of regulation but would require retrograde signaling from the organelles to the nucleus to adjust NGE according to the demands of the organelles. Whereas chloroplast-nucleus signaling has been extensively investigated in plants (Rodermel and Park, 2003; Beck, 2005; Nott et al., 2006; Pogson et al., 2008; Woodson and Chory, 2008), mitochondrial retrograde signaling has been predominantly studied in yeast (Liu and Butow, 2006), and relatively little is known about mitochondrion-nucleus signaling in plants (Rhoads and Subbaiah, 2007). Our analysis of the coregulation of genes that code for proteins with similar functions but are distributed between different genetic compartments showed that nucleus-organelle coregulation at the transcript level is characteristic for chloroplast but not for mitochondrial functions (Fig. 2D). This conclusion is underpinned by the large number of photosynthesis genes that are positively coregulated under all conditions examined (77 nucleus-organelle gene pairs; Supplemental Table S5H). In contrast, nuclear and organellar genes coding for the respiratory chain or OGE even show some degree of negative coregulation (Fig. 2D). This is in line with the observation that, during mitochondrial biogenesis, no coordination of the expression of nucleus- and mitochondrion-encoded proteins at the transcript level was found in response to sugar starvation and refeeding, prompting the hypothesis that nucleus-mitochondrion coordination under these conditions occurs predominantly at the level of protein complex assembly rather than mRNA synthesis (Giegé et al., 2005). Our data corroborate the hypothesis that, in general, nuclear-mitochondrial coordination does not occur at the transcriptome level. However, we found that under certain conditions ("chloroplast" and "ROS"), which impinge upon the energy supply of the cell, nuclear-organelle coordination of mtOGE or mtOEP gene expression indeed occurs (Fig. 3; Supplemental Fig. S3B). Under these circumstances, some gene pairs for components of mtOGE actually showed the highest coregulation (two gene pairs with a PCC of 1) of all pairs investigated (Supplemental Table S5A).

The control of gene expression in the chloroplast is generally thought to be dominated by posttranscriptional (Somanchi and Mayfield, 1999; Barkan and Goldschmidt-Clermont, 2000; Rochaix, 2001) and translational autoregulatory and transregulatory mechanisms (Choquet and Wollman, 2002), and also for nucleus-encoded proteins, doubts have been expressed as to whether transcriptional regulation plays a major role in modulating their abundance (Kleffmann et al., 2004). However, the high degree of nucleus-chloroplast coexpression of photosynthesis and cpOGE genes implies that transcriptional

regulation in the organelle also has a part to play, particularly under conditions such as "nutrient supply" and "general stresses" (Supplemental Fig. S3B).

Different Layers of Regulation: Intranuclear Versus Nuclear-Chloroplast Versus Chloroplast-Mitochondrion Coregulation

The hierarchical clustering of gene set expression for all eight individual conditions, as well as over all conditions (Fig. 3), showed that the tightest coregulation generally occurs for genes that are located in the same genetic compartment and code for products targeted to the same organelle. However, it also emerges that coregulation between genetic compartments is characteristic for chloroplasts and occurs at a basal level also for mitochondria. In chloroplasts, nucleus-organelle coregulation can actually predominate over intracompartmental networks, as exemplified by the coexpression of nuclear and organellar photosynthesis genes under "general stresses" conditions (Fig. 3).

The general dominance of intranuclear and nuclear-organelle coregulation of genes for chloroplast proteins can be overridden under certain conditions. Thus, organelle-organelle (i.e. chloroplast-mitochondrion) coregulation prevails under "sugars" conditions (Fig. 3). This may be an incidental consequence of the fact that the photosynthetic function of the organelle is essentially idle under these conditions and, therefore, coordination of photosynthesis gene expression regulation would be superfluous, but it could imply that photosynthesis itself serves as a signal emitter (Pfannschmidt, 2003). A similar dominance of organelle-organelle coordination is observed when the perception or transduction of light signals is perturbed ("light signaling") but not in retrograde signaling mutants ("*gun* signaling"). This suggests either that anterograde mechanisms are mostly responsible for the nuclear-organelle coregulation events described here or that additional retrograde signaling pathways can compensate for the loss of GUN functions.

