Cluster Analysis and Comparison of Various Chloroplast and Nuclear Transcriptomes in *Arabidopsis thaliana*

Dissertation

zur Erlangung des Doktorgrades der Fakultät für Biologie

der Ludwig-Maximilians-Universität München

vorgelegt von

Won Kyong Cho

aus Seoul, Korea

2007

Erstgutachter : PD Dr. J. Meurer

Zweitgutachter : Prof. Dr. R.G. Herrmann

Datum der mündlichen Prüfung: 22. 06. 2007

Contents

ABBREVIATIONS

µE	microeinstein (1 E = 1 mol of photons)
ATP	adenosine 5′-triphosphate
bp	base pairs
B light	blue light
cDNA	complementary DNA
Ci	curie
cpm	counts per minute
DNA	deoxyribonucleic acid
dCTP	2'-deoxycytidine 5'-triphosphate
dNTPs	deoxynucleoside triphosphates
EDTA	ethylenediaminetetraacetic acid
EMS	ethyl methanesulfonate
ESTs	expressed sequence tags
g	gravity force, gramme
FR light	far-red light
hcf	high chlorophyll fluorescence
kb	kilobases
Lowess	locally weighted regression
Mb	megabases
MES	2-Morpholinoethanesulfonic acid
MOPS	3-[N-Morpholino]propanesulfonic acid
mRNA	messenger RNA
MS	murashige and skoog medium
NCBI	national center for biotechnology information
NEP	nuclear encoded RNA polymerase
ORF	open reading frame
PAM	pulse amplitude–modulated fluorometer
PCR	polymerase chain reaction

PEP	plastid encoded RNA polymerase
Phy	phytochrome
PSI	photosystem I
PSII	photosystem II
qP	photochemical chlorophyll *a* fluorescence quenching
R light	red light
RNA	ribonucleic acid
rpm	revolutions per minute
rRNA	ribosomal RNA
RT-PCR	reverse transcription PCR
S	svedberg unit
SD	standard deviation
SDS	sodium dodecyl sulfate
SOM	self organizing map
T-DNA	transferred DNA
Tm	annealing temperature
Tris	tris-(hydroxymethyl)-aminomethane
tRNA	transfer RNA
U	unit, enzyme activity
UTR	untranslated region
UV	ultra violet
v/v	volume per volume
w/v	weight per volume

1 INTRODUCTION

1.1 Origin of the Plastid Genome

The chloroplast evolved as a result of an endosymbiotic event in which a cyanobacterial ancestor was taken over by a eukaryotic cell. Although most chloroplast genes have been lost or transferred to the nucleus and the majority of plastid proteins is encoded by nuclear genes, the organelle still retained the coding capacity for a number of genes and possesses its own gene expression machinery (Race et al., 1999) (Figures 1 and 2). Due to its endocytobiotic parentage, chloroplast gene expression represents a unique chimeric system assembled from multiple origins (Barkan and Goldschmidt-Clermont, 2000; Monde et al., 2000; Rochaix, 2001).

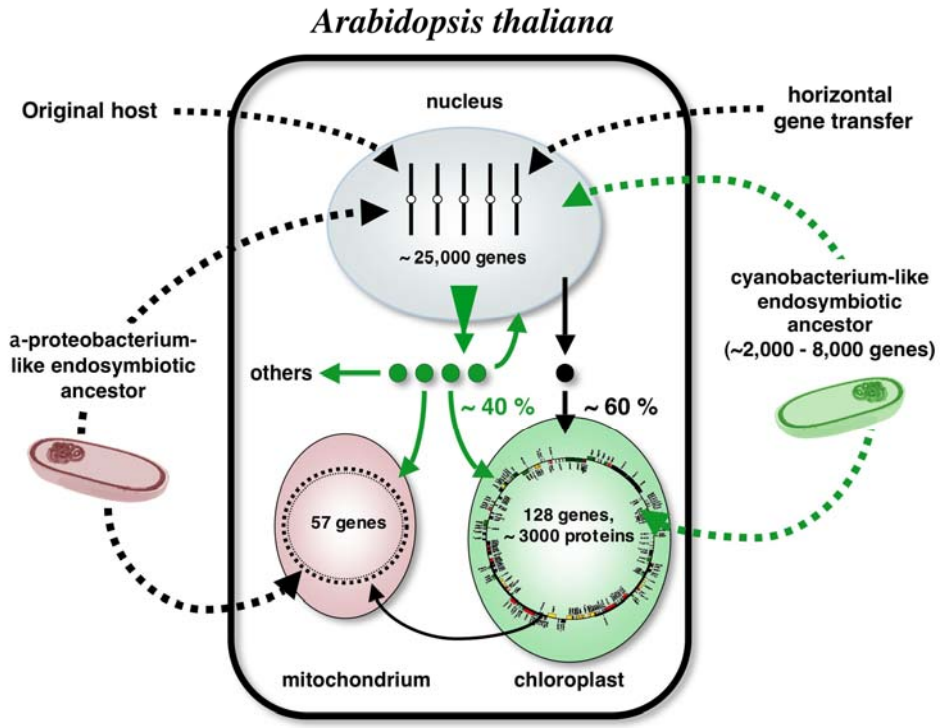


Figure 1. Endosymbiotic Rearrangement of Genes and Proteins in Plants.
The estimated numbers and the localisation of genes and proteins of endosymbiotic origin. Dotted lines, genes of endosymbiotic origin. Solid lines: localisation of nuclear encoded proteins. Green: genes and proteins of cyanobacterial origin; black: genes and proteins of other origin.

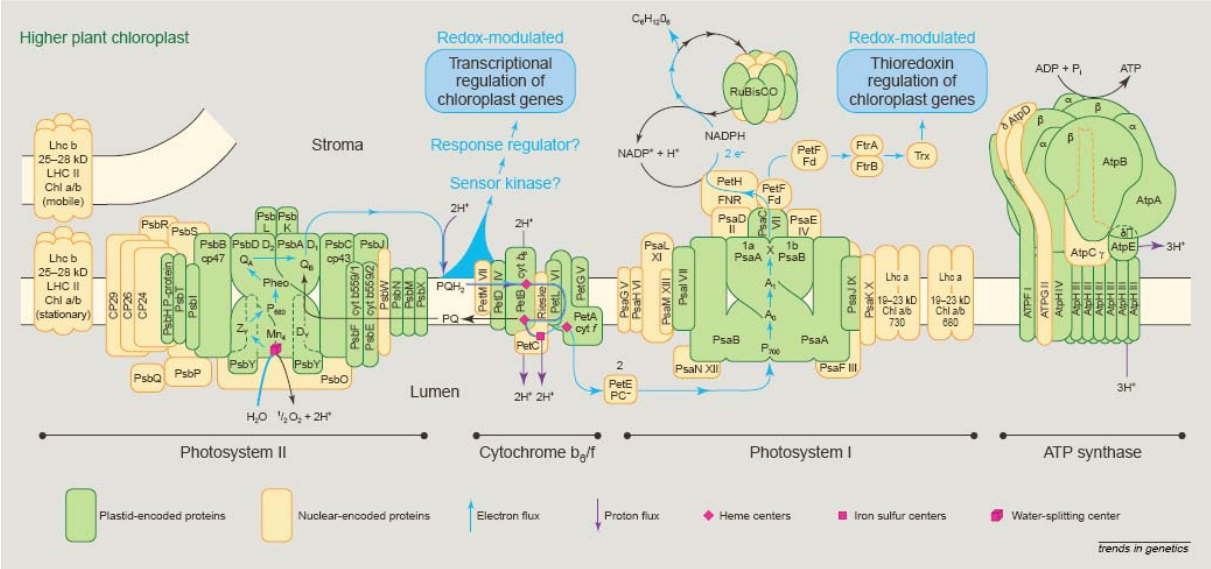


Figure 2. Functional Phylogenetic Model of the Photosynthetic Membrane.
The four major complexes are shown. A fifth minor abundant complex, the NADH-dehydrogenase complex, is also present in the thylakoid membrane (not shown). Yellow and green colours indicate nuclear and plastid encoded proteins for photosynthetic complexes (Race et al., 1999).

Plastid genes are embedded into regulatory networks that enable an adaptive and developmentally dependent chloroplast biogenesis at various levels (Figure 2; Table 1). An intriguing number of plastid transcriptional and posttranscriptional events were acquired in the result of endosymbiosis, which are not found at that extent in currently living cyanobacteria. Knowledge about this regulation is not only increasingly relevant for studying phylogenetic and ontogenetic aspects of chloroplast biogenesis but also for applying transplastomic approaches, since plastomes of several model plants as well as of agriculturally important plants will be genetically engineered in the near future (Bock, 2006). Numerous nuclear-encoded factors with yet unidentified functions play a crucial role in the regulation of the complex chloroplast transcript homeostasis.

1.2 Regulation of Plastid Gene Expression at the Transcriptional Level

Transcription rates depend on light, tissue- and cell-type. They also undergo endogenous circadian rhythms and possibly redox regulation (Allison, 2000). The activity of the plastid-encoded RNA polymerase (PEP) is regulated by nuclear-encoded sigma factors which are involved in the global and specific environmental and developmental dependent as well as tissue-specifc regulation of plastid transcription (Tanaka et al., 1996; Allison, 2000; Privat et al., 2003; Ichikawa et al., 2004; Favory et al., 2005; Zghidi et al., 2007). Sigma factors

analyzed so far are dispensable for photoautotrophic growth. Moreover, newly acquired nuclear-encoded phage-type RNA polymerases (NEP) are involved in the regulation of plastid transcription thereby increasing the complexity of the chloroplast transcript metabolism (Hedtke et al., 2002; Kanamura and Tanake, 2004; Emanuel et al., 2004; Liere et al., 2004). However, unlike eubacteria and the nuclear system in eukaryotes, chloroplast genes rarely undergo individual regulation at the transcriptional level (Shiina et al., 2005; Liere and Börner, 2006). The occurrence of different polymerases led to development of three promoter types: PEP (-10 and -35), consensus type (YRTA, class I, NEP), and exceptional promoters (class II) (Weihe and Börner, 1999; Shiina et al., 2005; Liere and Börner, 2006). The consensus type is mostly active in non-photosynthetic tissues and the other two are mainly active in green tissues (Liere and Maliga, 1999).

1.3 Regulation of Plastid Gene Expression at the Post-Transcriptional Level

An important characteristic of chloroplast gene regulation is the predominance of posttranscriptional control, which is exerted at both the gene-specific and global level. Steady state transcript levels are therefore mostly determined by fine tuning mRNA stability. Application of DNA arrays was demonstrated for studies on the chloroplast transcriptome in *Chlamydomonas* (Hihara et al., 2001; Erickson et al., 2005), tobacco (Legen et al., 2002; Nakamura et al., 2003), and *Arabidopsis* (Nagashima et al., 2004; Pfalz et al., 2006). These data revealed general and specific changes of the plastid RNA metabolism in response to mutations and environmental conditions. However, analysis of plastid gene expression in response to environmental and genetic signals was so far largely restricted to individual genes and mutations affected in chloroplast functions. Co-regulated expression of groups of plastid genes has not yet been intensively studied.

The effect of nuclear-encoded factors on chloroplast gene expression in higher plants is thought to be less gene-specific than in *Chlamydomonas,* which often leads to pleiotropic phenotypes of nuclear mutants (Barkan and Goldschmidt-Clermont, 2000). However, there is increasing evidence that nuclear genes act in a gene-specific manner in higher plants as well (Felder et al., 2001; Meurer et al., 2002; Meierhoff et al., 2003; Lezhneva and Meurer, 2004; Schmitz-Linneweber et al., 2006). Taken together, coordination of plastid and nuclear gene expression is necessary on both, the transcriptional and posttranscriptional level (Barkan and Goldschmidt-Clermont, 2000; Rochaix, 2001; Gray et al., 2002).

1.4 Microarray Techniques

Since the first microarrays containing only 45 *Arabidopsis thaliana* genes have been successufully used (Schena et al., 1995), rapid advances are being made to understand transcript regulation of several model organisms (Reinke White, 2002; Wellmer et al; 2005). Microarrays are simply a method for visualizing which genes are likely to be used in a particular tissue at a particular time under a particular set of conditions. The output of a microarray experiment is called a gene expression profile (Gibson, 2003). Microarrays can be divided cDNA microarray and oligonucleotide microarray by their probe type. cDNA microarrays are capable of profiling gene expression patterns of tens of thousands of genes in a single experiment. DNA targets, in the form of 3´ expressed sequence tags (ESTs), are arrayed onto glass slides (or membranes) and probed with fluorescent- or radioactively-labelled cDNAs (Figure 3) (Duggan et al., 1999). Especially, cDNAs are labelled with radioactivity, they are called as Macroarrays.

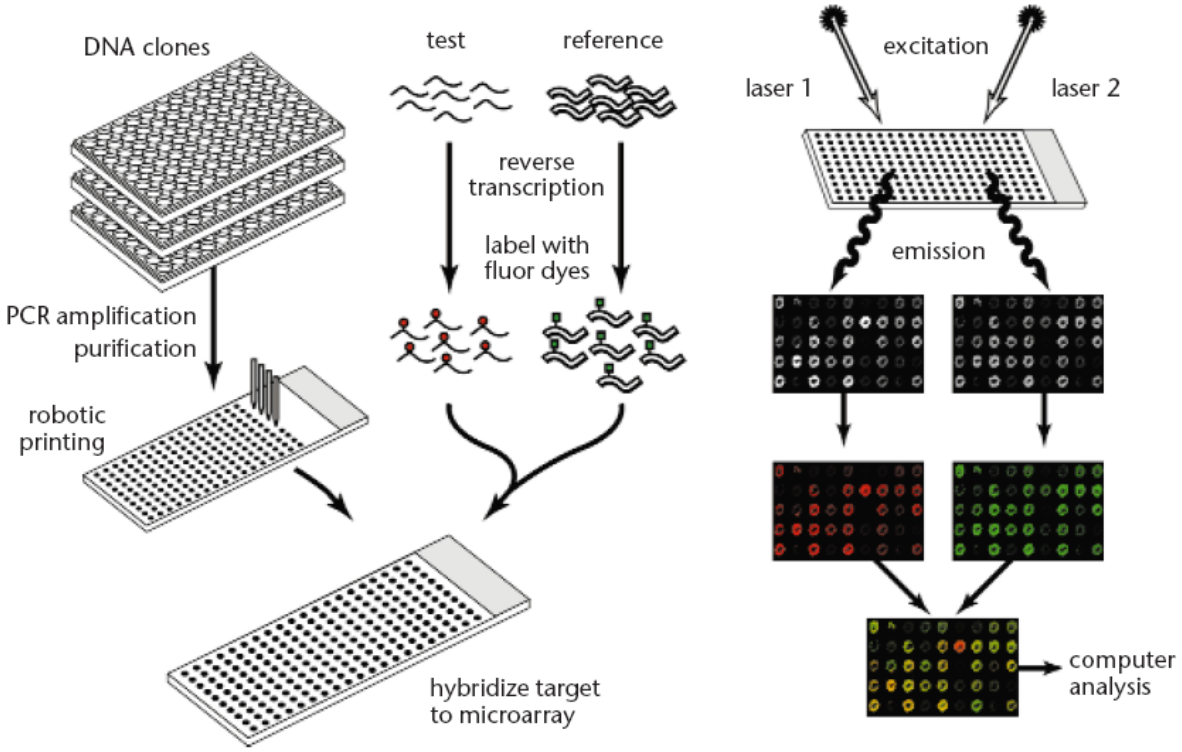


Figure 3. The cDNA Microarray Scheme.
DNA clones are spotted on microscope slides using a robot. After hybridization with differentially labelled cDNAs or RNAs the slides are scanned using laser excitation. The two images obtained from the test and the reference on the same slide are merged and used for further bioinformatics analysis (Duggan et al., 1999).

One of most favourite oligonucleotide microarray is Affymetrix DNA chips. The oligonucleotide arrays, developed by the Affymetrix Company, are a new approach in microarray technology, based on hybridization to small, high-density arrays containing tens of thousands of synthetic oligonucleotides. The arrays are designed based on sequence information alone and are synthesized in situ using a combination of photolithography and oligonucleotide chemistry. RNAs present at a frequency of 1:300,000 are unambiguously detected, and detection is quantitative over more than three orders of magnitude. This approach provides a way to use directly the growing body of sequence information for highly parallel experimental investigations. Because of the combinatorial nature of the chemistry and the ability to synthesize small arrays containing hundreds of thousands of specifically chosen oligonucleotides, the method is readily scalable to the simultaneous monitoring of tens of thousands of genes. The Affymetrix integrated GeneChip arrays contain up to 500,000 unique probes corresponding to tens of thousands of gene expression measurements (Figure 4).

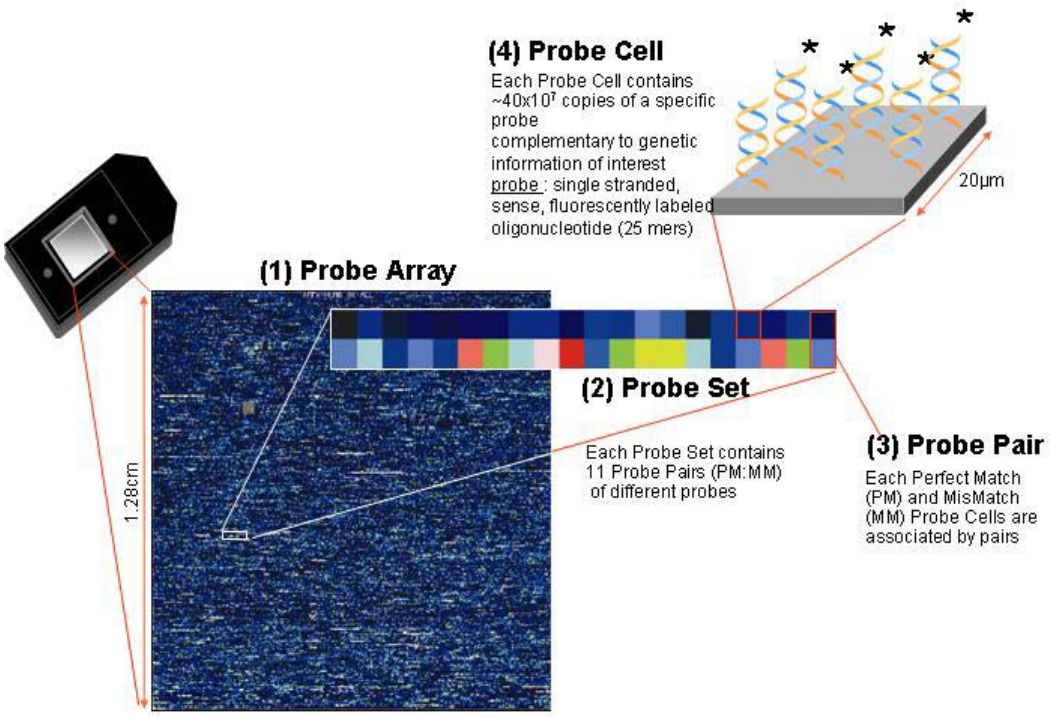


Figure 4. A Typical Experiment with an Oligonucleotide Chip; Preparation of Sample for Genechip Arrays.
Messenger RNA (mRNA) is extracted from the cell and converted to cDNA. It then undergoes amplification and labeling step before fragmentation and hybridization to 25-mer oligos on the surface of the chip. After washing of unhybridized material, the chip is imaged by a confocal laser scanner and the hybridisation signals are further analyzed by computer programs (http://cnx.org/content/m12387/latest/).