Identification of Tentative Chloroplast Target Genes of σ Factors SIG1 to SIG6

Considerable advances have been made in elucidating the role of σ factors for specific promoter recognition and selected transcription of some plastid genes (Lerbs-Mache, 2011). However, the unambiguous identification of target genes was hindered for two reasons. (1) One σ factor can have both specialized roles (at certain promoters and certain times/tissues) and overlapping redundant functions (at other promoters, times, and situations; Schweer et al., 2010). (2) The promoter of one plastid gene can be the target of multiple σ factors (Schweer et al., 2010; Lerbs-Mache, 2011). Our coregulation analysis of chloroplast genes (Supplemental Fig. S4C) has the potential to provide the basis to determine which σ factors might be involved in the transcriptional regulation of certain PEP-transcribed chloroplast genes. Thus, the promoters of *rbcL* and *psbA* are recognized by SIG2, whereas the SIG3-PEP holoenzyme transcribes specifically *psbN* and *atpH* (Lerbs-Mache, 2011). Indeed, *rbcL* and *psbA* were found in adjacent clusters, and *psbN* and *atpH* grouped in the same cluster. Therefore, it is tempting to speculate that the transcription of genes present in the same cluster like *rbcL* and *psbA*, such as *ndhJ*, *psbG*, *psaI*, and *orf31*, is also regulated by SIG2; accordingly, promoters of other genes from the *psbN-atpH* cluster might be recognized by SIG3.

The Paucity of Marker Genes Might Emanate from the Complexity of Intercompartmental Signaling

All genes investigated in this study were assigned to different classes according to their extent of differential expression, ranging from "very low" to "very high" responsiveness (Supplemental Table S6). Sets of 127, 170, or 120 genes were classified as "very highly," "highly," or "very low" responsive, respectively. The categories of "medium" and "low" responsiveness included genes differentially regulated in 36 to 70 and six to 35 experiments, respectively, and contained the vast majority of genes (583 and 1,891 genes, respectively). These two classes were exploited to identify marker genes specifically regulated under certain classes of conditions. However, not even one marker gene could be identified, strongly suggesting that the specific stimuli for plastid signaling considered in each of the investigated eight categories are often overlapping or interconnected. This supports the previous notion that the effects of individual stimuli on transcription are difficult to separate and that the pathways of communication between various organelles of a plant cell are complex and interdependent (Leister, 2005; Koussevitzky et al., 2007; Giraud et al., 2009).

Toward the Identification of Transcription Factors That Coordinate the Expression of Nuclear Genes for Organellar Proteins

In our study, several Leu-zipper transcription factors were identified that might account for coregulation of the three nuclear chloroplast gene sets and belong to the group of transcription factors binding the ACGT core element (Fig. 4). Proteins binding to this element have been associated with light-dependent regulation of the expression of photosynthesis genes (Donald and Cashmore, 1990; Meier and Gruissem, 1994; Sun and Ni, 2011) and play a role in hormone signaling (Liu and Lam, 1994). ARR10, the binding site for which was overrepresented in several nuclear chloroplast gene sets, binds the core sequence AGAT (Hosoda et al., 2002) and belongs to response regulators in His-to-Asp phosphorelays, which are involved particularly in the response of hormones (Urao et al., 2000). ABI4-binding sites were also enriched in nuclear chloroplast gene sets. Accordingly, ABI4 was shown

to bind the sequence CACCG in ABA and sugar response genes in maize (*Zea mays*; Niu et al., 2002) and was postulated to be a component of chloroplast (Koussevitzky et al., 2007) as well as mitochondrial (Giraud et al., 2009) retrograde signaling. Actually, ABI4 is thought to bind to the CCAC element of the promoter of the gene for the light-harvesting chlorophyll *a*/*b*-binding protein *LHCB1.2* in response to the GUN1-derived signal, which in turn prevents the binding of G box-binding factors required for the light-induced expression of nuclear photosynthetic genes (Koussevitzky et al., 2007).

Binding sites for the transcription factors ABI4 and MNB1A were overrepresented in the mt_n^{all} and mt_n^{resp} gene sets (Fig. 4). Indeed, binding of ABI4 to the AOX1a promoter has been demonstrated by electromobility shift and yeast one-hybrid assays (Giraud et al., 2009). Dof transcription factors, of which MNB1A (Dof1) is one, are involved in tissue-specific and light-regulated gene expression (Yanagisawa and Sheen, 1998). MNB1A is associated with the expression of multiple genes involved in carbon metabolism in maize (Yanagisawa and Sheen, 1998; Yanagisawa, 2000), but no further targets are yet known in plants. Thus, the predicted target promoters of the mt_n^{all} and mt_n^{resp} gene sets might represent a good starting point for their identification. However, in contrast to the nuclear chloroplast gene sets, no known transcription factor-binding site common to all nuclear mitochondrial gene sets could be identified. The mtCoReg2 element identified here (HCGGRYHD; Fig. 4) strongly resembles the site II element (TGGGCC/T), which is the main determinant of the expression levels of the three Arabidopsis *COX6b* genes for cytochrome *c* oxidase subunit 6b (Mufarrege et al., 2009), the *COX5b-2* gene (Comelli and Gonzalez, 2009), and the *CYTC-2* gene, encoding an isoform of cytochrome *c* (Welchen and Gonzalez, 2005). Only one of the *COX6b* genes was found among the putative targets of mtCoReg2, implying that the mtCoReg2-binding site is an independent cis-regulatory element from the site II element. Taken together, the newly identified cpCoReg and mtCoReg elements and their putative target genes (Supplemental Table S7) represent attractive targets for further studies on retrograde signaling and the transcription factors involved in it.

CONCLUSION

The concept of retrograde signaling includes that organelles convey information on their developmental and metabolic state to the nucleus, thus enabling NGE to be appropriately modified. On the contrary, chloroplast gene expression was thought to be mainly regulated by posttranscriptional mechanisms. However, the notion that chloroplast gene expression is also regulated at the transcriptional level in higher plants became accepted more recently (Eberhard et al., 2008; Lerbs-Mache, 2011). In this study, we show that all chloroplast and mitochondrial organellar gene sets are highly coregulated at the transcript level, even higher than their nuclear counterparts. Moreover, the coregulation of genes that code for proteins with similar functions but are distributed between different genetic compartments has only been studied for a few instances of singular genes, whereas our global analysis clearly shows that nucleus-organelle coregulation at the transcript level is characteristic for chloroplast but not for mitochondrial functions. In contrast, nuclear and organellar genes coding for the respiratory chain or OGE even show some degree of negative coregulation.

Retrograde signaling was discovered more than 30 years ago, but knowledge of the involved nuclear transcription factors is still scarce. We provide here a comprehensive list of conserved motifs in the promoter regions of gene sets with a high level of coregulation, serving as a starting point to identify the transcription factors involved.

MATERIALS AND METHODS

Gene Sets and Microarray Expression Data

Protein-coding Arabidopsis (*Arabidopsis thaliana*) genes located in plastids or mitochondria were retrieved from Affymetrix ATH1 array elements (Table I; Supplemental Table S1). The set of 1,476 nuclear genes encoding chloroplast proteins was extracted from a previously compiled list of 1,808 chloroplast proteins (Yu et al., 2008), selecting those proteins that had been reliably assigned to chloroplasts and adding known chloroplast proteins not already on the list. The set of 1,323 nuclear genes for mitochondrial proteins was compiled by extracting mitochondrial proteins validated by the presence of at least one EST in the SUBA (Heazlewood et al., 2007) and the MitoP2 (Elstner et al., 2009) databases. Assignment of nuclear and organellar genes to the functional subclasses OGE, photosynthesis, respiration, and tetrapyrrole biosynthesis was done according to Giegé et al. (2005) using the SUBA database and the mapping files of MapMan (Thimm et al., 2004).

A total of 101 processed Affymetrix ATH1 microarray data sets comprising 1,290 hybridization experiments and covering eight different categories of genetic or environmental conditions/perturbations (Table II; Supplemental Table S2) were down-loaded from ArrayExpress, the public expression data repository (http://www.ebi.ac.uk/microarray-as/ae/browse.html?keywords=arabidopsis).

Analysis of Absolute Expression Data

All computational analyses were either performed in the statistical programming language R (http://www.r-project.org) or using custom-made PERL scripts on a Sun Grid Engine cluster with 64-bit architecture running on a Linux operating system. Gene identifiers were remapped to The Arabidopsis Information Resource 9 annotation to ensure consistency of expression data for each gene between the various experiments and groups (Poole, 2007). Expression values were scaled to mean 0 and variance 1 by a z-transformation. In this study, mean expression values for a gene set, whether a gene class, an experimental group, or a combination of both, are defined as an ensemble average (i.e. as a mean z-score of all genes contained in that class, group, or combination). Monte Carlo simulations were employed to assess the significance of differences between the average background expression (the null hypothesis) and expression levels of a particular combination. One thousand samples, each containing the same number of genes as the respective combination, were randomly and uniformly drawn from all genes without replacement. A Mersenne-Twister implementation was used in the PYTHON system package "random" for the random number generator to minimize biased selections. Mean expression levels of these random samples were calculated as described above and used to assess the background distribution. Significance levels were estimated as the fraction of random samples with

means greater than or equal to the mean of the gene class-experimental group combination.

Analysis of Expression Similarities

Similarity between genes in expression response upon genetic/environmental perturbation was quantified by computing the PCC of their z-transformed absolute expression values. Expression similarity within gene sets was defined as the mean of the PCCs of all possible gene pairs within a given gene set. Expression similarity between gene sets was defined as the mean PCC of all interset pairs derived from the cross-product of the two gene sets. Significance was assessed by Monte Carlo simulations as described above on the basis of 100 random samples.

Clustering

The R package pvclust (http://www.is.titech.ac.jp/~shimo/prog/pvclust/) was used to carry out hierarchical clustering with average linkage (Suzuki and Shimodaira, 2006). *P* values were either derived as approximately unbiased *P* values from multiscale or as bootstrap probability *P* values from standard bootstrap resampling. Both methods supported identical topologies and conclusions for all trees.