1.5 Gene Expression Analysis in Higher Plants Using Microarray Technique

Along with the complete sequence of the chondriome, plastome and nuclear genome of *Arabidopsis thaliana* (Unseld et al., 1997; Sato et al., 1999; The *Arabidopsis* genome initiative, 2000), large-scale expression data from various array analyses provide a useful tool for exploring and comparing the three genetic compartments at the expression level (Rensink et al., 2005). Nuclear gene expression in higher plants is regulated mostly at transcriptional level. Genome wide transcript profiles have been monitored in previous investigations to study the response to endogenous and exogenous signals often by using mutants deficient in signaling pathways, like those affected in the photoreceptors, photomorphogenesis, and retrograde signaling (Chory et al., 1989; Tepperman et al., 2001, 2004, 2006; Ma et al., 2003; Ohgishi et al., 2004; Mao et al., 2005) These data revealed that generally the expression of several hundred genes is significantly changed indicating a complex gene regulatory network. Surprisingly, macroarray analysis of 3289 nuclear *Arabidopsis* genes mostly encoding chloroplast proteins revealed a regulatory master switch showing either a preferential up- or down-regulation of most genes (Richly et al., 2003; Biehl et al., 2005). In contrast to in-depth studies on nuclear gene expression relatively little genome wide information for regulation of the plastid genetic compartment is available. Since expression of nuclear and chloroplast genes has to be tightly coupled comparative array based analysis of both genomes may provide a framework for the understanding of the integrated gene regulatory network.

1.6 Aim of the Project

In order to elucidate regulatory dynamics and clustering of chloroplast gene expression, a macroarray system subjected to 78 nuclear mutants affected in a variety of plastid functions and to the wild type (WT) under diverse environmental conditions had to be established. The utility of various computer programs for the genome wide expression analysis should be proven. The data should be verified by representative RNA gel blot analysis. Furthermore, the photosynthetic performance and the phenotype of the mutants available in the collection of PD Dr. J. Meurer should be described.

In addition, the data obtained should be compared with those extracted from public available affymetrix 22K ATH1 expression dataset of all plastid coding genes and 1,430 nuclear genes encoding bona fide chloroplast proteins in *Arabidopsis* (Biehl et al., 2005). In

average, 83 biological conditions and 136 mutants should be taken into consideration. Distinct expression signatures and clusters of coexpressed plastid genes that are potential targets for the concerted nuclear control should be identified using an array-based approach. Specific responses of the chloroplast and nuclear transcriptomes could provide insights into the coordinated regulation of several defined clusters of nuclear and plastid genes, even if they are not co-transcribed. The identification of co-regulated genes may point to the integration of gene expression into common pathways and to a concerted response. These data represent the first comprehensive cluster analysis of the chloroplast transcriptome in higher plants and potentially indicate a crucial role of the chloroplast to cope with abiotic as well as biotic stress conditions.

Table 1. Functional Classification of *Arabidopsis thaliana* Plastid Genes.
Plastid genes are classified by their function according to NCBI gene annotation.

Classification	Number	Gene	Description
ATP synthase	6	*atpA*	ATPase alpha subunit
		atpB	ATP synthase CF1 beta chain
		atpE	ATP synthase CF1 epsilon chain
		atpF	ATP synthase CF0 B chain
		atpH	ATP synthase CF0 C chain
		atpI	ATP synthase CF0 A chain
NADH dehydrogenase	11	*ndhA-K*	NADH dehydrogenase subunit A-K
Cytochrome b_6f	7	*orf31*	cytochrome b_6/f complex subunit VI
		petA	cytochrome f
		petB	cytochrome b_6
		petD	cytochrome b_6/f complex subunit IV
		petG	cytochrome b_6/f complex subunit V
		ycf5	cytochrome c biogenesis protein
		ycf6	cytochrome b_6/f complex subunit VIII
Photosystem I	7	*psaA*	photosystem I P700 apoprotein A1
		psaB	photosystem I P700 apoprotein A2
		psaC	photosystem I subunit VII
		psaI	PSI I subunit VIII
		psaJ	photosystem I subunit IX
		ycf3	photosystem I assembly protein YCF3
		ycf4	photosystem I assembly protein YCF4
Photosystem II	15	*psbA, B, C, D, E, F, H, I, J, K, L, M, N, T*	photosystem II protein A, B, C, D, E, F, H, I, J, K, L, M, N, T
Ribosomal protein	21	*rps11,12, 14, 15, 16, 18, 19, 2, 3, 4, 7, 8*	ribosomal protein small subunit 11, 12, 14, 15, 16, 18, 19, 2, 3, 4, 7, 8
		rpl14, 16, 2, 20, 22, 23, 32, 33, 36	ribosomal protein large subunit 14, 16, 2, 20, 22, 23, 32, 33, 36
RNA polymerase	4	*rpoA*	RNA polymerase alpha chain
		rpoB	RNA polymerase beta chain
		rpoC1	RNA polymerase beta' chain
		rpoC2	RNA polymerase beta'' chain

Table 1. Continued

Classification	Number	Gene	Description
Ribosomal RNA	4	*rrn16S*	16S ribosomal RNA
		rrn23S	23S ribosomal RNA
		rrn4.5S	4.5S ribosomal RNA
		rrn5S	5S ribosomal RNA
Transfer RNA	21	*trnA*	Ala tRNA
		trnC	Cys tRNA
		trnD	Asp tRNA
		trnE	Glu tRNA
		trnF	Phe tRNA
		trnfM	Met tRNA
		trnG	Gly tRNA
		trnH	His tRNA
		trnI	Ile tRNA
		trnK	Lys tRNA
		trnL	Leu tRNA
		trnM	Met tRNA
		trnN	Asn tRNA
		trnP	Pro tRNA
		trnQ	Gln tRNA
		trnR	Arg tRNA
		trnS	Ser tRNA
		trnT	Thr tRNA
		trnV	Val tRNA
		trnW	Trp tRNA
		trnY	Tyr tRNA
Other	4	*accD*	acetyl-CoA carboxylase beta subunit
		clpP	ATP-dependent Clp protease proteolytic subunit
		matK	maturase
		rbcL	riblose 1,5-bisphosphate carboxylase/oxygenase large chain
Unknown	5	*ycf1*	hypothetical protein
		ycf10/cemA	hypothetical protein
		ycf2	hypothetical protein
		ycf9	hypothetical protein
		orf77	hypothetical protein
Total	104		

2 MATERIALS AND METHODS

2.1 Materials

2.1.1 Chemicals, Enzymes, Radioactive Substances, and Devices

Chemicals used in this work were usually of p.a. quality and, if not mentioned, were purchased from the following companies: Applichem (Darmstadt, Germany), Biozym (Oldendorf, Germany), Fluka (Steinhein, Germany), ICN Biomedicals GmbH (Eschwege, Germany), Merck (Darmstadt, Germany), Pharmacia (Uppsala, Sweden), Roth (Karlsruhe, Germany), Serva (Heidelberg, Germany), Sigma-Aldrich Chemie GmbH (Taufkirchen, Germany), and USB (Cleveland, USA).

Enzymes were obtained from Clontech (Palo Alto, USA), Invitrogen (Karlsruhe, Germany), MBI Fermentas (St. Leon-Rot, Germany), New England Biolabs (Frankfurt/Main, Germany), Promega (Mannheim, Germany), Qiagen (Hilden; Germany), Roche Diagnostics (Mannheim, Germany), and Stratagene (Heidelberg, Germany).

[α-32P] dCTP Radioactive nucleotides were obtained from Hartmann analytic GmbH (Braunschweig, Germany).

Other materials were obtained from Biomol (Hamburg, Germany), Eppendorf (Hamburg, Germany), Greiner Bio-One GmbH (Frickenhausen, Germany), Millipore (Eschborn, Germany), Pall Bio Support Division (Dreieich, Germany), Qiagen (Hilden, Germany), and Schleicher and Schüll (Dassel, Germany). The manufacturers of commercial devices are mentioned in the text.

2.1.2 Source of *Arabidopsis* Mutants, Phenotypes, and Growth Conditions

Arabidopsis mutant plants were characterized by the appearance of their color and maximum photosystem II quantum yield, (Fv/Fm) (Schreiber, 1986). These parameters, accessions, as well as references of EMS-induced and T-DNA insertion lines used in this work are described in Table 2. WT and mutant seeds were surface sterilized before plating on MS medium (Murashige and Skoog, 1962) supplemented with 15 g sucrose/L. Followed by cold treatment (4°C) for 2 days in darkness, plants were grown in climate chamber under continuous light (60 μmol photons m^{-2} s^{-1}).

Table 2. Names, Growth Conditions, Phenotypes, and Photosynthetic Parameter Fv/Fm of Mutants as Well as Biological Conditions Used.

Mutant	Growth conditions	Phenotype	Fv/Fm	Control
ins1	grown on MS medium	pale green, hcf	0.677	WT
ins2	grown on MS medium	pale green, hcf	0.682	WT
ins3	grown on MS medium	albino	0.75	WT
ins4	grown on MS medium	albino	0.752	WT
ins5	grown on MS medium	albino	0.207	WT
ins6	grown on MS medium	albino	0.207	WT
ins7	grown on MS medium	pale green, hcf	0.755	WT
ins8	grown on MS medium	pale green	0.72	WT
ins9	grown on MS medium	pale green	0.715	WT
ins10	grown on MS medium	pale green	0.611	WT
ins11	grown on MS medium	albino	0.111	WT
ins12	grown on MS medium	pale green, hcf	0.414	WT
ins13	grown on MS medium	pale green	0.662	WT
ins14	grown on MS medium	albino	0.2	WT
ins15	grown on MS medium	pale green	0.784	WT
ins16	grown on MS medium	albino	0.575	WT
ins17	grown on MS medium	pale green, hcf	0.68	WT
ins18	grown on MS medium	albino	0.2	WT
ins19	grown on MS medium	pale green	0.288	WT
ins20	grown on MS medium	albino	0.13	WT
ins21	grown on MS medium	albino	0.15	WT
ins22	grown on MS medium	albino	0.188	WT
ins23	grown on MS medium	albino	0.078	WT
ins24	grown on MS medium	albino	0	WT
ins25	grown on MS medium	albino	0	WT
ins26	grown on MS medium	albino	0	WT
ins27	grown on MS medium	albino	0	WT
ins28	grown on MS medium	albino	0	WT
ins29	grown on MS medium	pale green	0.73	WT
ins30	grown on MS medium	pale green	0.65	WT
hcf102	grown on MS medium	hcf	0.314	WT
crp135	grown on MS medium	hcf	0.48	WT
ins31	grown on MS medium	pale green, hcf	0.451	WT
ins32	grown on MS medium	pale green, hcf	0.6	WT
ins33	grown on MS medium	albino	0	WT
ins34	grown on MS medium	albino	0	WT
ins35	grown on MS medium	pale green	0.793	WT
ins36	grown on MS medium	pale green	0.578	WT
ins37	grown on MS medium	pale green	0.598	WT
ins38	grown on MS medium	albino	0	WT

Table 2. Continued

Mutant	Growth conditions	Phenotype	Fv/Fm	Control
ins39	grown on MS medium	albino	0	WT
ins40	grown on MS medium	pale green, hcf	0.489	WT
crp102	grown on MS medium	hcf	0.53	WT
hcf145	grown on MS medium	hcf	0.65	WT
atprfB1-1	grown on MS medium	hcf	0.18	WT
pac	grown on MS medium	pale	0.12	WT
atprfB1-2	grown on MS medium	hcf	0.2	WT
crp160	grown on MS medium	hcf	0.56	WT
ntt1	grown on soil	4 days old *ntt1* mutant	ND	WT
ntt2	grown on soil	4 days old *ntt2* mutant	ND	WT
RNAi-ntt	grown on soil	4 days old *rnai* mutant of ntt	ND	WT
nm-ntt	grown on soil	4 days old *ntt* double mutants	ND	WT
ntt1	grown on soil	4 days old *ntt1* mutant	ND	WT
ntt2	grown on soil	4 days old *ntt2* mutant	ND	WT
RNAi-ntt	grown on soil	4 days old *rnai* mutant of ntt	ND	WT
nm-ntt	grown on soil	4 days old *ntt* double mutants	ND	WT
ntt1	grown on soil	6 weeks old *ntt1* mutant	ND	WT
ntt2	grown on soil	6 weeks old *ntt2* mutant	ND	WT
RNAi-ntt	grown on soil	6 weeks old *rnai* mutant of ntt	ND	WT
nm-ntt	grown on soil	6 weeks old *ntt* double mutants	ND	WT
hcf203	grown on MS medium	hcf	ND	WT
hcf208	grown on MS medium	hcf	ND	WT
hcf210	grown on MS medium	hcf	ND	WT
hcf219	grown on MS medium	hcf	ND	WT
hcf221	grown on MS medium	hcf	ND	WT
hcf222	grown on MS medium	hcf	ND	WT
hcf229	grown on MS medium	hcf	ND	WT
hcf232	grown on MS medium	hcf	ND	WT
hcf240	grown on MS medium	hcf	ND	WT
vipp-w	grown on MS medium	3 weeks old *vipp1* has albino phenotype		WT
vipp-g	grown on MS medium	3 weeks old *vipp1* has pale green phenotype		WT
vipp-r	grown on MS medium	3 weeks old *vipp1* which are rescued		WT
sol8-10d	grown on MS medium	pale green, 10 days old *soldat8*		WT
sol8-4d	grown on MS medium	pale green, 4 days old *soldat8*		WT
alb3	grown on MS medium	albino		WT

Table 3. Use of three different tissues and description of stress conditions.

Experiments	Growth conditions	Treatments	Comparison
Dark/Light	MS medium	harvest 3 weeks plant after 8 hours light in long day condition	harvest 3 weeks plant after 4 hours dark in long day condition
Without Sugar/With Sugar	MS medium	1 week plant grown on MS medium without sugar	1 week plant grown on MS medium with sugar
Cytokinin	MS medium	2 weeks old plants transferred to MS medium with cytokinin (0.5 mg/l) for 1 week	2 weeks old WT plants
Auxin	MS medium	2 weeks old plants transferred to MS medium with auxin (0.5 mg/l) for 1 week	2 weeks old WT plants
GA	MS medium	2 weeks old plants transferred to MS medium with GA (0.5 mg/l) for 1 week	2 weeks old WT plants
ABA	MS medium	2 weeks old plants transferred to MS medium with ABA (0.5 mg/l) for 1 week	2 weeks old WT plants
Flowers/Leaves	MS medium	flowers from 6 weeks old plants	leaves from 6 weeks old plants
Stems/Leaves	MS medium	stems from 6 weeks old plants	leaves from 6 weeks old plants
DCMU 50µm	MS medium	2 weeks old plants transferred to MS medium with DCMU (50 µM) for 1 week	2 weeks old WT plants
MV 50µm	MS medium	2 weeks old plants transferred to MS medium with Methyl viologen (50 µM) for 1 week	2 weeks old WT plants
MV 2µm	MS medium	2 weeks old plants transferred to MS medium with Methyl viologen (2 µM) for 1 week	2 weeks old WT plants
Cold Stress	MS medium	2 weeks old plants transferred to 4 degrees for 3 days	2 weeks old WT plants
Light Stress	MS medium	2 weeks old plants transferred to high light (800 µE) for 3days	2 weeks old WT plants
Heat Stress	MS medium	2 weeks old plants transferred to 32 degrees for 3days	2 weeks old WT plants

2.1.3 Media, Solutions and Buffers

TE buffer: 10 mM Tris-HCl, pH 8.0
1 mM Na_2EDTA

10x TBE buffer: 108 g/l Tris, pH 8.2-8.4

55 g/l M boric acid
7.4 g/l Na_2EDTA

MOPS buffer: 20 mM MOPS
5 mM Na-acetate
1 mM Na_2EDTA, pH 7.0

20x SSC buffer: 3 M NaCl
0.33 M Na-citrate, pH 7.0

MS-medium: 1x MS-salts (Murashige and Skoog, 1962)
1.5% sucrose
2.5 mM MES-NaOH, pH 5.7
0.3% gelrite

Hybridisation buffer: 250 mM Na_2HPO4, pH 7.2
7% (w/v) SDS
2.5 mM EDTA (Church and Gilbert, 1984)

Washing solution 1: 2.5x SSC 1% SDS
Washing solution 2: 1.0x SSC 1% SDS
Washing solution 3: 0.5x SSC 1% SDS
Washing solution 4: 0.2x SSC 1% SDS

Solutions and buffers which are not mentioned otherwise were prepared as described in Sambrook et al., (1989).

2.1.4 Softwares

AIDA Image Analyzer (3.52) and Aida Array Compare (4.15) (http://www.raytest.de)
R program (http://www.r-project.org)
Simpleaffy (Wilson et al., 2005) (http://bioinformatics.picr.man.ac.uk/index.jsp)
R/MAANOVA version 0.98.8 (Wu et al., 2002)
(http://www.jax.org/staff/churchill/labsite/software/Rmaanova/index.html)
Microarray analysis toolbox (http://nbc11.biologie.uni-kl.de)

Genesis (Sturn et al., 2002) (http://genome.tugraz.at)

Genevestigator (http://www.genevestigator.ethz.ch)

AtGenExpress (http://www.arabidopsis.org/info/expression/ATGenExpress.jsp)

TAIR (http://www.arabidopsis.org)

Athena (http://www.bioinformatics2.wsu.edu/cgi-bin/Athena/cgi/home.pl)

FunCat (http://mips.gsf.de/projects/funcat)

Photoshop 7.0 (http://www.adobe.com)

Microsoft Excel 2002 (http://www.microsoft.com/excel)

Spotfire (http://www. spotfire.com)

2.1.5. Oligonucleotides used for the generation and spotting macroarray gene probes

Oligonucleotides used in the present work are indicated in Table 4.

Table 4. Gene Name, Primer-sequcences, Position of 3' End, Length of Primer, Annealing Temperature (Tm), and Length of PCR Product Are Described.