Identification of Coregulated Gene Pairs

The degree of correlation of transcriptional activity among genes from two gene sets was calculated by computing the PCCs for pairs of genes chosen from the different gene sets under the respective conditions. To identify significantly coregulated gene pairs, all-against-all Pearson correlation matrices from 23,913 Arabidopsis genes were calculated under each relevant experimental condition, and the background distributions of pairwise expression similarities from the correlation matrices were determined.

Identification of cis-Elements

The program PScan (Zambelli et al., 2009) was used to locate binding sites for known plant transcription factors. For the identification of novel cis-elements, the expectation maximization-based Amadeus program (Linhart et al., 2008), in which a list of 28,446 background promoters is implemented, was used.

Supplemental Data

The following materials are available in the online version of this article.

Supplemental Figure S1. Absolute mRNA expression of gene sets and expression similarities in all eight categories of conditions.

Supplemental Figure S2. Expression similarities between different gene sets from the same genetic compartment that code for products destined for the same organelle in all eight different condition categories.

Supplemental Figure S3. Expression similarities between different gene sets from the same genetic compartment that code for products destined for different organelles (A) or from different genetic compartment and with products directed to the same organelle (B) in all eight categories of conditions.

Supplemental Figure S4. Absolute mRNA expression, expression similarities, and coregulation of class I, II, and III chloroplast genes.

Supplemental Table S1. List of all nuclear and organellar genes for chloroplast and mitochondrial proteins.

Supplemental Table S2. List of all data sets and hybridization experiments.

Supplemental Table S3. Monte Carlo simulation for expression similarities within gene sets.

Supplemental Table S4. Monte Carlo simulation for expression similarities between gene sets.

Supplemental Table S5. Gene pairs derived from different gene sets that are coregulated under different conditions.

Supplemental Table S6. List of all differentially regulated genes and the number of experiments in which they are regulated.

ACKNOWLEDGMENTS

We thank Paul Hardy for critical comments on the manuscript.

Received April 5, 2011; accepted July 19, 2011; published July 20, 2011.

LITERATURE CITED

Alabadí D, Blázquez MA (2009) Molecular interactions between light and hormone signaling to control plant growth. Plant Mol Biol **69:** 409–417

Apel K, Hirt H (2004) Reactive oxygen species: metabolism, oxidative stress, and signal transduction. Annu Rev Plant Biol **55:** 373–399

Barkan A, Goldschmidt-Clermont M (2000) Participation of nuclear genes in chloroplast gene expression. Biochimie **82:** 559–572

Beck CF (2005) Signaling pathways from the chloroplast to the nucleus. Planta **222:** 743–756

Biehl A, Richly E, Noutsos C, Salamini F, Leister D (2005) Analysis of 101 nuclear transcriptomes reveals 23 distinct regulons and their relationship to metabolism, chromosomal gene distribution and co-ordination of nuclear and plastid gene expression. Gene **344:** 33–41

Bonardi V, Pesaresi P, Becker T, Schleiff E, Wagner R, Pfannschmidt T, Jahns P, Leister D (2005) Photosystem II core phosphorylation and photosynthetic acclimation require two different protein kinases. Nature **437:** 1179–1182

Bräutigam K, Dietzel L, Kleine T, Ströher E, Wormuth D, Dietz KJ, Radke D, Wirtz M, Hell R, Dörmann P, et al (2009) Dynamic plastid redox signals integrate gene expression and metabolism to induce distinct metabolic states in photosynthetic acclimation in *Arabidopsis*. Plant Cell **21:** 2715–2732

Cherniad'ev II (2000) [Ontogenetic changes in the photosynthetic apparatus and effect of cytokinins]. Prikl Biokhim Mikrobiol **36:** 611–625

Cho WK, Geimer S, Meurer J (2009) Cluster analysis and comparison of various chloroplast transcriptomes and genes in *Arabidopsis thaliana*. DNA Res **16:** 31–44

Cho YH, Yoo SD, Sheen J (2006) Regulatory functions of nuclear hexokinase1 complex in glucose signaling. Cell **127:** 579–589

Choquet Y, Wollman FA (2002) Translational regulations as specific traits of chloroplast gene expression. FEBS Lett **529:** 39–42

Comelli RN, Gonzalez DH (2009) Identification of regulatory elements involved in expression and induction by sucrose and UV-B light of the *Arabidopsis thaliana COX5b-2* gene, encoding an isoform of cytochrome c oxidase subunit 5b. Physiol Plant **137:** 213–224

Davletova S, Schlauch K, Coutu J, Mittler R (2005) The zinc-finger protein Zat12 plays a central role in reactive oxygen and abiotic stress signaling in Arabidopsis. Plant Physiol **139:** 847–856

Donald RG, Cashmore AR (1990) Mutation of either G box or I box sequences profoundly affects expression from the Arabidopsis *rbcS-1A* promoter. EMBO J **9:** 1717–1726

Eberhard S, Finazzi G, Wollman FA (2008) The dynamics of photosynthesis. Annu Rev Genet **42:** 463–515

Elstner M, Andreoli C, Klopstock T, Meitinger T, Prokisch H (2009) The mitochondrial proteome database: MitoP2. Methods Enzymol **457:** 3–20

Escobar MA, Franklin KA, Svensson AS, Salter MG, Whitelam GC, Rasmusson AG (2004) Light regulation of the Arabidopsis respiratory chain: multiple discrete photoreceptor responses contribute to induction of type II NAD(P)H dehydrogenase genes. Plant Physiol **136:** 2710–2721

Flügge UI (1999) Phosphate translocators in plastids. Annu Rev Plant Physiol Plant Mol Biol **50:** 27–45

Giegé P, Sweetlove LJ, Cognat V, Leaver CJ (2005) Coordination of nuclear and mitochondrial genome expression during mitochondrial biogenesis in *Arabidopsis*. Plant Cell **17:** 1497–1512

Giraud E, Van Aken O, Ho LH, Whelan J (2009) The transcription factor ABI4 is a regulator of mitochondrial retrograde expression of *ALTERNATIVE OXIDASE1a*. Plant Physiol **150:** 1286–1296

Gray JC, Sullivan JA, Wang JH, Jerome CA, MacLean D (2003) Coordi-

nation of plastid and nuclear gene expression. Philos Trans R Soc Lond B Biol Sci **358:** 135–144; discussion 144–145

Hajdukiewicz PT, Allison LA, Maliga P (1997) The two RNA polymerases encoded by the nuclear and the plastid compartments transcribe distinct groups of genes in tobacco plastids. EMBO J **16:** 4041–4048

Hanaoka M, Kanamaru K, Takahashi H, Tanaka K (2003) Molecular genetic analysis of chloroplast gene promoters dependent on SIG2, a nucleus-encoded sigma factor for the plastid-encoded RNA polymerase, in *Arabidopsis thaliana*. Nucleic Acids Res **31:** 7090–7098

Heazlewood JL, Verboom RE, Tonti-Filippini J, Small I, Millar AH (2007) SUBA: the Arabidopsis Subcellular Database. Nucleic Acids Res **35:** D213–D218

Hosoda K, Imamura A, Katoh E, Hatta T, Tachiki M, Yamada H, Mizuno T, Yamazaki T (2002) Molecular structure of the GARP family of plant Myb-related DNA binding motifs of the *Arabidopsis* response regulators. Plant Cell **14:** 2015–2029

Ishizaki Y, Tsunoyama Y, Hatano K, Ando K, Kato K, Shinmyo A, Kobori M, Takeba G, Nakahira Y, Shiina T (2005) A nuclear-encoded sigma factor, Arabidopsis SIG6, recognizes sigma-70 type chloroplast promoters and regulates early chloroplast development in cotyledons. Plant J **42:** 133–144

Jarvis P (2008) Targeting of nucleus-encoded proteins to chloroplasts in plants. New Phytol **179:** 257–285

Jensen LT, Nielsen MM, Loeschcke V (2008) New candidate genes for heat resistance in *Drosophila melanogaster* are regulated by HSF. Cell Stress Chaperones **13:** 177–182

Jung SH, Lee JY, Lee DH (2003) Use of SAGE technology to reveal changes in gene expression in Arabidopsis leaves undergoing cold stress. Plant Mol Biol **52:** 553–567

Kim C, Lee KP, Baruah A, Nater M, Göbel C, Feussner I, Apel K (2009) $^1O_2$-mediated retrograde signaling during late embryogenesis predetermines plastid differentiation in seedlings by recruiting abscisic acid. Proc Natl Acad Sci USA **106:** 9920–9924

Kimura M, Yamamoto YY, Seki M, Sakurai T, Sato M, Abe T, Yoshida S, Manabe K, Shinozaki K, Matsui M (2003) Identification of Arabidopsis genes regulated by high light-stress using cDNA microarray. Photochem Photobiol **77:** 226–233

Kleffmann T, Russenberger D, von Zychlinski A, Christopher W, Sjölander K, Gruissem W, Baginsky S (2004) The *Arabidopsis thaliana* chloroplast proteome reveals pathway abundance and novel protein functions. Curr Biol **14:** 354–362

Kleine T, Maier UG, Leister D (2009a) DNA transfer from organelles to the nucleus: the idiosyncratic genetics of endosymbiosis. Annu Rev Plant Biol **60:** 115–138

Kleine T, Voigt C, Leister D (2009b) Plastid signalling to the nucleus: messengers still lost in the mists? Trends Genet **25:** 185–192

Koussevitzky S, Nott A, Mockler TC, Hong F, Sachetto-Martins G, Surpin M, Lim J, Mittler R, Chory J (2007) Signals from chloroplasts converge to regulate nuclear gene expression. Science **316:** 715–719