Gene Name	Primer-sequence	Positioin of 3'end	Length	Tm °C	Length of PCR product
accD.f	**GGATGAAGACATGGTCTCTGCGG**	**844**	**23**	**67**	**421**
accD.r	**ACCAAAACTAGCTGTCACTCCACC**	**1219**	**24**	**67**	
atpA.f	**TTACAGGAACGAGGGGCAATGG**	**866**	**22**	**68**	**330**
atpA.r	**GGTCTCGACGATTGGTAAGGCAG**	**580**	**23**	**68**	
atpB.f	**GGTCTGACTGCCTTAACTATGGC**	**728**	**23**	**67**	**399**
atpB.r	**CTCGCCAACGATTCGAGGTTGTAG**	**375**	**24**	**67**	
atpE.f	**TTAGCAAACCACGCGCCGATTG**	**279**	**22**	**67**	**266**
atpE.r	**GCTCGTCTGAGAGCTAGATTAGCC**	**58**	**24**	**68**	
atpF.f	**GCCGGGAGTTTCGGATTTAATACC**	**483**	**24**	**66**	**223**
atpF.r	**CACGCGAAACTTATCCGCTTCCG**	**286**	**23**	**67**	**958**
atpH.f	**ATCCACTGGTTTCTGCTGCTTCG**	**220**	**23**	**67**	**136**
atpH.r	**CCTTCTGCCTCAGGTTGTCTCG**	**128**	**22**	**68**	
atpI.f	**TTTCCAGGTCCATGCCCAAGTTC**	**627**	**23**	**67**	**300**
atpI.r	**TTCGTTGGTGCTGCTAACTCCC**	**371**	**22**	**67**	
clpP.f	**GCCTATTGGCGTTCCAAAAGTACC**	**566**	**24**	**67**	**364**
clpP.r	**CCTAGCGTGAGGGAATGCTATACG**	**249**	**24**	**68**	**879**
matK.f	**CTGGCAACAAAGGATACGCCGC**	**645**	**22**	**69**	**333**
matK.r	**CCGTCCAGGTTGCTTTACTAATGGG**	**360**	**25**	**69**	
ndhA.f	**TTCAAGCTCTAGCCGATGGGAC**	**845**	**22**	**67**	**416**
ndhA.r	**TGACGCCACAAATTCCATCCCC**	**473**	**22**	**67**	
ndhB.f	**TAGCTGCTTCAGCTTCAGCCAC**	**698**	**22**	**68**	**413**
ndhB.r	**TAGTGGAGGAAGACCTCCTAGG**	**328**	**22**	**68**	
ndhC.f	**ATCTCCAATTAGGAAGGGGCCG**	**250**	**22**	**68**	**260**
ndhC.r	**TCCAATGCTCCTTTTCGCCATGC**	**33**	**23**	**67**	
ndhD.f	**AGCGGCTTTTCCAGTTACTCGG**	**1177**	**22**	**67**	**427**
ndhD.r	**ACTGTAATGTGCCTCACCGTGG**	**793**	**22**	**67**	
ndhE.f	**GATCACAAGTCGAAATATGGTTAGGGCTC**	**216**	**29**	**68**	**175**
ndhE.r	**AGCTAATCCAATAGCTGCTTCAGCG**	**94**	**25**	**68**	

Table 4. Continued

Gene Name	Primer-sequence	Positioin of 3'end	Length	Tm °C	Length of PCR product
ndhF.f	TGTTAGCTCTAGGTATGGGGTC	1229	22	66	338
ndhF.r	TGCGGTTAATCCCGCTGTTGAAC	935	23	67	
ndhG.f	GGTCTGGGAGTGGTATTACTTCC	443	23	67	304
ndhG.r	TAACCCCGTACCATGACGTATCGAG	186	25	68	
ndhH.f	TTCGTATTGGAGGAATAGCGGCTG	687	24	67	377
ndhH.r	GGAAGTCCTTCCAGAGCCTGTTG	357	23	68	
ndhI.f	TGGTCAACAAACCCTACGAGCTG	465	23	67	414
ndhI.r	ACTGACATTGGTAAACGACCCAAAGC	99	26	67	
ndhJ.f	GCGGGCTGGTTCATAGATCGTTG	421	23	68	359
ndhJ.r	GATCCGTTTCAGTCGTGGATGGC	107	23	67	
ndhK.f	CAGCCTATGGCCGCTTCTTTATG	561	23	67	336
ndhK.r	TCTGGTTTAGGTGGACAACCCG	269	22	67	
orf31/petL	TCGGTTTTCTACTAGCAGCTTTAAC	47	25	63	56
orf31/petL	GCTTAGACCAATAAACAGAACTGAGG	53	26	64	
orf77.fycf15	ACACTATGTATGGATGGTATGAACTGCC	71	28	67	160
orf77.rycf15	GATAGCAACAACAATTTAATCCGACATGCG	174	30	65	
petA.f	GGAGGGTTTTGAATTAGCCCCC	360	22	68	307
petA.r	ACTTATTATCCCTCCTGCCGTAGC	622	24	68	
petB.f	ACCGTCCAACCGTTACTGAAGC	194	22	67	318
petB.r	GGATTGTCCAACACTAGCACTTCC	511	24	66	
petD.f	TTATTACGGGGAACCCGCATGG	96	22	68	324
petD.r	AAGACGGTTGTCGCTACTGGAC	377	22	67	
petG.f	TGGAATCGTCTTAGGTCTAATTCCT	45	25	63	77
petG.r	CACCACGTCTGTATTGTAAATAAGCAG	71	27	63	
psaA.f1	ACCACGCCCGCTGAATAGAAAC	205	22	67	355
psaA.f2	AATGGCTGCATGCGCAGCAG	1049	20	68	399
psaB.f1	CGAGGTGGTACTTGTGATATTTCGGC	491	26	67	333
psaB.f2	GATCCTCATTTTGGTCAACCGGCTG	1905	25	68	317
psaC.f	AGGATGTACTCAATGTGTCCGAGC	188	24	66	144
psaC.r	GTTGGACAGGCGGATTCACATCTC	91	24	67	
psaI.f	GACAACTTTCAATAACTTACCCTCT	27	25	60	87
psaI.r	TGAAGAAATAAAGAAGCCATTGC	67	23	58	
psaJ.f	CATATCTTTCCGTAGCACCG	36	20	60	62
psaJ.r	TAATAAACCTGCTAACGAACCG	57	22	60	
psbA.f	TGGCTATACAACGGCGGTCC	731	20	68	
psbA.r	AAGGGAGCCGCCGAATACAC	452	20	68	318
psbB.f	TTGTTCCGGGAGGAATAGCCTC	599	22	68	301
psbB.r	GCCAGTCCAGCACTAACTCTTCG	856	23	68	
psbC.f	GCGTATATGATACCTGGGCTCC	572	22	67	432
psbC.r	CGTTAGCTCCAAGACGTTGGTCTC	959	24	67	
psbD.f	CTATGGGCTTTTGTTGCTCTCCAC	354	24	67	316
psbD.r	CATGAATAGCGCATAGCAGAGCC	624	23	67	
psbE.f	TGTCTGGAAGCACAGGAGAAC	231	21	64	191
psbE.r	TGGAATGCCTTGTCGGCTCTC	81	21	67	
psbF.f	CCTATCCAATTTTTACAGTGCGCTGG	79	26	66	89
psbF.r	GAACTGCATTGCTGATATTGACCCCA	38	26	67	
psbH.f	GTCCAAGAAGCACTACTGTAGGG	60	23	67	98
psbH.r	AACACCCATCAAAGGGGTCGTTC	113	23	67	
psbI.f	ACTGTAGTTATATTCTTTGTTTCTCTCTTC	57	30	60	68
psbI.r	GGATTACGTCCTGGATCATTAG	74	22	61	
psbJ.f	CTGGAAGGATTCCTCTTTGGG	90	21	64	110
psbJ.r	CTACAGGGATGAACCTAATCCTG	23	23	64	

Table 4. Continued

Gene Name	Primer-sequence	Positioin of 3'end	Length	Tm °C	Length of PCR product
psbK.f	TTTAGTCGCCAAATTGCCAGAGG	85	23	65	118
psbK.r	AAAACTTACAGCGGCTTGCCAAAC	157	24	66	
psbL.f	ACACAATCAAATCCGAACGAACAAAGTGT	86	29	64	85
psbL.r	CAGCAAGTACAAAAATGAGTAATAACCCC	58	29	64	
psbM.f	ATATTCTTGCATTTATTGCTACTGCACT	68	28	62	57
psbM.r	GCAAAAAAGCAGTAGGAACGAGAATG	64	26	63	
psbN.f	GAAACAGCAACCCTAGTCGCCATC	106	24	68	95
psbN.r	TGTTGAGAGGGTTGCCCAAAGG	56	22	68	
psbT.f	TCTCTTAGTATCCACTTTAGGG	45	22	60	56
psbT.r	GTGGTTCCCGAAAAAAAATAGCG	57	23	61	
rbcL.f	CTAGAGGATCTGCGAATCCCTCC	425	23	68	434
rbcL.r	CTAGTATTTGCGGTGAATCCCCC	814	23	67	
rpl14.f	CAGACAACAGCGGGGCTAGAGAA	316	23	69	289
rpl14.r	TGTCTCAATTCCCGTGGGATCGC	69	23	68	
rpl16.f	CTCGTGGGAATCGTATTTGTTTTGGCAG	317	28	67	279
rpl16.r	TTGCTATATTTTCGGGTACACCACCC	91	26	68	
rpl2.f	GGGGAGGTCATAAGCGTCTATAC	646	23	67	333
rpl2.r	CAATCAGTTTCGCTACAGCACCCG	359	24	68	
rpl20.f	AGCTCGGAGGCGTAGAACAAAAC	306	23	67	196
rpl20.r	TACCCCCATTTCGTGAATTACGGC	156	24	67	
rpl22.f	TCTGCCCATAAAGCACGGAGAG	396	22	67	362
rpl22.r	CTGTTTTCATTGCTGCATCCAGG	78	23	64	
rpl23.f	TCGAATCAGGATCAACTAGGACAG	183	24	63	110
rpl23.f	CATTCTTTTAACCTTTCCGGGGAG	120	24	64	
rpl32.f	GGCAGTTCCAAAAAAACGTACTTCTATCTC	32	30	65	120
rpl32.r	CCTGTAGAAAGTGATTTCCCTAATG	98	25	61	
rpl33.f	GCCAAGGGTAAAGATGTTCGAG	25	22	64	193
rpl33.r	TCTTGATTTCCCCGTGAATTGTATG	172	25	62	
rpl36.f	TGAAAATAAGGGCTTCCGTTCGTA	90	24	63	94
rpl36.r	TGCCTCGGGTTGGAACAAATTACTAT	45	26	66	
rpoA.f	CGCTTATTTGTGTCCAAGGTCCTGG	650	25	68	343
rpoA.r	CGGGAGGCTTGATGAAGTGCTTC	354	23	68	
rpoB.f1	CCCTCACGCATGAATGTAGGACAG	619	24	68	361
rpoB.f2	ACGTATTCGCTCTGTAGCGGATC	2262	23	66	343
rpoC1.f	TTCTTCCTCCCGAGTTGAGACC	1211	22	66	357
rpoC1.f	GTGAAGGTCCAACGACAATCACCG	899	24	67	
rpoC2.f1	CGGTGATATAACCCAAGGTCTTCC	706	24	66	371
rpoC2.f2	ATAGAAAGGCAGGATGCCCGTG	3412	22	68	380
rps11.f	TAACCGTTACAGATGTACGGGGTC	273	24	67	260
rps11.r	ATGTGGCATAGGGGTTACATCTCG	60	24	67	
rps12.f	TGCCAGAGTACGATTAACCTCGG	195	23	67	171
rps12.r	TCCTTTACTCCGACAGCATCTAGG	70	24	67	
rps14.f	TCCGTCGCTAAGTGAGAAATGG	181	22	64	179
rps14.r	CTGGCAACAAACATGCCTGAAC	45	22	64	
rps15.f	TGAAGAACAAAAAGAAGAAAGCAGGGG	215	27	64	138
rps15.r	GAATTTTTCGTAGACCCCGCTGAG	128	24	65	
rps16.f	CGATGTGGTAGAAAGCAACGAGCC	196	24	68	181
rps16.r	GAGCTGTTCTTGTTGGTTGAGCTCC	63	25	68	
rps18.f	AGCGACTTTTTACTAAATCCAAGCG	38	25	62	156
rps18.r	TTAAAGTCACTCTATTCACCCGTC	146	24	62	
rps19.f	GGGCATCCACCATTATACCCAC	152	22	68	75
rps19.r	AAACGGGTAAGTGTTCCCTTCC	120	22	65	

Table 4. Continued

Gene Name	Primer-sequence	Positioin of 3'end	Length	Tm °C	Length of PCR product
rps2.f	TGGAAGAGATGATGAGAGCAGGAG	660	24	66	373
rps2.r	GGAAGACGGTTGAATCCCTCTGTC	334	24	67	
rps3.f	AGAATATCCTCTGGTATGGAGGG	482	24	65	362
rps3.r	CTTTTCCGTCGATACGTCCTGC	165	22	65	
rps4.f	GGCAGGAAGCGATCTTAGAAACC	504	23	66	320
rps4.r	CTCGGGGTTTGCAGCGATAACTTG	230	24	68	
rps7.f	ACGTCAAGCAATACGTGGAGTAACTC	258	26	66	167
rps7.r	AAGCCATATTTCGACCCGGACG	138	22	67	
rps8.f	GGAACGGTTCGGATAGGATCGAC	320	23	67	302
rps8.r	AGTCGAGCTTCTCGGTCTGTCATT	64	24	67	
rrn16S.f	TAAGCATCGGCTAACTCTGTGCC	467	23	68	365
rrn16S.r	TACAGCACTGCACGGGTCGATAC	787	23	69	
rrn23S.f	TGTGGTTAGGGGTGAAATGCCAC	800	23	68	398
rrn23S.r	ATCGCTTAGCCCCGTTCATCTTC	1153	23	68	
rrn4.5S.f	ACGAGCCGTTTATCATTACGATAGGTGTC	44	29	68	77
rrn4.S5.r	CCGGTCTGTTAGGATGCCTCAG	71	22	68	
rrn5S.f	GGCGTAGAGGAACAACACCAATCC	39	24	68	95
rrn5S.r	AGCTATTTTTCCGCAGGACCTCC	87	23	68	
trnA-ugc.f	GGGGATATAGCTCAGTTGGTAG	22	22	65	72
trnA-ugc.r	GGAGATAAGCGGACTCGAACC	52	21	66	873
trnE-uuc.f	CCCCCATCGTCTAGTGGTTCAG	51	22	68	65
trnE-uuc.r	GGGAAGTCGAATCCCCGCTG	29	20	68	
trnF.f	GGATAGCTCAGTTGGTAGAG	24	20	61	60
trnF.r	CCAGATTTGAACTGGTGACAC	44	21	60	
trnfM-cau.f	CGGGGTAGAGCAGTTTGGTAG	52	21	67	61
trnfM-cau.r	GGATTTGAACCCGTGACCTCAAGG	35	24	68	
trnI-cau.f	CATGGCTGAATGGTTAAAGCG	49	21	61	64
trnI-cau.r	CCAGTAGGAATTGAACCTACGA	27	22	62	798
trnI-gau.f	GCTATTAGCTCAGTGGTAGAGCG	25	23	67	69
trnI-gau.r	GGGCCATCCTGGATTTGAACC	51	21	67	
trnK-uuu.f	GGTTGCTAACTCAACGGTAGA	49	21	62	42
trnK-uuu.r	CTAGTCGGTTAAAAGCCGAGTA	51	20	62	
trnL-uag.f	GCTATGGTGAAATTGGTAGACACGCTG	30	27	68	76
trnL-uag.r	GCCGCTACTCGGACTCGAAC	60	20	68	
trnP.f	GGATGTAGCGCAGCTTGGTAG	52	21	67	70
trnP.r	GGGATGACAGGATTTGAACCCGTG	26	24	68	
trnR-acg.f	GGGCTTGTAGCTCAGAGGATTAG	23	23	67	52
trnR-acg.r	CCCGACACCGTGGTTCGTAG	33	20	68	
trnV-gac.f	GGGAAGGGATATAACTCAGCGG	25	22	67	75
trnV-gac.r	AGGGATAATCAGGCTCGAACTG	57	22	65	
ycf1.f1	GCCTCTGCATTTAGCATTGGGTAG	269	24	67	544
ycf1.r1	TGTTTAGTCCCACCCGTTTCTGAG	765	24	67	
ycf10.fcemA	CCCTGGTTGATCTCTCTCTGCTG	80	23	68	450
ycf10.rcemA	AACCGTGAGGCGAATGGAATCCG	486	23	68	
ycf2.f1	CTAACTGGAGTTCGCGGTGGTG	542	22	68	348
ycf2.f2	TCCCAGGTAAGATCGGTTCAGG	5707	22	67	350
ycf3.f	GCTATGCGACTAGAAATTGACCCC	304	24	66	202
ycf3.r	CTCCTTGTTGAATGGCCTGTTCTC	149	24	66	989
ycf4.f	GTAATTTCTGCTGGGCCTTTATCC	82	24	65	374
ycf4.r	CATATAAAGGACACGCCGGGC	412	21	67	
ycf5.f	CAGTCCCAGTGGTTAATGATGCACG	427	25	68	347
ycf5.r	GCTCCTGAAAGAATGCCGACGG	728	22	68	

Table 4. Continued

Gene Name	Primer-sequence	Positioin of 3'end	Length	Tm °C	Length of PCR product
ycf6.f/petN	CTCGCATGGGCTGCTTTAATGG	37	22	67	69
ycf6.r/petN	TCCACTTCTTCCCCACACTACG	63	22	67	
ycf9.f/psbZ	CAATCTTACTGATTAGTGTACCCGT	74	25	62	59
ycf9.r/psbZ	GTTACTCGACCAACCATCAGGA	87	22	64	

2.2 Methods

2.2.1 RNA Isolation and Gel Blot Analysis

Total RNA was extracted from frozen plant tissue using TRIzol® Reagent following the manufacturer's instruction (Invitrogen, Karsruhe, Germany). DNase I (RNase free) (Roche, Mannheim, Germany) was added to the solutions to remove DNA contamination. Concentration and purity were measured by photometery and equal amounts of RNA from each sample were loaded on 1.2% MOPS agarose gels to check their quality and quantity for RNA gel blot and macroarray analysis. RNA was denaturated with 30% glyoxal at 42°C before electrophoresis. Separated RNA was capillary transferred onto a Biodyne A nylon membrane (0.2 µm; Pall, Dreieich, Germany) and immobilized by UV radiation (UV-Stratalinker 1800, Startagene, La Jolla, USA). Probes for hybridization were amplified by PCR using gene specific primers. Each probe was labelled with [α-^{32}P] dCTP using Random Primed DNA Labeling Kit (Roche, Mannheim, Germany) and hybridized at 65°C at least for 12 hours in hybridization buffer (0.25M Na_2HPO_4, pH 7.2, 7% SDS). The hybridized membranes were washed three times in 0.5 x SSC, 0.1% SDS at 65°C for 30 min. Hybridization signals were obtained with the FLA-3000 phosphoimager (Fuji, Tokyo, Japan).

2.2.2 Preparation of Macroarray Filters

94 probes of genes for plastid proteins, tRNAs and rRNAs were amplified by using gene-specific primer pairs (data available upon request) and DNA of WT, accession Columbia, as template. Intron-containing genes were amplified with Titan One Tube RT-PCR Kit (Roche, Mannheim, Germany). The size and quality of PCR products purified with PCR Purification Kit (Qiagen, Hilden, Germany) were checked on 1.2% agarose gels by electrophoresis and diluted to 3 different concentrations (30.0, 7.5 and 1.87 ng/µl). Probes were spotted on 11.9 cm x 7.8 cm positively charged nylon membranes (Hybond™-N+ Amersham Pharmacia

Biotech, Munich, Germany) by using a 96-pin tool (0.4 mm pins with a BioGrid spotting Device-Roboter (BioRobotics, UK). Each probe was spotted 20 times in duplicate to get final DNA quantities of 1.25, 5 or 20 ng per spots (Figures 6A and 6B). The spotted DNA was denatured in 1.5 M NaCl, 0.5 M NaOH and neutralized in 0.5M Tris-HCl pH 7.2, 1M NaCl. Filters were cross-linked with UV light (120 mJ, 302 nm, UV-Stratalinker 1800) and dried before hybridization.

2.2.3 Hybridization of Labelled cDNAs to Macroarray Filters

Before hybridization, macroarray filters were incubated at 65°C for 1 hour in 10 ml buffer (0.25 M Na_2HPO_4 pH 7.2, 7% SDS). [α-^{32}P]-dCTP labeled cDNAs were synthesized at 50°C for one hour with hexanucleotides (Roche, Mannheim, Germany) using the SuperScript™ III RNase H$^-$ Reverse Transcriptase (Invitrogen, Karlsruhe, Germany) and 20 µg of total RNA as template. After inactivation of transcriptase at 70°C for 20 min, the labeled cDNAs were incubated at 37°C for 20 min with RNase H (Invitrogen, Karlsruhe, Germany) to remove RNA. cDNAs were purified by MicroSpin™ G-25 columns (Amersham, Freiburg, Germany) and used for hybridization for 12 hours at 65°C in hybridization buffer. Filters were washed separately at 65°C for 20 min in three different washing buffers (2 x SSC, 0.1% SDS; 1.0 x SSC, 0.1% SDS; 0.5 x SSC, 0.1% SDS).

2.2.4 Normalization and Statistical Analysis

To increase the accuracy for the array data, probes were spotted in three different concentrations and in duplicate. The radioactive images were scanned with a FLA-3000 phosphoimager (Fuji, Tokyo, Japan) and the obtained signals were imported to the AIDA Image Analyzer (3.52) software for background correction and normalization. The mean value of three selected background dots within each sub-grid was calculated for background subtraction. Background-corrected hybridization signals were normalized using R/MAANOVA version 0.98.8 (Wu et al., 2002) implemented in R program (www.r-project.org). Robust locally weighted regression (lowess) method was applied in R/MAANOVA for normalization (Cleveland et al., 1988). After performing standard t-tests, adjusted p-values for each gene were calculated using web-based microarray analysis toolbox (http://nbc11.biologie.uni-kl.de).

2.2.5 Microarray Data Analysis

ATH1 (22k) expression data from *Arabidopsis thaliana* were obtained from Genevestigator, GEO, AtGenExpress databases using default parameters and selecting data from mutants and a variety of biological conditions (Zimmermann et al., 2004; Barrett and Edgar, 2006). All microarray data from Genevestigator were normalized by the MAS 5.0 method using Simpleaffy implemented in R program (Wilson et al., 2005). Expression profiles of 79 plastid and 1,430 nuclear genes were selected for cluster analysis. For cluster analysis fold changes were first converted to $\log_2$ and then normalized with mean values. Several clustering methods (Hierarchical, SOM, K-means, and Terrain Clustering) were performed using program Genesis (Sturn et al., 2002).

2.2.6 Promoter Analysis

Promoter analysis was performed using the web-based promoter database Athena (http://www.bioinformatics2.wsu.edu/cgi-bin/Athena/cgi/home.pl) (O'Connor et al., 2005).

3 RESULTS

3.1 Plant Growth and Mutant Phenotypes

If not otherwise indicated leaves from three-week-old WT and mutant plants grown in the climate chamber (Percival, Iowa, USA) under continuous light (60 µmol photons m^{-2} s^{-1}) were used for the analysis. Detailed information of mutant names, sources, phenotypic characteristics and photosynthetic parameters as well as variations in growth and stress conditions can be obtained from Tables 2 and 3. Most of the affected genes in the mutants are essential for photoautotrophic growth and plants only survived when grown on medium supplemented with sucrose indicating primary photosynthetic defects. Mutant plants exhibited three major phenotypes i.e., albino mutants arrested at an early stage of chloroplast development, *hcf* (*high chlorophyll fluorescence*) mutants with impaired photosynthetic electron transport (Meurer et al., 1996b), and yellow to pale green lines with deficiencies in various often unknown chloroplast functions (Table 2).

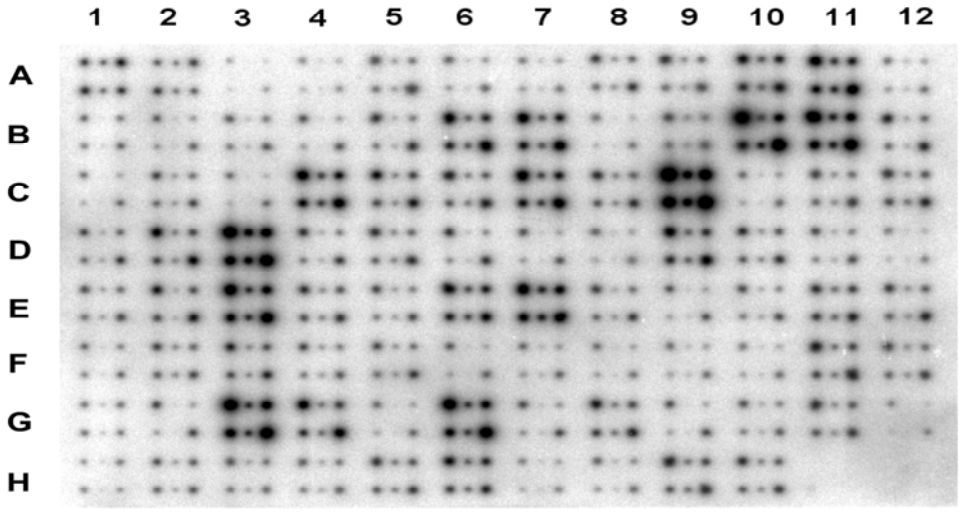
3.2 Establishment of Plastid Macroarrays

DNA macroarrays containing probes of genes encoding all plastid proteins, ribosomal RNAs, and 11 tRNAs were generated for comprehensive expression analyses of various *Arabidopsis* mutants affected in chloroplast development and function, and under different environmental conditions (Figures 5A and 5B). Names and gene functions are described in Table 1. To reliably apply and statistically estimate the use of macroarrrays for the evaluation of expression levels of plastid genes three different concentrations of each gene probe were spotted in duplicate onto filters, background subtraction for each gene grid was performed, p-values have been calculated, and several repetitions of selected experiments were used (Table 1). Moreover, scatter plot analyses have been performed routinely (Figure 5C). Basically, RNA gel blot analysis of several mutants with representative gene probes confirmed the data obtained by array-based expression profiling (Figure 6). Due to the relative high expression level of most chloroplast genes the intensity of statistically calculated hybridization signals was much higher than the background level indicating reliable use of the macroarray-based approach (Figures 5A and 5B). Furthermore, macroarrays obtained from *Arabidopsis* chloroplast genes can also be applied heterologously in a wide range of plant species due to the high conservation of plastid genes (data not shown).

A

	1	2	3	4	5	6	7	8	9	10	11	12
A	*psbA*	*matK*	*trnK-uuu*	*rps16*	*psbK*	*psbI*	*rps2*	*atpI*	*atpH*	*atpF*	*atpA*	*rpoB*
B	*rpoC1*	*rpoC2*	*petN*	*trnE-uuc*	*psbM*	*psbD*	*psbC*	*psbZ*	*ycf3*	*psaA*	*psaB*	*rps14*
C	*trnfM-cau*	*rps4*	*trnF-gaa*	*atpB*	*atpE*	*ndhC*	*ndhK*	*ndhJ*	*rbcL*	*accD*	*psaI*	*ycf4*
D	*cemA*	*petA*	*psbE*	*psbF*	*psbL*	*psbJ*	*petL*	*petG*	*psaJ*	*rpl33*	*rps18*	*trnP-ugg*
E	*rpl20*	*clpP*	*psbB*	*psbT*	*psbH*	*petB*	*petD*	*psbN*	*trnI-cau*	*rpl23*	*rpl12*	*rps19*
F	*rpl22*	*rps3*	*rpl16*	*rpl14*	*rps8*	*rpl36*	*rps11*	*rpoA*	*ycf2.1*	*ycf15*	*rps12*	*rps7*
G	*ndhB*	*trnV-gac*	*rrn16S*	*trnI-gau*	*trnA-ugc*	*rrn23S*	*rrn4.5S*	*rrn5S*	*trnR-acg*	*ndhF*	*rpl32*	*trnL-uag*
H	*ycf5*	*ndhI*	*ndhG*	*ndhE*	*psaC*	*ndhD*	*rps15*	*ndhH*	*ndhA*	*ycf1*	*Water*	*Water*

B



C

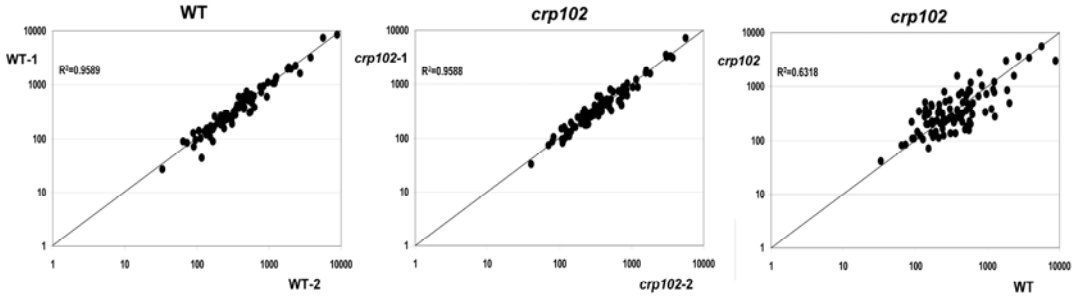


Figure 5. Spotting Scheme for the 94 Plastid Gene Probes Applied to Macroarrays, Representative Hybridization Signals for the WT, and Scatter Plot Analysis.
(A) The spotting scheme for the 94 plastid gene probes used for plastid macroarrays is shown.
(B) Representative hybridization signals obtained with cDNAs from WT. Each plastid gene probe was arrayed in three different concentrations (1:1, 1:4, 1:16) and in duplicate.
(C) Linear regression ($R^2 > 0.95$) of representative scatter plot analysis of repeated experiments using WT and *crp102* RNA demonstrates the reliability of the array-based approach. The scatter plots for WT and *crp102* ($R^2 = 0.63$) show that the expression profile is significantly changed in the mutant.

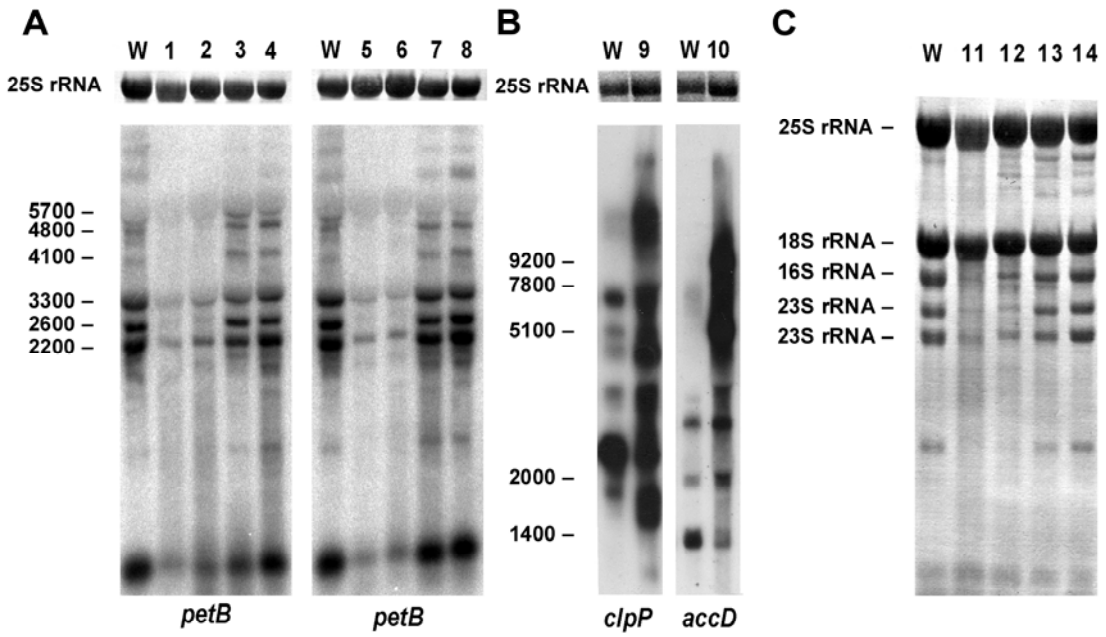


Figure 6. RNA Gel Blot Analysis of the Plastid *petB*, *clpP* and *accD* Genes in Mutants and Wild Type.
Each lane was loaded with 10 µg of total leaf RNA isolated from 3-week-old mutant (lanes 1 to 14) and wild-type (W) seedlings that had been grown on sucrose-supplemented agar medium. The numbers on the left indicate RNA sizes in bases.
(A) RNA gel blot analysis of *petB* (encoding cytochrome b_6) in the mutants *ins24* (lane 1), *ins27* (2), *ins17* (3), *ins23* (4), *ins28* (5), *ins25* (6), *ins4* (7), and *ins11* (8).
(B) RNA gel blot analysis of *clpP* (caseinolytic protease) (9) and *accD* (carboxytransferase beta subunit of the acetyl-CoA carboxylase) (10) in the mutant *crp135*.
(C) rRNAs of *ins24* (11) and *ins27* (12) mutants of group A, and *ins17* (13) and *ins23* (14) mutants of group B, stained with methylene blue.

3.3 Expression Profiling of Plastid Genes under Various Biological Conditions

In concert with photosynthetic mutants and different types of tissues (stems, leaves and flowers), various hormones (gibberellin, abscisic acid, auxin, and cytokinin), herbicides (DCMU (3-(3', 4' - dichlorphenyl) - 1,1 - dimethylurea) and methylviologen (*N*,*N*'-dimethyl-4,4'-bipyridinium, MV), biological stress (heat, light, cold, and dark), and sucrose applied for a defined time were used to study the effect on plastid gene expression patterns (Tables 2 and 3). To compare the expression profile obtained under the various biological conditions with that of the mutants a prolonged treatment was chosen to reach steady-state mRNA levels. This also avoided detection of transcript levels in the course of induction or repression of gene expression. In the reproductive flower organ, photosynthetic genes tend to be down-regulated

and non-photosynthetic genes were preferentially up-regulated when compared to leaves (Figure 7A). Most of the chosen stress conditions induced a significant deviation from the expression pattern observed under standard conditions (Figures 7B, 7C, 7D, 7E and 7F). Similarly, excess light inducing photo-inhibition and formation of reactive oxygen species caused reduction of photosynthetic gene expression and induction of genetic system genes (Figure 7B). DCMU, known to block the electron flow from Q_A- to Q_B and thus preventing reduction of plastoquinone (Yamagishi et al., 1987) did not show a significant differential expression of most plastid genes indicating that transcript steady state levels are not changed upon adaptation to the herbicide (Figure 7C). Auxin plays an essential role in coordination of numerous developmental processes in plant life cycle (Grebe, 2005). Although auxin treatment had a strong bleaching effect, the expression of most plastid genes was not significantly changed compared to other stress conditions. Exceptionally, the ribosomal RNAs *rrn4.5S* and *rrn5S* were much more abundant upon auxin treatment (Figure 7D). In contrast, several other stress conditions like heat (Figure 7E) and cold (Figure 7F) as well as sucrose depletion significantly affected the expression of numerous plastid genes (Table 1). Differential expression of many plastid genes was more pronounced under various stress conditions (e.g. light, heat, cold) than that in several non-photosynthetic mutants (e.g. *hcf145* and *crp160*) (Figures 7 and 8).

3.4 Cluster Analyses of Plastid Genes deduced from 89 Transcriptomes

From all datasets, p-values were statistically calculated demonstrating the significance of the differential expression of plastid genes in 78 mutants, 3 tissues and under 12 stress conditions. Filtering genes by p-values below 0.05 and 1.5 fold-changes did not significantly affect clustering of co-expressed plastid genes. Therefore, all datasets were included in this clustering analysis for a more comprehensive and subtle expression analysis of the plastid genome. The fold-changes were converted to $\log_2$ for simplicity, expressed relative to the mean value for normalization and then hierarchical clustering was performed using default parameters of the Genesis program (Eisen et al., 1998; Sturn et al., 2002).

Hierarchical clustering of 89 transcriptomes identified two groups, I and II, which displayed reversal expression patterns (Figure 9A). Mutants of group I preferentially exhibited albino (~70%) and less often pale or yellow phenotypes. Most up-regulated genes of group I had non-photosynthetic function, whereas most down-regulated genes encode photosynthetic proteins. Many of these mutants are severely affected either primarily or

secondarily in the general mRNA metabolism. In average the expression of about 13 genes is more than four times fold changed in the mutants.

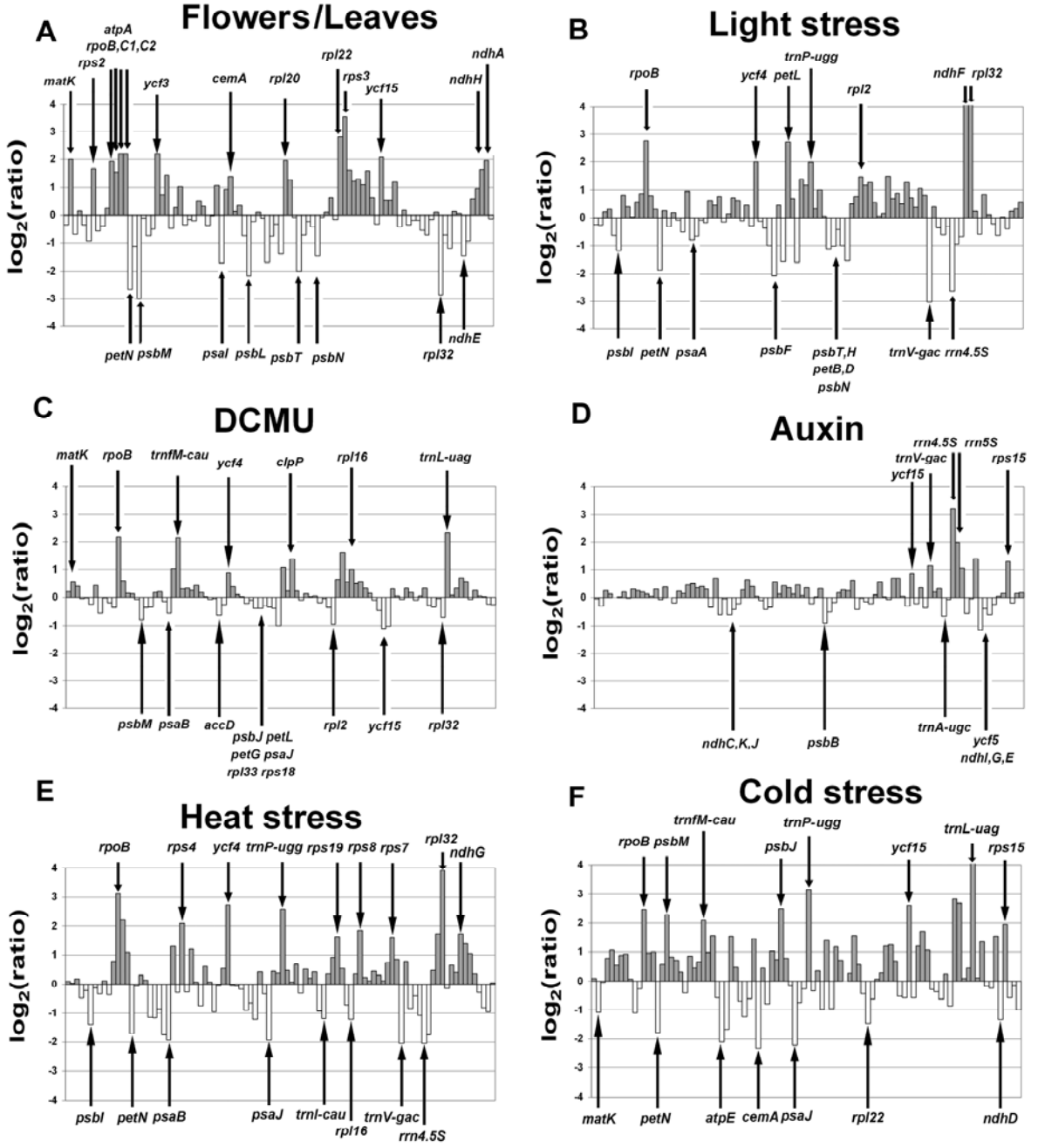


Figure 7. Changes in Plastid Transcript Levels in Tissues and Leaves Exposed to the Indicated Stressors.