Laloi C, Stachowiak M, Pers-Kamczyc E, Warzych E, Murgia I, Apel K (2007) Cross-talk between singlet oxygen- and hydrogen peroxide-dependent signaling of stress responses in *Arabidopsis thaliana*. Proc Natl Acad Sci USA **104:** 672–677

Larkin RM, Ruckle ME (2008) Integration of light and plastid signals. Curr Opin Plant Biol **11:** 593–599

Lau OS, Deng XW (2010) Plant hormone signaling lightens up: integrators of light and hormones. Curr Opin Plant Biol **13:** 571–577

Leister D (2005) Genomics-based dissection of the cross-talk of chloroplasts with the nucleus and mitochondria in Arabidopsis. Gene **354:** 110–116

Lerbs-Mache S (2011) Function of plastid sigma factors in higher plants: regulation of gene expression or just preservation of constitutive transcription? Plant Mol Biol **76:** 235–249

Li JM, Chory J (1999) Brassinosteroid actions in plants. J Exp Bot **50:** 275–282

Liere K, Weihe A, Börner T (2011) The transcription machineries of plant mitochondria and chloroplasts: composition, function, and regulation. J Plant Physiol **168:** 1345–1360

Linhart C, Halperin Y, Shamir R (2008) Transcription factor and microRNA motif discovery: the Amadeus platform and a compendium of metazoan target sets. Genome Res **18:** 1180–1189

Litonin D, Sologub M, Shi Y, Savkina M, Anikin M, Falkenberg M, Gustafsson CM, Temiakov D (2010) Human mitochondrial transcription revisited: only TFAM and TFB2M are required for transcription of the mitochondrial genes in vitro. J Biol Chem **285:** 18129–18133

Liu XJ, Lam E (1994) Two binding sites for the plant transcription factor ASF-1 can respond to auxin treatments in transgenic tobacco. J Biol Chem **269:** 668–675

Liu Z, Butow RA (2006) Mitochondrial retrograde signaling. Annu Rev Genet **40:** 159–185

Loschelder H, Schweer J, Link B, Link G (2006) Dual temporal role of plastid sigma factor 6 in Arabidopsis development. Plant Physiol **142:** 642–650

Maclean D, Jerome CA, Brown AP, Gray JC (2008) Co-regulation of nuclear genes encoding plastid ribosomal proteins by light and plastid signals during seedling development in tobacco and Arabidopsis. Plant Mol Biol **66:** 475–490

Mehler AH (1951) Studies on reactions of illuminated chloroplasts. I. Mechanism of the reduction of oxygen and other Hill reagents. Arch Biochem Biophys **33:** 65–77

Meier I, Gruissem W (1994) Novel conserved sequence motifs in plant G-box binding proteins and implications for interactive domains. Nucleic Acids Res **22:** 470–478

Melis A (1991) Dynamics of photosynthetic membrane-composition and function. Biochim Biophys Acta **1058:** 87–106

Meskauskiene R, Nater M, Goslings D, Kessler F, op den Camp R, Apel K (2001) FLU: a negative regulator of chlorophyll biosynthesis in *Arabidopsis thaliana*. Proc Natl Acad Sci USA **98:** 12826–12831

Mochizuki N, Brusslan JA, Larkin R, Nagatani A, Chory J (2001) Arabidopsis *genomes uncoupled 5* (*GUN5*) mutant reveals the involvement of Mg-chelatase H subunit in plastid-to-nucleus signal transduction. Proc Natl Acad Sci USA **98:** 2053–2058

Mufarrege EF, Curi GC, Gonzalez DH (2009) Common sets of promoter elements determine the expression characteristics of three Arabidopsis genes encoding isoforms of mitochondrial cytochrome c oxidase subunit 6b. Plant Cell Physiol **50:** 1393–1399

Niu X, Helentjaris T, Bate NJ (2002) Maize ABI4 binds coupling element1 in abscisic acid and sugar response genes. Plant Cell **14:** 2565–2575

Nott A, Jung HS, Koussevitzky S, Chory J (2006) Plastid-to-nucleus retrograde signaling. Annu Rev Plant Biol **57:** 739–759

Ohgishi M, Saji K, Okada K, Sakai T (2004) Functional analysis of each blue light receptor, cry1, cry2, phot1, and phot2, by using combinatorial multiple mutants in Arabidopsis. Proc Natl Acad Sci USA **101:** 2223–2228

Okada S, Brennicke A (2006) Transcript levels in plant mitochondria show a tight homeostasis during day and night. Mol Genet Genomics **276:** 71–78

Oelmüller R, Mohr H (1986) Photooxidative destruction of chloroplasts and its consequences for expression of nuclear genes. Planta **167:** 106–113

op den Camp RG, Przybyla D, Ochsenbein C, Laloi C, Kim C, Danon A, Wagner D, Hideg E, Gobel C, Feussner I, et al (2003) Rapid induction of distinct stress responses after the release of singlet oxygen in *Arabidopsis*. Plant Cell **15:** 2320–2332