Log_2-transformed fold changes in plastid RNA levels were determined in flowers as compared to leaves **(A)**, and in leaves subjected to high light stress **(B)**, DCMU treatment **(C)**, auxin treatment **(D)**, heat **(E)**, and cold stress **(F)** as compared to untreated control leaves. Significant expression ratios are indicated by the arrows. Detailed information on the stress conditions employed is given in Table 3. In all six histograms, genes are listed according to their positions on the plastid chromosome.

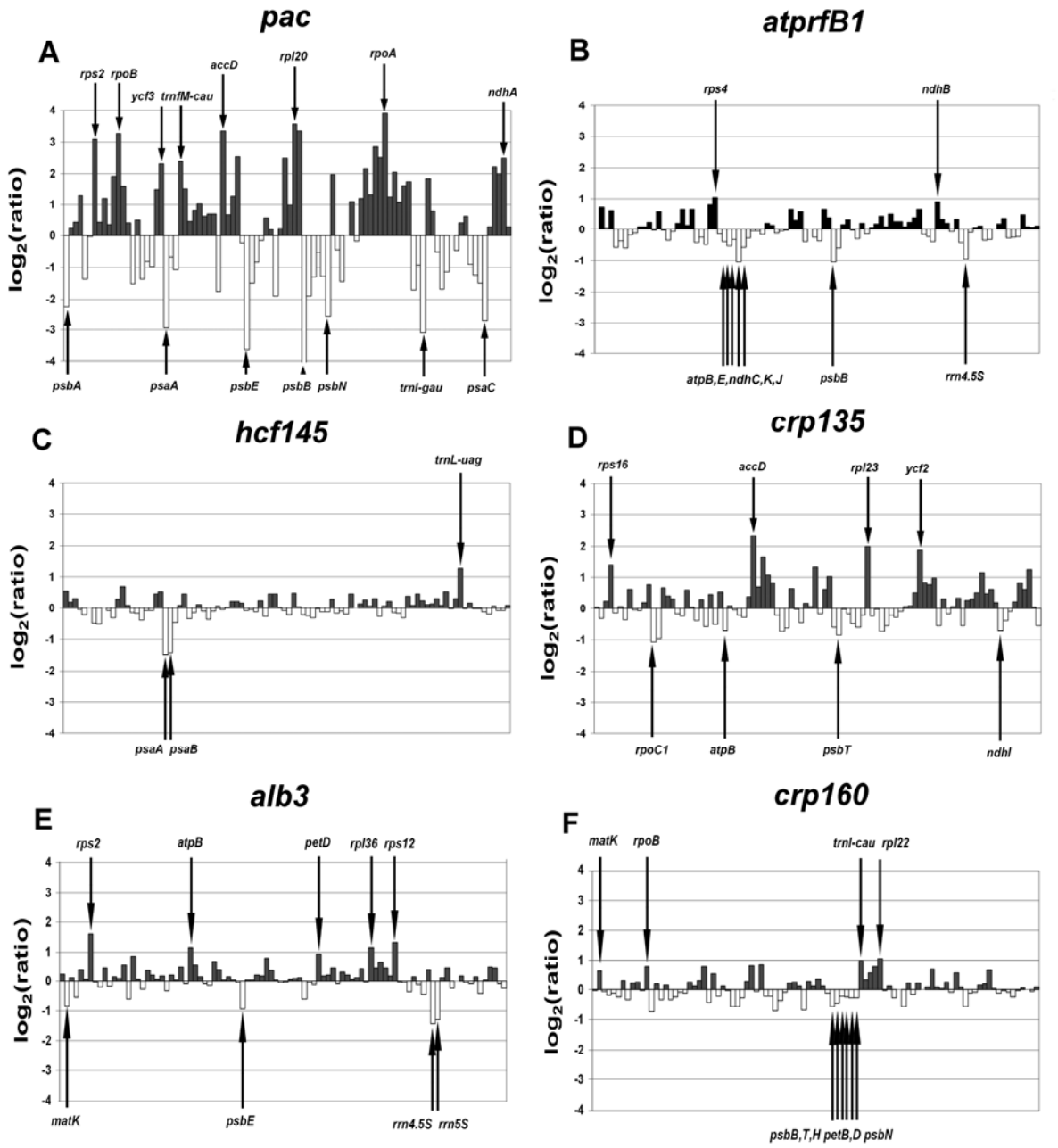


Figure 8. Plastid Transcript Levels in Six Representative Nuclear Mutants Compared to WT.

Log_2-transformed fold changes in RNA levels in each mutant are expressed relative to WT. Up-regulated genes have positive, down-regulated genes negative values. Significant expression ratios are indicated by the arrows. Genes are listed as in Figure 7. **(A)** *pac*, **(B)** *atprfB1*, **(C)** *hcf145*, **(D)** *crp135*, **(E)** *alb3*, **(F)** *crp160.*

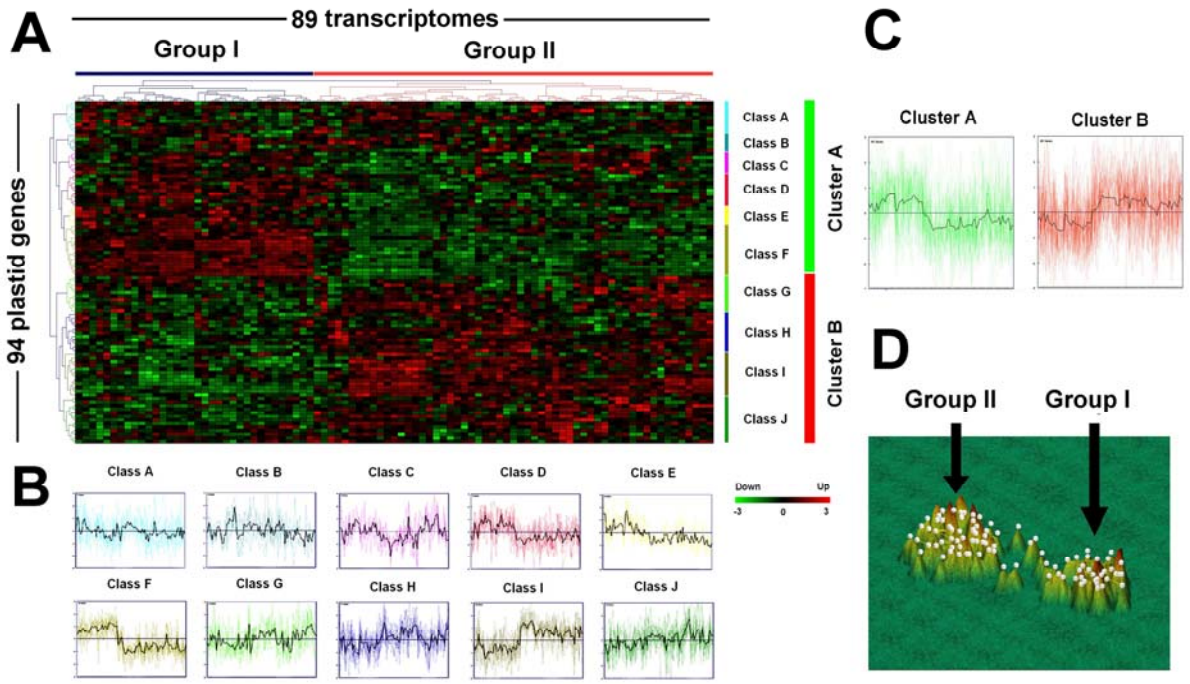


Figure 9. Expression Profiles of 94 Plastid Genes in 89 Transcriptomes.

(A) Transcript levels in 75 mutants and in WT plants exposed to 14 different biological conditions were determined by macroarray analysis. Fold change values were transformed to $\log_2$ and normalized releative to the mean value of genes and experiments. Hierarchical clustering was performed using Euclidean distances and complete linkage. Fold changes close to, higher and lower than the mean values are represented by black, red and green colors, respectively. Co-expressed plastid genes were distributed into two major clusters (A and B), which were further divided into 10 classes (A - J). Cluster A (green bar) and cluster B (red bar) contain six and four classes, respectively. Detailed information can be found in Table 6.

(B) Average expression views of plastid genes in each class show 10 distinct expression patterns of plastid genes in 89 transcriptomes. The colors used correspond to the classes in Figure 9A. The mean expression pattern within each gene class is shown by the black line. The x and y axes represent the 89 transcriptomes and $\log_2$-transformed fold changes of plastid genes, respectively. The sequence along the x-axis is according to Figure 9A.

(C) Non-hierarchical K-means clustering (K=2) was performed as described in Methods and identified almost the same clusters A (green) and B (red) as shown in panel A. The order of the 89 transcriptomes is identical to that shown in panel A. The average expression views of the two clusters are shown.

(D) Here expression profiles were used to cluster the 89 transcriptomes rather than genes using non-hierarchical terrain clustering as described in Methods. The terrain map illustrates the correlation of the 89 transcriptomes in three dimensions. Peak height corresponds to the density of transcriptomes, denoted by red, yellow and green colors. The white cube on each peak indicates an individual or a group of transcriptomes and neighboring peaks have similar expression profiles. The arrows indicate the two distinct transcriptome groups.

Group II contains transcriptomes of WT plants exposed to a variety of stress conditions as well as of mutants affected in various chloroplast functions. Mutants of group II often showed a pale or *hcf* phenotype. The appearance of albino mutants in group II (~21%), like *alb3* and *vipp1*, demonstrates a deviation from the signature typical for most albino mutants clustering in group I. Different from mutants in group I, albino mutants in group II did not show a severe differential expression of plastid genes. Presumably, an albino phenotype and an early arrest of chloroplast development are not necessarily correlated with the expression signature of group I. Most mutants of group II exhibited similar expression patterns with less striking fold-changes and a lower number of significantly differentially expressed genes than mutants of group I (Figure 9A) indicating primary deficiencies in the general gene expression system in mutants of group I. The data also imply that related genetic defects and stress conditions both affecting chloroplast functions result in similar expression patterns. Reference mutants, *hcf145 and atprfB1* (previously designated *atprfB*) were clustered together with other pale and non-photosynthetic mutants as well as with transcriptomes of various stress conditions in group II. The degree of up- and down-regulation of genes was generally higher (> 2.5-fold) in the reference mutants than in mutants not supposed to be impaired in mRNA abundance in group II.

Table 5. Distribution of Co-expressed Plastid Genes within 10 Identified Classes Shown in Figure 9A.

Class A	9 genes	*ndhB, ndhC, ndhF, petG, psbJ, trnfM-cau, trnR-acg, ycf2, ycf5*
Class B	5 genes	*petL, psbC, psbD, rps1, rrn4.5S*
Class C	6 genes	*rpoB, rps7, rps8, trnL-uag, trnP-ugg, ycf4*
Class D	9 genes	*atpB, atpE, atpH, atpI, cemA, ndhJ, ndhK, rpl16, rpl20*
Class E	5 genes	*accD, psbZ, rps18, trnF-gaa, trnK-uuu,*
Class F	14 genes	*clpP, ndhA, ndhH, rpl14, rpl36, rpoA, rps11, rps15, rps2, rrn23S, trnA-ugc, trnI-cau, ycf15, ycf3*
Class G	10 genes	*ndhE, ndhG, ndhI, petA, rpl22, rpoC1, rpoC2, rps3, rps4, trnV-gac*
Class H	11 genes	*petN, psaJ, psbA, psbF, psbL, psbM, rbcL, rrn16S, rrn5S, trnE-uuc, trnI-gau*
Class I	12 genes	*matK, ndhD, petB, petD, psaC, psbB, psbE, psbH, psbK, psbN, psbT, rpl2*
Class J	13 genes	*atpA, atpF, psaA, psaB, psaI, psbI, rpl23, rpl32, rpl33, rps12, rps14, rps19, ycf1*

Table 6. Functional categories and distribution of plastid genes within the two clusters identified by K-means clustering (K=2) of 89 transcriptomes shown in Figure 9.

Functional categories	Cluster A	Cluster B
ATP synthase	*atpA, atpB, atpE, atpF, atpH, atpI*	
NADH dehydrogenase	*ndhA, ndhB, ndhC ndhH, ndhK, ndhJ*	*ndhD, ndhE, ndhF,ndhG, ndhI*
Cytochrome b_6f		*petA, petB, petD, petG, petL, petN*
Photosystem I		*psaA, psaB, psaC, psaI, psaJ*
Photosystem II		*psbA, psbB, psbC, psbD, psbE, psbF,psbH, psbI, psbJ,psbK, psbL, psbM, psbN, psbT, psbZ*
Ribosomal proteins	*rpl14, rpl16, rpl20, rpl36, rps11, rps12, rps15, rps16, ps18, rps2, rps4, rps7, rps8*	*rpl2, rpl22, rpl2, rpl32, rpl33, rps14, rps19, rps3*
RNA polymerase	*rpoA, rpoB, rpoC1, rpoC2*	
rRNAs	*rrn23S, rrn4.5S*	*rrn16S, rrn5S*
tRNAs	*trnA-ugc, trnF-gaa trnfM-cau, trnI-cau trnK-uuu, trnL-uag trnP-ugg*	*trnE-uuc, trnI-gau trnR-acg, trnV-gac*
Unknown	*ycf15, ycf3, ycf4*	
Others	*accD, cemA, clpP, ycf2 , ycf5*	*matK, rbcL, ycf1*

Hierarchical clustering, which considers both, transcriptomes and plastid genes, identified two gene clusters, A and B, of similar size showing in average a reversal expression of their genes (Figure 9A). Cluster A contains 31 genetic system genes, four ATP synthase and seven NDH genes, whereas cluster B represents mostly genes encoding 25 components of the linear and cyclic electron transport, two ATP synthase genes as well as *rbcL*, and 19 non-photosynthetic genes. Remarkably, almost all genes present in cluster A contain promoters recognized by the NEP (Shiina et al., 2005; Liere and Börner, 2006; Kuhn et al., 2007), whereas genes known to be transcribed preferentially by the PEP are found in cluster B.

Hierarchical clustering further sub-classified plastid genes into 10 distinct co-regulated and often functionally associated gene classes, classes A - F are present in cluster A and classes G - J in cluster B (Figure 9B; Table 5). The averaged expression ratios for each class showed a differential expression view (Figure 9B). For example, classes D, E and F, which contain 19 genes with non-photosynthetic function (ribososomal proteins and RNAs, RNA polymerase, tRNAs, *clpP*, *accD*), *psbZ* and four genes of each the ATP synthase and the NDH show a very similar average expression view (Figure 9B; Table 5).

A similar tendency albeit less pronounced could be observed in classes A - C. Gene members of classes C and H, as well as D and I are reversely regulated. The expression of genes in classes G and J was rather complex and did not strictly follow the general tendency of class B members. Clustered genes often derived from same polycistronic operons or they were tightly functionally associated even when they were not co-transcribed as is the case for NDH genes in class A, ATP synthase genes in class D, photosynthetic genes in classes H and I, as well as ribosomal genes in class J (Table 5).

As a matter of fact, the number and gene members of clusters chosen from hierarchical analysis highly depend on the program and parameters utilized. Therefore, additional clustering methods were performed to confirm the identification of co-regulated genes. Using SOM (Self Organizing Map) (Tamayo et al., 1999) and K-means clustering methods (Soukas et al., 2000), the appearance of two distinct reversely expressed gene clusters (A and B) containing almost identical gene members was confirmed (Figure 9C; Table 6). With the exception of few genes SOM and K-means clustering showed a still tighter functional and transcriptional association of clustered genes. Genes for the NDH, ribosomal proteins, ribosomal RNAs, and tRNAs are present in both clusters with a higher representation in cluster A. Genes with non-photosynthetic or unknown functions like, *ycf2, ycf3, ycf4, ycf5, ycf15, accD, cemA*, and *clpP* are present in cluster A. Interestingly, cluster A consists of all genes encoding the ATP synthase and the RNA polymerase (Table 6). Cluster B contains all genes of photosynthetic complexes, PSI, PSII and cytochrome b_6f. Expectedly, many co-expressed plastid genes within clusters are present in same polycistronic operons (Table 7) indicating that co-transcription plays an important role in determining transcript abundance.

Terrain cluster analysis also confirmed two distinguishable groups of transcriptomes (Figure 9D). Each spot represents a transcriptome and their heights display the dynamics of gene expression. The two-dimensional orientations and spot distances, which reflect the relationship between transcriptomes, indicate that members of group I are quite unrelated to those of group II (Figure 9D).

Table 7. Co-transcribed Plastid Genes, which were Found to Be Tightly Co-Expressed as Determined by Hierarchical Clustering Analysis of 89 Transcriptomes.

Clusters of homogeneous function	Clusters of heterogeneous function
psbD, psbC	*rpl14, rpl36, rps11, rpoA*
atpI, atpH	*trnA-ugc, rrn23S*
atpB, atpE	*rps15, ndhH, ndhA*
ndhK, ndhJ	*trnE-uuc, psbM*
rpoC1, rpoC2	*rrn16S, trnI-gau, rrn5S*
rpl22, rps3	*psbB, psbT, psbH, petB, petD*
ndhI, ndhG, ndhE	*ndhD, psaC*
psbF, psbL	*psaA, psaB, rps14*
atpF, atpA	

3.5 Identification of Mutants Affected in the Chloroplast mRNA Metabolism

Defects in photosynthesis may result in a distinctive transcript pattern. To discern mutants primarily involved in the plastid mRNA metabolism from those affected in other plastid functions, the control mutants *pac* (Meurer et al., 1998), *atprfB1* (Meurer et al., 1996a, 2002), and *hcf145* (Lezhneva and Meurer, 2004), already known to be specifically affected in plastid mRNA degradation and processing were used. The data obtained from macroarray analysis of these mutants fitted quite well with those previously described by RNA gel blot analysis (Figure 5). For example, the expression of most photosynthetic genes was severely reduced (> 2.5-fold) but many genetic system genes encoding ribosomal proteins, subunits of the RNA polymerase, *ycf3,* and *accD* were significantly up-regulated (> 2.5-fold) in *pac* (Figure 8A). Expectedly, in *atprfB1*, which is affected in the stabilization of UGA stop codon-containing transcripts, such as *psbB, psbT, atpB, atpE, ndhK,* and *ndhJ*, significant down-regulation of

these genes could also be detected by the macroarray-based approach (Figure 8B and data not shown). The nuclear-encoded factor HCF145 is involved in the stabilization of the tricistronic *psaA-psaB-rps14* transcript in *Arabidopsis* (Lezhneva et al., 2004). Accordingly, the histogram of the *hcf145* mutant shows that although the expression of most genes was not severely changed, transcript abundance of *psaA* and *psaB* was highly reduced (Figure 8C) confirming for the first time the specificity by which the mRNA metabolism is affected in this mutant.

Table 8. Identified 14 Putative mRNA Metabolism Mutants.

Name	Group	Identified differentially expressed genes
ins20	A	*psbD, psbC, ndhC, ndhK, rpoC1, rpoC2, ndhG, ndhI, ndhD*
ins24	A	*psbD, psbC, ndhI, ndhG, ndhE*
ins25	A	*psbD, psbC*
ins26	A	*psbD, psbC*
pac	A	*psbD, psbC*
ins15	A	*rpl16, rps8*
ins36	A	*ndhK, ndhC*
ins21	A	*psaJ, psaC, psbA rbcL, rpl15, rps8*
ins34	B	*rrn4.5S, rrn5S*
hcf145	B	*psaA, psaB*
crp102	B	*ndhI, ndhG, ndhE, atpH, atpI, atpF, atpA*
crp135	B	*psbB, psbT, rpoC1, rpoC2, accD, cemA, petA, clpP*
hcf109	B	*psbB, psbT*
ins33	B	*ndhA, rpoA*

By analyzing transcript profiles of 78 nuclear *Arabidopsis* mutants 25 lines showed at least four times fold changes for at least 10 genes as compared to the WT. In 43 mutants, like in the *pac* control or the newly identified mutants *ins20, ins24,* and *ins25*, the number of genes, which are significantly differentially expressed (2.5 fold changes), was higher than 10. In six mutants only one or few transcript levels were either increased and decreased or only decreased, like the *psaA-psaB-rps14* transcript in the control mutant *hcf145* **(**Figure 8D). In the remaining mutants, like *vipp1* and *alb3*, the degree of up- and down-regulation was less

striking (> 2.5-fold) suggesting that the expression profile in these mutants reflects secondary or milder primary effects (Figure 8E). Both proteins, VIPP1 (Aseeva et al., 2007) and ALB3 (Ossenbühl et al., 2004), have been shown to be responsible primarily for assembly of membranes and thylakoid complexes, respectively, but not for the plastid mRNA metabolism. The defect of *petB* mRNA processing in *crp160,* which was found by RNA gel blot analysis (Figure 5) and real time RT PCR (data not shown) could not be evaluated by expression profiles (Figure 8F) indicating limitations of the array system for the detection of only splicing or processing variants. This is clearly due to the polycistronic nature of transcripts and various posttranscriptional processing events in chloroplasts.

Since almost all chloroplast transcriptomes responded rather sensitive to mutations and displayed complex patterns the characteristic expression signatures of groups, I and II, as well as fold changes could not be used to identify primary defects in chloroplast mRNA metabolism. Therefore, mutants showing a departure from the general expression signature within the groups I and II, e.g. mutants in group I and II showing a down-regulation of co-transcribed genes in cluster A and B, respectively, or vice versa were selected. This allowed identification of 14 putative mRNA metabolism mutants, among them the three control mutants, *pac, hcf145, hcf109*, and 11 new mutants (Table 8).

Representative RNA gel blot analysis confirmed the specificity by which the selected mutants are affected in plastid gene expression (Figure 5). For instance, although *crp135* (*chloroplast RNA processing*) clustered in group II non-photosynthetic genes, like *clpP* and *accD*, were several fold up-regulated. Furthermore, this mutant also showed severe RNA processing defects (Figure 5) indicating a functional linkage between mRNA processing and abundance. This confirmed the specificity and strength by which the mRNA processing and abundance is affected in this line, and that the array-based approach is well suited to identify mutants directly involved in the plastid mRNA metabolism.

3.6 Evaluation of Nuclear and Plastid Gene Expression using Affymetrix Microarray Data

ATH1 microarrays of *Arabidopsis thaliana* contain 22,500 probes representing about 24,000 nuclear and 79 plastid genes. Using different normalization methods often results in different evaluations of gene expression. For detection of differentially expressed genes, RMA or GCRMA normalization methods are recommended (Shedden et al., 2005). For identification of co-expressed genes located within operons, the MAS 5.0 method is most suitable (Harr et

al., 2006). The usability of the methods for normalization of plastid gene expression was checked. When using GCRMA and RMA it appeared that compared to the expression of most nuclear genes plastid gene expression is only moderately changed in the transcriptomes analyzed (data not shown). Differential expression and co-expression of plastid genes could at best be performed with the MAS 5.0 method due to its higher sensitivity. Therefore, MAS 5.0 normalized data from Genevestigator were selected to study and compare plastid and nuclear gene expression.

3.7 Use of Microarrays for the Evaluation of Plastid Gene Expression in Various *Arabidopsis* Mutants

Expression data for plastid genes from 136 different transcriptomes defined by mutants mostly affected in non-chloroplast functions, like morphogenesis of plant organs and signaling pathways were selected. Fold changes were obtained, converted to $\log_2$, normalized and clustered as described in Methods. Using these data, plastid transcriptomes and genes have not been classified in two groups and two clusters, respectively (Figure 10). Nevertheless, in accordance to the operon organization of plastid genes 35 genes co-transcribed in 14 polycistronic units were found to be co-expressed (Table 9). Hierarchical clustering identified nine groups of plastid transcriptomes (I - IX) and six co-regulated plastid gene clusters (A - F) (Table 10). The average expression view of cluster A and B displayed more prominent alternations as compared to other clusters (Figure 10B). This demonstrates that genetic defects unrelated to the chloroplast did not identify discernable gene expression signatures and clusters of genes which are under transcriptional control.

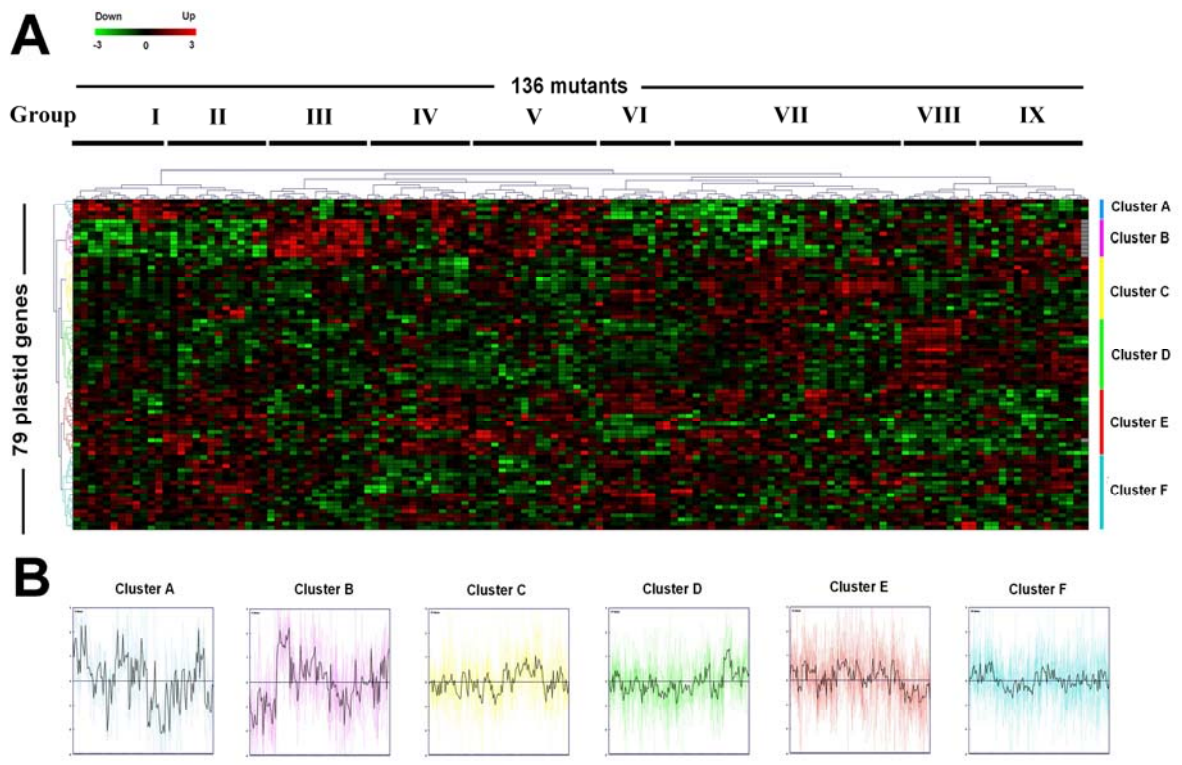


Figure 10. Expression Map of 79 Plastid Genes in 136 Different Mutants Generated from Data Obtained from Genevestigator.
(A) Hierarchical clustering identifies six co-regulated gene clusters, as illustrated by different colored bars. Up-regulated, down-regulated, and unchanged transcript levels are indicated by red, green, and black colors, respectively.
(B) Average expression views of plastid genes in each cluster. The mean expression pattern within each gene clusters is shown by the black line. The x- and y-axes represent 136 mutants and $\log_2$-transformed fold changes of plastid genes, respectively.

Table 9. Co-transcribed Plastid Genes, which were Found to Be Tightly Co-expressed as Determined by Hierarchical Clustering Analysis of Transcriptomes of 136 Mutants not Related to Chloroplast Functions.

Clusters of homogeneous function		Clusters of heterogeneous function
atpB, atpE	*psbB, psbT*	*ndhI, ndhG, ndhE, psaC*
petB, petD	*rpoC1, rpoC2*	*rps15, ndhH, ndhA*
atpF, atpA	*rpl23, rpl2*	*psaJ, rpl33, rps18*
psbJ, psbL	*psbI, psbK*	
psbE, psbF	*ndhJ, ndhK*	
rpl16, rpl14, rps8, rpl36, rps11		

Table 10. Distribution of Co-expressed Plastid Genes within Six Identified Clusters Shown in Figure 10.
Data were deduced from hierarchical cluster analysis of plastid genes in 136 mutants not primarily affected in chloroplast functions.

Cluster A	5 genes	*orf77, ycf2, ndhF, ndhB, rps14*
Cluster B	8 genes	*psaC, ndhE, ndhG, ndhI, ndhA, ndhH, rps15, rps12*
Cluster C	15 genes	*atpE, atpB, petA, clpP, rps7, rps4, petB, petD, psaB, ndhD, psaI, ycf5, psbC, atpA, atpF*
Cluster D	17 genes	*psbM, psbJ, psbL, ndhC, orf31, ycf9, psaJ, rpl32, matK, rpl33, rps18, rpoA, ycf1, rps3, rps19, rps2, rps16*
Cluster E	16 genes	*psbD, psaA, accD, psbH, ycf4, ycf10, rpoB, psbN, ycf6, atpH, psbF, psbE, rbcL, ndhJ, ndhK, psbA*
Cluster F	18 genes	*rps8, rpl14, rpl16, rpl22, rpl22, rpl23, rpoC2, rps11, rpl36, petG, psbB, psbT, rpl20, rpoC1, atpI, ycf3, psbI, psbK*

3.8 Use of Microarrays for the Evaluation of Plastid Gene Expression under Various Stress Conditions.

Fold changes of plastid gene expression were calculated from plants subjected to 83 stress conditions, including various abiotic, biotic, pathogen, chemical, nutrient, hormone, and light conditions (Zimmermann et al., 2004). However, it appeared that plastid transcriptomes and genes were clustered when various stress conditions were applied, which all affect the chloroplast (Figure 11). Again two groups of generally reversely regulated transcriptomes, I and II, could be defined. Group I includes biotic, light (different light qualities except UV-B stress), and chemical stress, and group II includes light (UV-B), abiotic, nutrient, and hormone stress. Genes were co-regulated in six clusters (A - F) (Table 11).

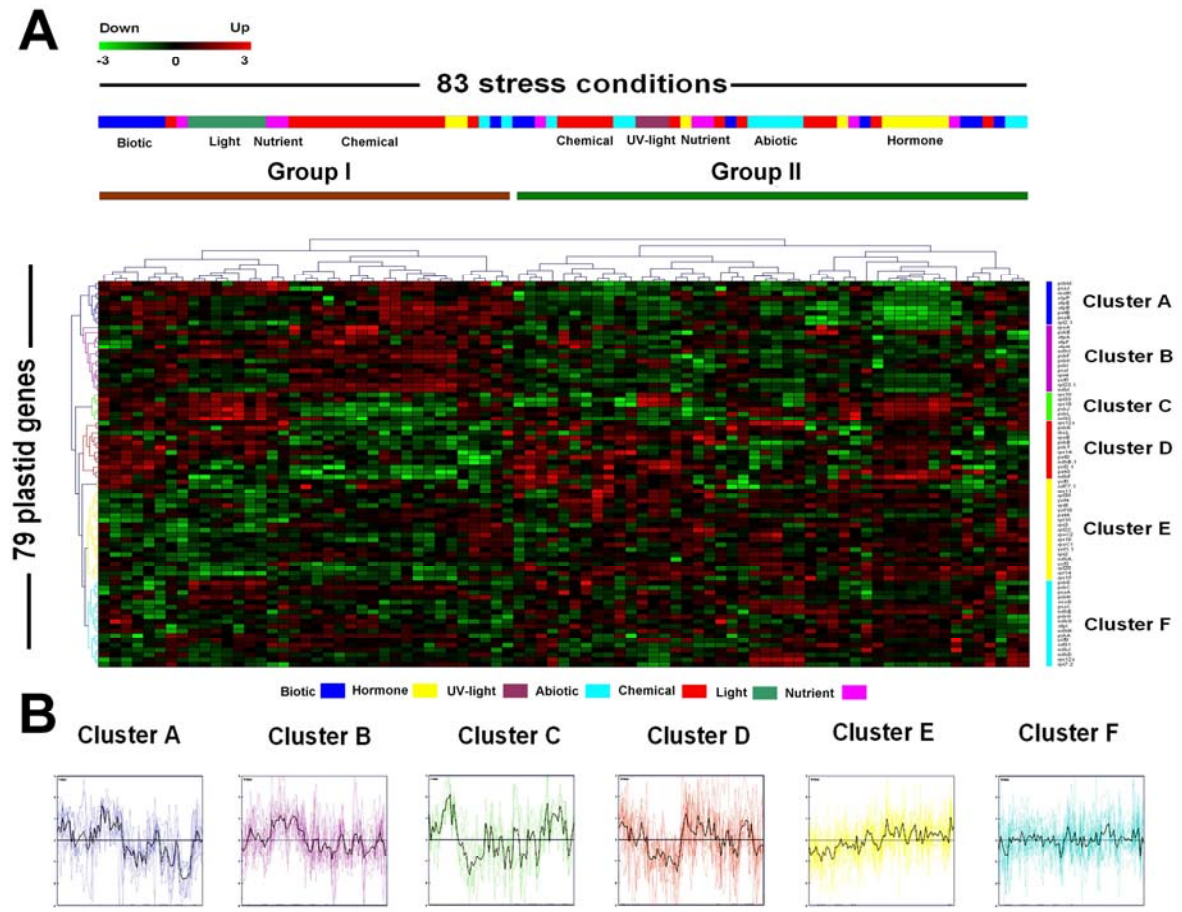


Figure 11. Expression Map of 79 Plastid Genes under 83 Various Stress Conditions Generated from Genevestigator.
(A) Hierarchical clustering identifies six co-regulated gene clusters as illustrated by different color bars. Up-regulated, down-regulated, and unchanged transcript levels are indicated by red, green, and black colors, respectively. The distribution into functional categories of genes in each cluster is shown by colored bars.
(B) The average expression view for plastid genes in each identified cluster is shown. The mean expression pattern within each gene clusters is shown by the black line. The x- and y-axes represent 83 different stress conditions and $\log_2$-transformed fold changes of plastid genes, respectively.

Although genes of clusters A - D encode quite heterogeneous functions, clusters E and F contain genes with almost exclusively non-photosynthetic (19 out of 21 genes) and photosynthetic (14 out of 18 genes) functions, respectively (Table 11). Therefore, clusters E and F resemble the two major transcriptionally determined clusters A and B, respectively, deduced from the mutant transcriptomes analyzed by the macroarray-based approach. Clusters E and F are reversely regulated predominantly under different light conditions. Gene expression of clusters A and B was highly induced in group I but reduced in group II (Figure 11B).

Table 11. List of 79 Plastid Genes in Six Clusters Identified by Hierarchical Clustering using 83 Various Stress Conditions.

Cluster A	9 genes	*psbM, psaJ, matK, clpP, atpE, atpB, petB, psaB, rpl2,*
Cluster B	14 genes	*rpoA, psbE, atpA, atpF, atpH, ndhC, psbF, psbK, psbI, psaI, rps4, ycf5, rpl23, ndhI*
Cluster C	6 genes	*rps16, rpl33, rps18, psbJ, psbL, rpl32*
Cluster D	12 genes	*rps12, psbN, rbcL, rpoB, psbB, psbT, rps14, petD, ndhB, ycf2, petG, ndhF*
Cluster E	21 genes	*ycf6, orf77, rps11, rpl36, ycf4, rps8, ycf10, petA, rpl16, rps3, rpl22, rpoC2, rps19, rpoC1, ycf10, rps2, ndhA, ycf3, rpl20, rpl14, rps15*
Cluster F	18 genes	*psbD, psbC, psaA, psbH, accD, psaC, ndhE, ndhK, ndhG, atpI, ndhH, psbA, ycf9, orf31, ndhJ, ndhD, rps12, rps7*

Table 12. Co-transcribed Plastid Genes, Which were Found to Be Tightly Co-Expressed as Determined by Hierarchical Clustering of 83 Transcriptomes Analysed under Various Stress Conditions.

Cluster of homogeneous function		**Cluster of heterogeneous function**
atpB, atpE	*psbB, psbT*	*ndhE, psaC*
atpH, atpF, atpA	*rpoC1, rpoC2*	
psbE, psbF	*rpl22, rps3*	
psbK, psbI	*psbD, psbC*	
psbL, psbJ	*rps12, rps7*	
rpl33, rps18	*ndhG, ndhE*	
rpl16, rps8, rpl36, rps11		

Transcript levels of genes in clusters C and D were relatively high under various hormone stress conditions but low under chemical stresses. Average expression view of cluster A - D displayed severe up-and down-regulation of plastid genes, however, those of clusters E and F did not show significant alterations of plastid gene expression (Figure 11B). Again, co-expression of 14 groups of genes originating from same transcription units was identified

indicating that transcriptional processes contribute in controlling the abundance of transcripts originating from same operons (Table 12).

3.9 Use of Microarrays for the Evaluation and Comparison of Nuclear and Plastid Gene Expression under Various Stress Conditions

To compare stress induced expression patterns between plastid and nuclear genes, 79 plastid genes represented by the ATH1 microarray and 1,430 selected nuclear genes, likely to encode chloroplast proteins according to the already established localization and/or predicted transit peptides for chloroplast import were analyzed (Biehl et al., 2005). Hierarchical and K-means clustering identified 13 groups of transcriptomes and eight distinct co-regulated gene clusters, A – H, respectively (Figures 12A and 12B). The number of genes in each cluster ranged from 168 to 240 except cluster A, which contains 80 genes. Genes of each cluster were also categorized by their functions to identify a relationship between gene function and their expression behavior under stress conditions (Figure 12C). Remarkably, the functional categories were not evenly distributed in clusters B to H and often showed a preferential over- or under-representation. Generally, 30 - 50% of gene functions within clusters are still unknown.

Most plastid genes (87.5%) were grouped in cluster A indicating a concerted expression of plastid genes relative to nuclear genes. The categories of genes in cluster A are separated into nuclear and plastid genes (Figure 12C). Nuclear genes of this cluster fall into four major categories, gene expression, secondary metabolism, protein phosphorylation, and amino acid metabolism (Figure 12C). Plants stressed with chemicals, i.e. 2,4-dichlorophenoxyacetic acid, 4 thiazolidinone/acetic acid, furyl acrylate ester and hydrogen peroxide highly induced expression of genes specifically in cluster A. Conversely, transition from dark to different light qualities and norflurazon treatment reduced transcript abundance of most plastid genes in cluster A. Expression of genes in cluster B was generally up-regulated but down-regulated after the onset of light and chemical treatment. Genes for protein modification, fate and synthesis were almost absent in this cluster. The average expression view of genes in cluster C revealed a constant transcript level in the transcriptomes except down-regulation under biotic stress conditions (Figure 12B). Genes of cluster D were highly up-regulated after the onset of nutrients, Cs and norflurazon treatment, however, cycloheximide, 6-benzyl adenine, biotic M persicae, anoxia, and hypoxia stresses reduced expression of genes in cluster D. Cluster D contains the lowest portion of genes for photosynthesis light reaction.