Penfield S, Li Y, Gilday AD, Graham S, Graham IA (2006) *Arabidopsis* ABA INSENSITIVE4 regulates lipid mobilization in the embryo and reveals repression of seed germination by the endosperm. Plant Cell **18:** 1887–1899

Peng M, Bi YM, Zhu T, Rothstein SJ (2007) Genome-wide analysis of Arabidopsis responsive transcriptome to nitrogen limitation and its regulation by the ubiquitin ligase gene NLA. Plant Mol Biol **65:** 775–797

Pesaresi P, Hertle A, Pribil M, Kleine T, Wagner R, Strissel H, Ihnatowicz A, Bonardi V, Scharfenberg M, Schneider A, et al (2009) *Arabidopsis* STN7 kinase provides a link between short- and long-term photosynthetic acclimation. Plant Cell **21:** 2402–2423

Pfannschmidt T (2003) Chloroplast redox signals: how photosynthesis controls its own genes. Trends Plant Sci **8:** 33–41

Pinheiro C, Chaves MM (2011) Photosynthesis and drought: can we make metabolic connections from available data? J Exp Bot **62:** 869–882

Pogson BJ, Woo NS, Förster B, Small ID (2008) Plastid signalling to the nucleus and beyond. Trends Plant Sci **13:** 602–609

Poole RL (2007) The TAIR database. Methods Mol Biol **406:** 179–212

Raghavendra AS, Padmasree K (2003) Beneficial interactions of mitochondrial metabolism with photosynthetic carbon assimilation. Trends Plant Sci **8:** 546–553

Rand DM, Haney RA, Fry AJ (2004) Cytonuclear coevolution: the genomics of cooperation. Trends Ecol Evol **19:** 645–653

Rhoads DM, Subbaiah CC (2007) Mitochondrial retrograde regulation in plants. Mitochondrion **7:** 177–194

Richly E, Dietzmann A, Biehl A, Kurth J, Laloi C, Apel K, Salamini F, Leister D (2003) Covariations in the nuclear chloroplast transcriptome reveal a regulatory master-switch. EMBO Rep **4:** 491–498

Rochaix JD (2001) Posttranscriptional control of chloroplast gene expression: from RNA to photosynthetic complex. Plant Physiol **125:** 142–144

Rodermel S, Park S (2003) Pathways of intracellular communication: tetrapyrroles and plastid-to-nucleus signaling. Bioessays **25:** 631–636

Rolland F, Baena-Gonzalez E, Sheen J (2006) Sugar sensing and signaling in plants: conserved and novel mechanisms. Annu Rev Plant Biol **57:** 675–709

Rossel JB, Wilson PB, Hussain D, Woo NS, Gordon MJ, Mewett OP, Howell KA, Whelan J, Kazan K, Pogson BJ (2007) Systemic and intracellular responses to photooxidative stress in *Arabidopsis*. Plant Cell **19:** 4091–4110

Ruckle ME, DeMarco SM, Larkin RM (2007) Plastid signals remodel light signaling networks and are essential for efficient chloroplast biogenesis in *Arabidopsis*. Plant Cell **19:** 3944–3960

Ruckle ME, Larkin RM (2009) Plastid signals that affect photomorphogenesis in Arabidopsis thaliana are dependent on GENOMES UNCOUPLED 1 and cryptochrome 1. New Phytol **182:** 367–379

Sardiello M, Tripoli G, Romito A, Minervini C, Viggiano L, Caggese C, Pesole G (2005) Energy biogenesis: one key for coordinating two genomes. Trends Genet **21:** 12–16

Schmülling T, Schafer S, Romanov G (1997) Cytokinins as regulators of gene expression. Physiol Plant **100:** 505–519

Schweer J, Türkeri H, Kolpack A, Link G (2010) Role and regulation of plastid sigma factors and their functional interactors during chloroplast transcription: recent lessons from *Arabidopsis thaliana*. Eur J Cell Biol **89:** 940–946

Shen YY, Wang XF, Wu FQ, Du SY, Cao Z, Shang Y, Wang XL, Peng CC, Yu XC, Zhu SY, et al (2006) The Mg-chelatase H subunit is an abscisic acid receptor. Nature **443:** 823–826

Somanchi A, Mayfield SP (1999) Nuclear-chloroplast signalling. Curr Opin Plant Biol **2:** 404–409

Stern DB, Goldschmidt-Clermont M, Hanson MR (2010) Chloroplast RNA metabolism. Annu Rev Plant Biol **61:** 125–155

Strodtkötter I, Padmasree K, Dinakar C, Speth B, Niazi PS, Wojtera J, Voss I, Do PT, Nunes-Nesi A, Fernie AR, et al (2009) Induction of the *AOX1D* isoform of alternative oxidase in *A. thaliana* T-DNA insertion lines lacking isoform *AOX1A* is insufficient to optimize photosynthesis when treated with antimycin A. Mol Plant **2:** 284–297