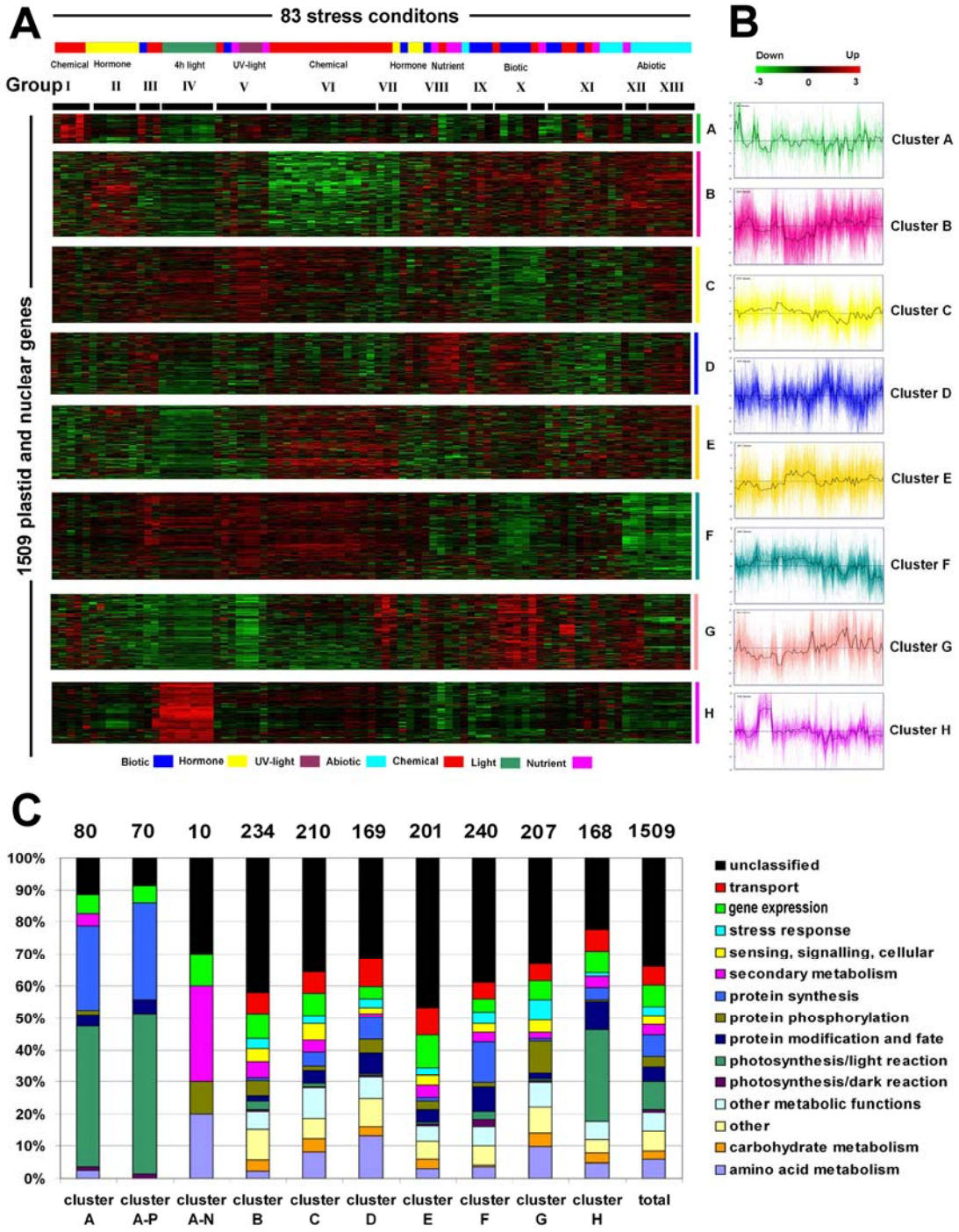


Figure 12. Expression Map and Distribution According to Functional Category of 79 Plastid and 1,430 Nuclear Genes Analyzed under 83 Stress Conditions.

(A) A total of 1,509 plastid and nuclear genes were clustered by genes and stress conditions using K-means (K=8) clustering based on the results of hierarchical clustering. Red, green, and black colors correspond to up-regulation, down-regulation, and unchanged, respectively. Clustering of biological conditions is indicated by different colors.

(B) Average expression views for each cluster are shown. The black line represents the average expression of genes in each cluster.

(C) The distribution into functional categories of genes in each cluster is shown (Biehl et al., 2005). The numbers of genes in each cluster are given above the histograms. Genes in cluster A are divided into plastid genes (A-P) and nuclear genes (A-N).

Various light treatments repressed gene expression and most of chemical stresses significantly induced transcript abundance of cluster E members. Cluster F contains 37.6% and 45.4% of genes for protein synthesis and photosynthesis dark reaction, respectively, but is almost lacking genes for carbohydrate metabolism. It showed similar numbers of up- and down-regulated genes under various conditions. In detail, genes of cluster F were up-regulated under different hormones, light qualities and chemical stresses but were down-regulated under abiotic, biotic and nutrient stress conditions. Cluster G comprises the highest number of genes for protein phosphorylation and stress response and is most surprisingly completely missing genes for photosynthesis light reaction. Only one gene encodes a protein for photosynthesis dark reaction. Transcript levels of this cluster were reduced under different light treatments, especially UV-B light stress. Genes were only up-regulated under biotic *P. infestans, B.cinerea, P. syringae* and ozone stress conditions. Cluster H is particularly intriguing because the three almost uniformly distributed categories within all clusters, stress response, protein phosphorylation and sensing are missing except two genes for stress response. Moreover, the categories are comparable to that of plastid genes in cluster A mostly due to its high proportion of nuclear photosynthetic light reaction genes (73.8%). With the exception of cluster G the remaining light reaction genes are distributed in clusters B to F. Surprisingly, genes for photosynthetic dark reaction are missing in this cluster. In accordance with this distribution expression of genes in cluster H was particularly highly induced under various light treatments and not remarkably differentially expressed under all other conditions chosen. Plastid genes of cluster A and nuclear genes of cluster H are rather reversely regulated especially under various light conditions. In summary, nuclear genes with chloroplast functions displayed tight co-regulation in eight distinctive clusters and are generally reversely regulated as compared to plastid genes enriched in a separate cluster. Clustering of genes was accompanied with an unequal distribution of functional categories indicating a tight correlation between gene function and expression pattern.

Clustering of genes implies that their promoters share common cis-elements allowing a coordinated expression mediated by transcription factors. Therefore, promoter motifs were searched in 1,000 bp upstream of the start codons of all genes in each cluster using Athena (http://www.bioinformatics2.wsu.edu/cgi-bin/Athena/cgi/home.pl). Whereas no significantly enriched promoter motifs were found in clusters A, C, E, and F, several motifs were significantly over-represented in clusters B, D, G, and H. Notably, eight among 13 highly enriched motifs in cluster H, like ABRE-like binding site motif, CACGTGMOTIF, ABRE

binding site motif, ABFs binding site motif, GBOXLERBCS, and ABREATRD22, are known to be responsible for light activation by G-box binding transcription factors (Table 13).

Table 13. Significantly Enriched Transcription Factor/Motif of Genes in Each Cluster.

Cluster	Transcription Factor/Motif	promoter's bound in subset		promoter's bound in genome		p-value
Cluster A	Not detected					
Cluster B	TATA-box Motif	83%	189	82%	24789	0.002
	ABRE-like binding site motif	24%	56	20%	6258	0.008
	CACGTGMOTIF	19%	44	15%	4546	0.007
	ABRE binding site motif	7%	18	4%	1422	0.006
Cluster C	Not detected					
Cluster D	MYB binding site promoter	33%	56	30%	8915	0.005
	TELO-box promoter motif	15%	26	10%	3062	0.001
Cluster E	Not detected					
Cluster F	Not detected					
Cluster G	TATA-box Motif	85%	174	82%	24789	< 10e-3
	MYB1AT	82%	168	85%	25733	0.002
	W-box promoter motif	71%	146	67%	20289	< 10e-4
	Ibox promoter motif	41%	85	40%	12259	0.004
Cluster H	ABRE-like binding site motif	50%	84	20%	6258	< 10e-10
	Ibox promoter motif	48%	81	40%	12259	< 10e-4
	ACGTABREMOTIFA2OSEM	43%	72	14%	4398	< 10e-10
	CACGTGMOTIF	33%	55	15%	4546	< 10e-10
	GADOWNAT	23%	39	8%	2579	< 10e-9
	ABRE binding site motif	22%	37	4%	1422	< 10e-10
	ABFs binding site motif	21%	35	3%	982	< 10e-10
	GBOXLERBCS	18%	31	2%	800	< 10e-10
	TGA1 binding site motif	13%	23	3%	951	< 10e-10
	UPRMOTIFIAT	13%	23	3%	951	< 10e-10
	UPRMOTIFIIAT	12%	21	3%	1121	< 10e-7
	ABREATRD22	11%	19	2%	736	< 10e-8
	GBF1/2/3 BS in ADH1	6%	11	1%	476	< 10e-4

Chromosomal positioning of genes within each cluster was rather random as has been shown by chromosome map tool (http://www.arabidopsis.org/) (data not shown) indicating that co-regulation of gene expression does not correlate with the chromosomal location as has already been suggested for genes encoding organelle proteins (Legen et al., 2001; Alexeyenko et al., 2006). Hierarchical clustering identified 13 groups of transcriptomes. Interestingly, related treatments were highly enriched within individual groups indicating that the responses were rather specific and that the signatures in the identified transcriptome clusters are indicative of the functional state of the plant and the stress applied.

3.10 Distribution of Plastid and Nuclear Gene Expression Ratios under Different Biological Conditions

No differences in the distribution of plastid gene expression ratios could be observed between transcriptomes of mutants and different biological conditions as deduced from the macroarray system (Figure 13A). The distribution of expression ratios of plastid and nuclear genes for chloroplast proteins obtained from transcriptomes of 136 mutants not affected in chloroplast functions also did not differ as calculated from ATH1 microarrays (Figure 13B). However, under 83 stress conditions, particularly the expression of nuclear genes displayed a wider range of alteration as compared to that of plastid genes (Figure 13C).

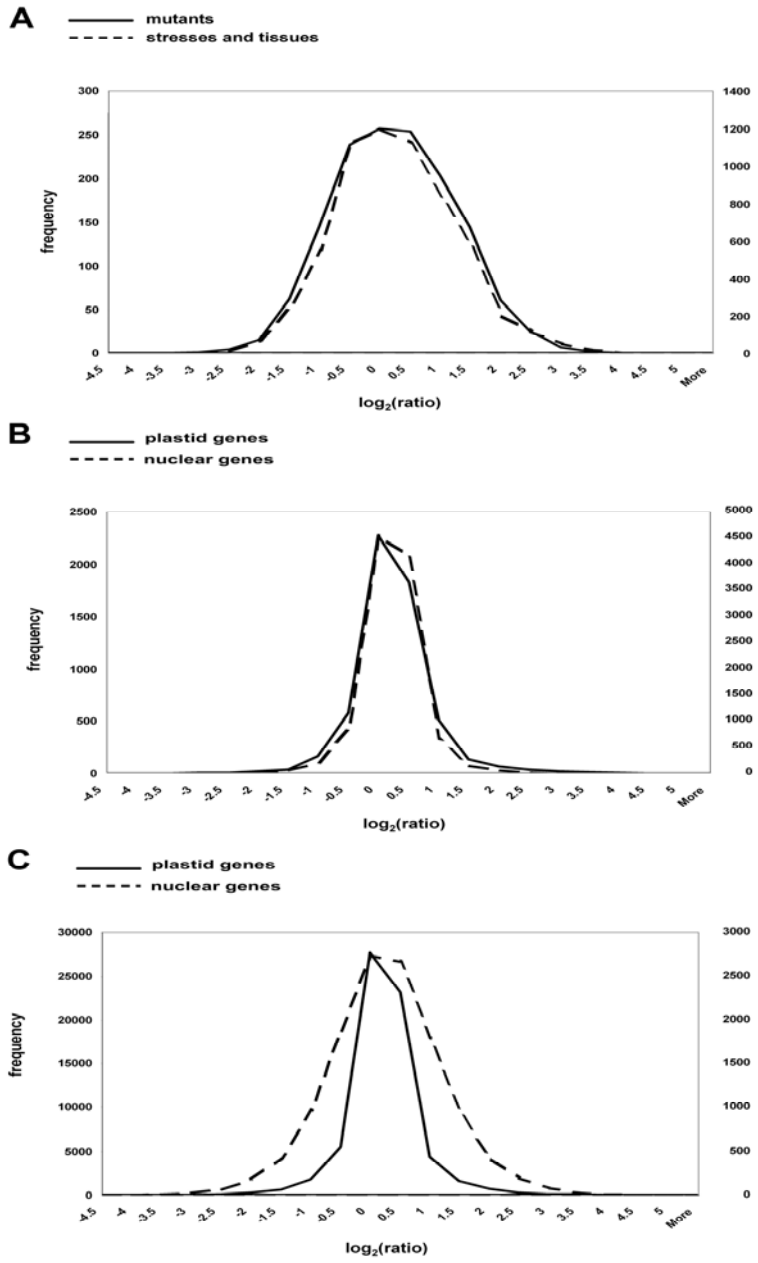


Figure 13. Distribution of Fold Change Values of Plastid and Nuclear Gene Expression in Mutant Transcriptomes and WT Plants Exposed to Various Biological and Stress Conditions.

(A) Distribution of fold change values of plastid gene expression using 75 mutants and 14 biological conditions (including tissue comparisons). The log_2-transformed fold change values were normalized. Left and right y-axes represent the frequency of mutants and biological conditions, respectively.

(B) Distribution of fold change values of plastid and nuclear gene expression using 136 mutants not primarily affected in chloroplast functions. The log_2-transformed fold change values were normalized. Left and right y-axes represent the frequency of plastid and nuclear genes, respectively.

(C) Distribution of fold change values of plastid and nuclear gene expression using 83 biological conditions. The log_2-transformed fold change values were normalized. Left and right y-axes represent the frequency of plastid and nuclear genes, respectively.

3.11 Plastid Gene Expression in Response to Different Light Qualities during Early Seedling De-etiolation

Light-dependent expression data of plastid and nuclear genes were extracted from AtGenExpress (http://www.arabidopsis.org/info/expression/ATGenExpress.jsp) and GEO databases (http://www.ncbi.nlm.nih.gov/geo). Four days dark-adapted seedlings were transferred to four different light qualities [far red, red (R), blue, and heterochromatic light] for 45 minutes and four hours. Irrespectively of light conditions expression ratios of the majority of plastid genes were within two times fold changes after 45 min irradiation (Figure 14A). Only few genes displayed expression ratios higher than 1.5. Surprisingly, when etiolated seedlings were transferred for four hours to different light regimes plastid genes tend to be significantly down-regulated. Only few plastid genes like *psbB*, *psbC, psbD,* and *rbcL* were up-regulated after four hours irradiation especially under far red and blue light. To test the involvement of phytochrome (Phy)A and PhyB signalling in regulation of plastid gene expression, the effect of one hour R light irradiation in *phyA*, *phyB*, and *phyAB* double mutants is analyzed (Tepperman et al., 2006).

Expression profiles of irradiated mutant and WT seedlings were compared with those obtained from four days dark-adapted seedlings (Figure 14A). The gene expression pattern of some genes after one hour R light irradiated WT plants was quite different from that after 45 minutes but close to that of four hours R light treatment. Some deviations in plastid gene expression obtained after 45 minutes (AtGenExpress) and one hour (Tepperman et al., 2006) may be explained either by a different experimental set up, like germination conditions, light exposure time, and light quantity and/or different normalization methods. However, the expression patterns of plastid genes after one and four hour R light irradiation were closely related to each other showing the tendency to be down-regulated with the exception of *rbcL, psbD,* and *psbC*, which are up-regulated after four hours R light treatment. Unexpectedly, most plastid genes were highly induced in *phyA, phyB* and *phyAB* knockout mutants already after application of one hour R light. The expression of only few genes, i.e. *psaJ, psbJ, psbL, petD*, and *rpl33,* were found to be slightly down-regulated or still unchanged in Phy mutants. This indicates that PhyA and PhyB signaling plays a crucial role in suppression of plastid gene expression during early seedling de-etiolation.

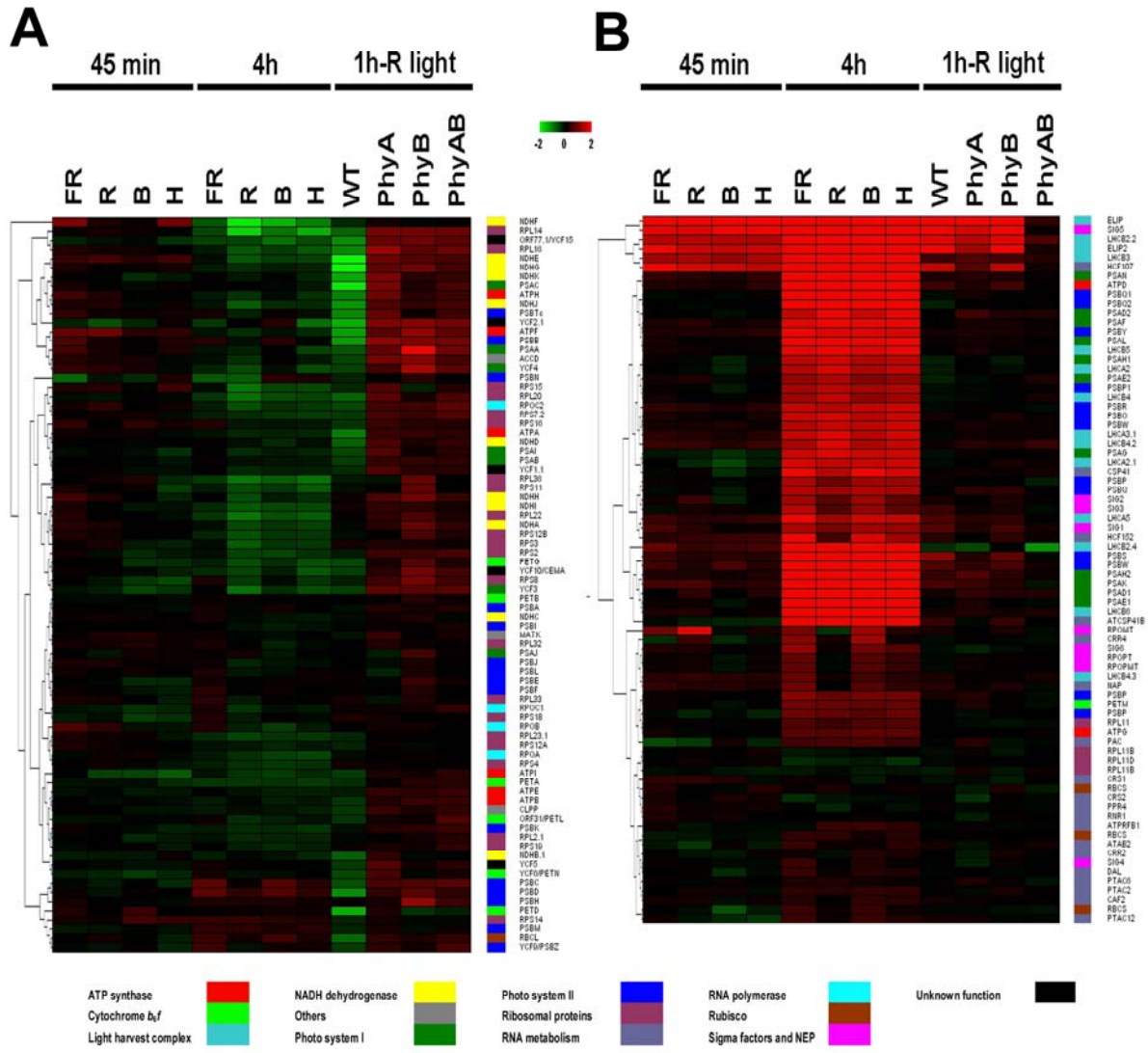


Figure 14. Expression Profiles of Plastid and Nuclear Genes for Chloroplast Function under Various Light Qualities during Seedling De-etiolation in WT and Phytochrome Mutants.