Sun XD, Ni M (2011) HYPOSENSITIVE TO LIGHT, an alpha/beta fold protein, acts downstream of ELONGATED HYPOCOTYL 5 to regulate seedling de-etiolation. Mol Plant **4:** 116–126

Susek RE, Ausubel FM, Chory J (1993) Signal transduction mutants of Arabidopsis uncouple nuclear *CAB* and *RBCS* gene expression from chloroplast development. Cell **74:** 787–799

Suzuki R, Shimodaira H (2006) pvclust: an R package for assessing the uncertainty in hierarchical clustering. Bioinformatics **22:** 1540–1542

Svensson AS, Rasmusson AG (2001) Light-dependent gene expression for proteins in the respiratory chain of potato leaves. Plant J **28:** 73–82

Swiatecka-Hagenbruch M, Liere K, Börner T (2007) High diversity of plastidial promoters in *Arabidopsis thaliana*. Mol Genet Genomics **277:** 725–734

Thimm O, Bläsing O, Gibon Y, Nagel A, Meyer S, Krüger P, Selbig J, Müller LA, Rhee SY, Stitt M (2004) MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. Plant J **37:** 914–939

Tian C, Chikayama E, Tsuboi Y, Kuromori T, Shinozaki K, Kikuchi J, Hirayama T (2007) Top-down phenomics of *Arabidopsis thaliana*: metabolic profiling by one- and two-dimensional nuclear magnetic resonance spectroscopy and transcriptome analysis of albino mutants. J Biol Chem **282:** 18532–18541

Timmis JN, Ayliffe MA, Huang CY, Martin W (2004) Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. Nat Rev Genet **5:** 123–135

Umbach AL, Fiorani F, Siedow JN (2005) Characterization of transformed Arabidopsis with altered alternative oxidase levels and analysis of effects on reactive oxygen species in tissue. Plant Physiol **139:** 1806–1820

Urao T, Yamaguchi-Shinozaki K, Shinozaki K (2000) Two-component systems in plant signal transduction. Trends Plant Sci **5:** 67–74

Vanderauwera S, Zimmermann P, Rombauts S, Vandenabeele S, Langebartels C, Gruissem W, Inzé D, Van Breusegem F (2005) Genome-wide analysis of hydrogen peroxide-regulated gene expression in Arabidopsis reveals a high light-induced transcriptional cluster involved in anthocyanin biosynthesis. Plant Physiol **139:** 806–821

Welchen E, Gonzalez DH (2005) Differential expression of the Arabidopsis cytochrome c genes *Cytc-1* and *Cytc-2*: evidence for the involvement of TCP-domain protein-binding elements in anther- and meristem-specific expression of the *Cytc-1* gene. Plant Physiol **139:** 88–100

Wingender E, Dietze P, Karas H, Knüppel R (1996) TRANSFAC: a database on transcription factors and their DNA binding sites. Nucleic Acids Res **24:** 238–241

Woodson JD, Chory J (2008) Coordination of gene expression between organellar and nuclear genomes. Nat Rev Genet **9:** 383–395

Yanagisawa S (2000) Dof1 and Dof2 transcription factors are associated with expression of multiple genes involved in carbon metabolism in maize. Plant J **21:** 281–288

Yanagisawa S, Sheen J (1998) Involvement of maize Dof zinc finger proteins in tissue-specific and light-regulated gene expression. Plant Cell **10:** 75–89

Yoshida K, Noguchi K (2009) Differential gene expression profiles of the mitochondrial respiratory components in illuminated Arabidopsis leaves. Plant Cell Physiol **50:** 1449–1462

Yoshida K, Watanabe C, Kato Y, Sakamoto W, Noguchi K (2008) Influence of chloroplastic photo-oxidative stress on mitochondrial alternative oxidase capacity and respiratory properties: a case study with Arabidopsis yellow variegated 2. Plant Cell Physiol **49:** 592–603

Yu QB, Li G, Wang G, Sun JC, Wang PC, Wang C, Mi HL, Ma WM, Cui J, Cui YL, et al (2008) Construction of a chloroplast protein interaction network and functional mining of photosynthetic proteins in *Arabidopsis thaliana*. Cell Res **18:** 1007–1019

Zambelli F, Pesole G, Pavesi G (2009) PScan: finding over-represented transcription factor binding site motifs in sequences from co-regulated or co-expressed genes. Nucleic Acids Res **37:** W247–W252

Zhang ZD, Shrager J, Jain M, Chang CW, Vallon O, Grossman AR (2004) Insights into the survival of *Chlamydomonas reinhardtii* during sulfur starvation based on microarray analysis of gene expression. Eukaryot Cell **3:** 1331–1348

Zhong GY, Burns JK (2003) Profiling ethylene-regulated gene expression in *Arabidopsis thaliana* by microarray analysis. Plant Mol Biol **53:** 117–131