The expression of 79 plastid and 76 nuclear genes, which are related to photosynthesis and chloroplast mRNA metabolism, were analyzed under different light qualities [FR (far red), R (red), B (blue), and H (heterochromatic) light] (AtGenExpress) and in *phyA, phyB* and *phyAB* null mutants. Fold changes are expressed relative to dark-adapted plants. The data of *phyA*, *phyB* and *phyAB* null mutants were retrieved from microarray analysis (Tepperman et al., 2006).

(A) Expression profiles of 79 plastid genes.

(B) Expression profiles of 76 nuclear genes for photosynthesis and chloroplast mRNA metabolism.

The impact of Phy signaling on the expression patterns of nuclear genes encoding components of the thylakoid membrane and the plastid transcript metabolism under different light qualities were analyzed (Figure 14B). References for chloroplast mRNA metabolisms genes are provided (Table 14). Transcripts of most nuclear genes tend to be highly up-

regulated with increasing exposure time from 45 minutes to four hours. Interestingly, ELIP, ELIP2, LHCB2.2, LHCB3, HCF107 and SIG5 were already highly up-regulated after 45 min irradiation and still up-regulated in the single but not in the double Phy knockouts indicating a partially redundant role and/or cross-talk of PhyA and PhyB on early up-regulation of these genes under the chosen conditions. A crosstalk of different phytochromes has well been documented for other signalling pathways (Casal, 1996; Canton and Quail, 1999; Hennig, et al., 2001; Torres-Galea et al., 2006).

Exceptionally, expression of few genes, i.e. SIG4 and RBCS, remained almost unchanged under all light conditions even after four hours illumination and that of the transcription machinery, i.e. SIG6, RPOPT, RPOPMT, RPOMT, encoding the mitochondrial NEP, and LHCB4.3was not responsive solely under red light. Furthermore, most nuclear genes required for plastid mRNA metabolism were somehow upregulated after the onset of four hours irradiation of all four light qualities except *CRS1, CRS2, PPR4,* and *RNR1* showing a light-induced down-regulation (Figure 14B). In contrast to photosynthesis related genes those for the plastid mRNA metabolism do not follow the apparent light response indicating that genes of the two categories are integrated into different regulatory expression networks (Figure 14).

Surprisingly, in sharp contrast to plastid genes, expression of nuclear genes in *phyA* and *phyB* mutants was comparable to the WT after one hour red light irradiation. Moreover, only minor differences in nuclear gene expression could be found between 45 min and one hour R light induction but plastid gene expression was severely downregulated after one hour as compared to 45 min light treatment. This indicates that the chloroplast transcriptome reacts rather sensitive to red light as compared to nuclear genes encoding photosynthetic proteins during early seedling de-etiolation and that PHY signaling is involved in the immediate response.

Interestingly, the expression response of plastid genes observed immediate after the onset of light in young seedlings is different from the later response and from the response in leaves indicating that regulation of mRNA stability and transcriptional activation are not synchronized. Phytochrome response is tissue specific (down-regulation is hypocotyl and up-regulation in cotyledons). On the other hand, the biomass of etiolated seedlings used for expression profiling (Tepperman et al., 2006) is mostly determined by the hypocotyl and less prominent by cotyledons indicating that phytochrome A and B signaling strongly and rapidly affects expression of plastid and nuclear genes in this tissue.

Table 14. Nuclear Factors Known to be Involved in Chloroplast mRNA Metabolism.

NAME	Source	AT-accession	Target	Function	References
CRR4	*Arabidopsis*	At2g45350	*ndhD*	RNA editing of *ndhD*	Okuda et al., 2006
HCF152	*Arabidopsis*	AT3G09650	*psbB-psbT-psbH-petB-petD RNAs*	Processing and splicing of *psbB-psbT-psbH-petB-petD* RNAs	Meierhoff et al., 2003
HCF107	*Arabidopsis*	AT3G17040	*psbH*	Processing of psbH	Felder et al., 2001
ATCSP41B	*Arabidopsis*	At1g09340	*petD*	*petD* RNA binding protein	Raab et al., 2006
RNR1	*Arabidopsis*	At5g02250	-	Maturation of chloroplast ribosomal RNAs	Bollenbach et al., 2005
CRR2	*Arabidopsis*	AT3G46790	*rps7, ndhB*	Intergenic processing between *rps7* and *ndhB*	Hashimoto et al., 2003
DAL	*Arabidopsis*	At2g33430	*rRNA*	Maturation of the plastid ribosomal RNAs	Bisaz et al., 2003
ATPRFB1	*Arabidopsis*	AT5G36170	*mRNAs*	Regulation of both mRNA stability and protein synthesis	Meurer et al., 2002
PAC	*Arabidopsis*	AT2G48120	*mRNAs*	Maturation of specific chloroplast mRNAs	Meurer et al., 1998
SIG3	*Arabidopsis*	AT3G53890	*psbN*	Sigma factor	Zghidi et al., 2007
SIG6	*Arabidopsis*	AT2G36990	PEP	Sigma factor	Ishizaki et al., 2005
SIG2	*Arabidopsis*	AT1G08540	-	Sigma factor	Kanamaru et al., 2001
SIG1	*Arabidopsis*	AT1G64860	-	Sigma factor	Privat et al., 2003
SIG5	*Arabidopsis*	AT5G24120	-	Sigma factor	Tsunoyama et al., 2004
SIG4	*Arabidopsis*	AT5G13730	*ndhF*	Sigma factor	Favory et al., 2005
ATAB2	*Arabidopsis*	AT3G08010	*psaB*	Translation of *psaB*	Barneche et al., 2006
PTAC2	*Arabidopsis*	At1g74850	-		Pfalz et al., 2006
PTAC6	*Arabidopsis*	At1g21600	-		Pfalz et al., 2006

Table 14. Continued

NAME	Source	AT-accession	Target	Function	References
PTAC12	*Arabidopsis*	At2g34640	-		Pfalz et al., 2006
HCF173	*Arabidopsis*	At1g16720	-	Initiation of Translation of the *psbA* mRNA	Schult et al., 2007
HCF153	*Arabidopsis*	At4g31560	-	post-translational step in biogenesis of the cytochrome bf	Lennartz et al., 2006
NAP	*Chlamydomonas*	At5g56950	*tscA*	A nucleosome assembly protein-like polypeptide binds to chloroplast group II	Glanz et al., 2006
MCD3	*Chlamydomonas*	-	-	RNA maturation and degradation	Rymarquis et al., 2006
MCD4	*Chlamydomonas*	-	-	RNA maturation and degradation	Rymarquis et al., 2006
MCD5	*Chlamydomonas*	-	-	RNA maturation and degradation	Rymarquis et al., 2006
MCD1-1	*Chlamydomonas*	-	*petD*	*petD* mRNA stability	Erickson et al., 2005
MCD1-2	*Chlamydomonas*	-	*atpA*	atpA mRNA stability	Erickson et al., 2005
RAA1	*Chlamydomonas*	-	*psaA*	Trans-splicing to mature *psaA* mRNA	Perron et al., 2004; Rochaix et al., 2004; Merendino et al., 2006
RAA2	*Chlamydomonas*	-	*psaA*	Trans-splicing to mature *psaA* mRNA	Perron et al., 2004; Rochaix et al., 2004; Merendino et al., 2006
RAA3	*Chlamydomonas*	-	*psaA*	Trans-splicing to mature *psaA* mRNA	Perron et al., 2004; Rochaix et al., 2004; Merendino et al., 2006
NAC2	*Chlamydomonas*	-	*psbD*	*psbD* mRNA stability	Boudreau et al., 2000
RBP40	*Chlamydomonas*	-	*psbD*	Interact with *Nac2* for translation of *psbD* mRNA	Ossenbühl et al., 2000
RB60	*Chlamydomonas*	-	-	modulates the binding of RB47	Kim and Mayfield, 1997; Trebitsh et al., 2000
RB47	*Chlamydomonas*	-	*psbA*	binding protein to the 5'-UTR of the *psbA* mRNA	Yohn et al., 1998a, 1998b
RB55	*Chlamydomonas*	-	-	-	Danon and Mayfield, 1991; Mayfield et al.,1994; Yohn et al., 1996

Table 14. Continued

NAME	Source	AT-accession	Target	Function	References
RB38	*Chlamydomonas*	-	-	-	Danon and Mayfield, 1991; Mayfield et al.,1994; Yohn et al., 1996
CRS1	*Maize*	AT5G16180	*atpF*	Splicing of the *atpF* intron	Jenkins et al., 1997
CRS2	*Maize*	AT5G19830	-	Splicing of many chloroplast introns	Jenkins et al., 1997
CAF1	*Maize*	AT3G25430	-	Promote splicing of group II introns in maize	Ostheimer et al., 2003
CAF2	*Maize*	AT1G23400	-	Promote splicing of group II introns in maize	Ostheimer et al., 2003
PPR2	*Maize*	At3g06430	*rRNA*	Plastid ribosome accumulation	Williams and Barkan, 2003
PPR4	*Maize*	AT5g04810	*rps12*	*rps12* trans-splicing factor	Schmitz-Linnenweber et al., 2006
P54	*Sinapis alba L.*		*mRNAs*	Endoribonuclease required for 3' end processing of plastid precursor transcripts	Liere and Link, 1997
CSP41	*Spinach*	AT3G63140	*petD*	Endoribonuclease required for 3' untranslated region of *petD* mRNA	Yang and Stern, 1997

4 DISCUSSION

4.1 Analysis of Chloroplast Transcriptomes

This study is the first comprehensive analysis of chloroplast gene expression in higher plants in response to mutations in various chloroplast functions and light treatments as well as abiotic and biotic stress conditions using macroarray and microarray data. To gain deeper insights chloroplast and nuclear expression profiles of WT grown under a variety of biological conditions and of mutants have been analyzed and compared with each other.

To confirm the reliable application of our macroarray system several control experiments have been performed, like the use of mutants, known to be affected primarily in the plastid mRNA metabolism or with deficiencies in other plastid functions, representative RNA gel blot analysis, different clustering methods and array systems. Moreover, scatter-plot analysis, repetitions and calculated p-values confirmed the reproducibility of the system. The high reliability of the macroarray-based approach is further demonstrated by clustering of almost all plastid genes, which are co-transcribed in same operons (Table 7). For example, co-transcribed genes with homogeneous functions, like *psbD*/*psbC*, *rpoC1*/*rpoC2*, and *atpF*/*atpA,* or heterogenous functions like *psaA*/*psaB*/*rps14* and *psbB*/*psbT*/*psbH*/*petB*/*petD* were identified as co-expressed genes. Many genes, which are not co-transcribed are also tightly co-expressed demonstrating co-regulation of different promoters or at the posttranscriptional level (Table 7).

4.2 Hierarchical Clustering of Plastid Mutant Transcriptomes Identified Two Distinguishable Signatures and Novel Mutants Impaired in mRNA Metabolism

Hierarchical clustering of transcriptomes of mutants affected in chloroplast functions revealed two groups, I and II. Mutants of group I mostly exhibited an albino phenotype with dramatic deficiencies in chlorophyll amount. Group II consists preferentially of expression profiles of *hcf* and pale green mutants, various tissues, and environmental conditions. Transcriptomes of group I were generally reversely regulated as compared to those of group II (Figure 9). Non-photosynthetic genes were preferentially up-regulated whereas photosynthetic genes tend to be down-regulated in transcriptomes of group I (Figure 9). The plastid gene expression profile of tobacco PEP mutants revealed that most genes for photosynthesis and ribosomal RNAs were downregulated and that for ribosomal proteins and RNA polymerase were up-regulated

(Legen et al., 2002). Interestingly, the signature of group I transcriptomes resembled the expression pattern of PEP mutants in tobacco (Legen et al., 2002), plastid gene expression mutants in *Arabidopsis* (Pfalz et al., 2006), and WT lines treated with lincomycin (Gray et al., 2002; Koussevitzky et al., 2007). Genes transcribed by the NEP and the PEP were generally severely up- and down-regulated, respectively. Therefore, mutants of group I are expected to be primarily affected in plastid gene expression at various levels allowing preferentially or only transcription of plastid genes by the NEP. The important role of NEP and PEP for the general gene expression system in all tissues during plant development is consistent with the assumption that both polymerases are active in non-photosynthetic tissue and important for early development of the chloroplast during germination (Demarsy et al., 2006). Several albino mutants are also present in group II, indicating that they are not primarily impaired in plastid gene expression, like *alb3* (Ossenbühl et al., 2004) and *vipp1* (Aseeva et al., 2007), known to have deficiencies at the posttranslational level. Therefore, albino mutants primarily affected in plastid gene expression and other posttranslational processes were discerned using plastid gene expression profiling.

Reduced levels of the plastid rRNAs (16S and/or 5S) in mutants of group I indicate a severe loss of translation and therefore loss of transcripts generated by the PEP. This makes it difficult to define the primary cause of the lesions just by comparing transcript levels. It appeared that not the magnitude and the diversity of gene expression changes are suitable to select mutants primarily affected in plastid mRNA metabolism but a departure from the expression signature characteristic for members of group I and II. Out of 78 lines analyzed 8 and 6 bona fide plastid mRNA metabolism mutants in group I and II, respectively, have been identified (Table 8), among them the control mutants *hcf145, pac* and *atprfB1*. Representative RNA gel blot analysis finally confirmed the defect in the plastid mRNA metabolism mutant *crp135* (Figure 6). Affected genes were either up- or down-regulated (Figure 6). The remaining members of group II are expected to be impaired in other plastid functions.

4.3 Hierarchical Clustering of Plastid Genes in Mutants of Chloroplast Functions Identified Two Transcriptionally Determined Gene Clusters

Clusters A and B represent genes, which are known to be preferentially transcribed by the NEP and the PEP, respectively (Shiina et al., 2005; Liere and Börner, 2006). Therefore, it appears that plastid gene expression in mutants impaired in chloroplast functions is mainly under transcriptional and less prominent under posttranscriptional control. It also shows that

NEP and PEP reversely control different sets of genes. Although posttranscriptional events are thought to be major players in the determination of plastid transcript abundance the mutant transcriptomes unequivocally demonstrate the predominant transcriptional determination of gene expression in the mutants (Figure 15).

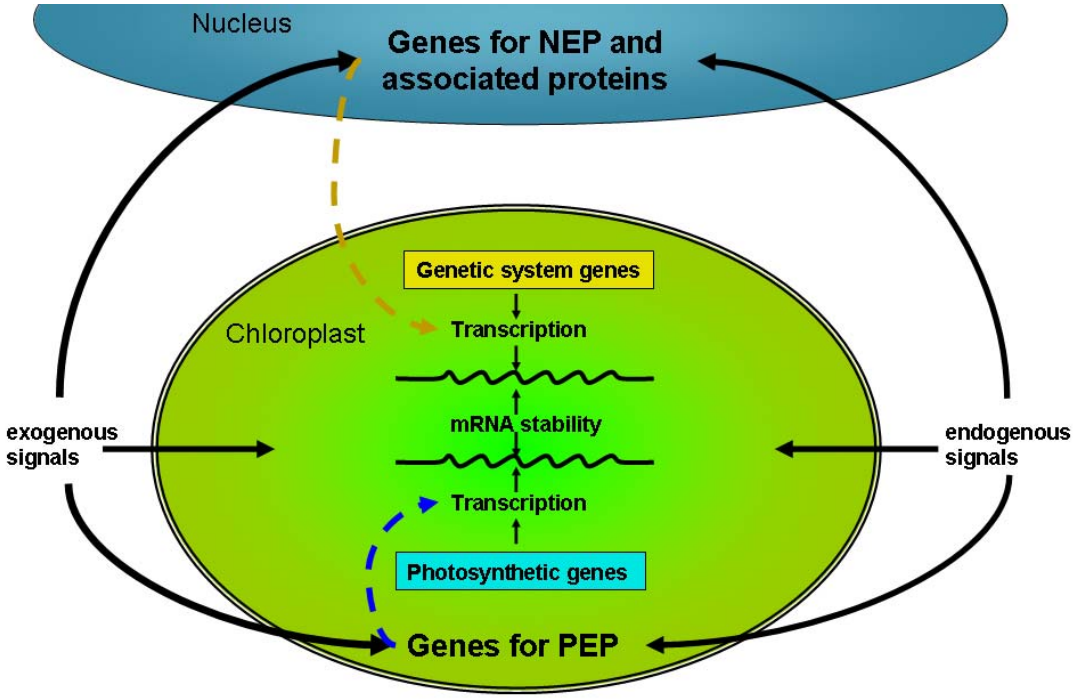


Figure 15. Determination of chloroplast transcript levels by regulation of RNA polymerases and mRNA stability.
Chloroplast genes contain promoters for two kinds of RNA polymerases, nuclear-encoded phage-like RNA polymerases (NEP) and a plastid-encoded RNA polymerase of cyanobactreial origin (PEP), which is regulated by nuclear encoded sigma factors. control The expression and activity of the RNA polymerases as well as plastid mRNA stability respond to exogenous and endogenous signals.

In contrast, clustering of plastid genes of WT exposed to various stress conditions, which also affect the chloroplast, failed to detect two major transcriptionally determined gene clusters. Only two, E and F, out of six clusters preferentially contain genes transcribed by the NEP and the PEP, respectively. However, most co-transcribed genes from same operons are also clustered as they are in general functionally tightly coupled (Table 12). Therefore, it is likely that plastid gene expression is mostly regulated on the posttranscriptional level under natural conditions for fine-tuning and adaptation to environmental changes (Gruissem and

Tonkyn, 1993). This is further supported by polymerase type dependent posttranscriptional processing of plastid transcripts (Legen et al., 2002). Posttranscriptional regulation seems to collapse in mutants severely and generally affected in chloroplast gene expression, like in numerous albino mutants of group I. This might also be the reason why fold changes and the diversity of gene expression changes were often lower in mutants affected in chloroplast functions than that under various stress situations applied to the WT (Figures 7 and 8).

4.4 Nuclear Genes for Plastid Components Displayed Dynamical Gene Expression Patterns, Eight Major Co-Regulated Clusters and 13 Transcriptome Groups

K-means clustering of 79 plastid and 1,430 nuclear genes for chloroplast components revealed eight different clusters (A-H) containing 80 - 240 tightly co-regulated genes under various stress conditions. 87.5 % of plastid genes grouped in cluster A suggesting that the gene expression machinery of the chloroplast is rather conserved as compared to that of the nucleus making differential expression of chloroplast genes difficult to analyze. 10 nuclear genes encoding are found in cluster A suggesting a concerted nuclear and plastid expression system. Reversely, 10 plastid genes followed the expression pattern of clusters B, C, E and F.

The predominance of individual functional categories, like gene expression, ribosomal subunits, protein phosphorylation, and photosynthesis in clusters E, F, G, and H, respectively, suggests that genes of related pathways in organelle biogenesis and function have been integrated into same regulatory networks that control related pathways. On the other hand, functional categories, like, photosynthesis light reaction, dark reaction, and carbohydrate metabolism were almost missing in clusters D, H, and F, respectively. Therefore, it was important for horizontally transferred and de novo evolved genes for chloroplast proteins to acquire specific promoter elements allowing the integration into the gene regulatory network and the adaptation of co-regulated expression systems to the requirements of plants under the respective conditions (Herrmann 1997; Martin and Herrmann, 1998; Herrmann and Westhoff, 2001). Inspection of co-regulated genes revealed the appearance of enriched promoter elements within the clusters (Table 13). Up to 50% of the genes present in cluster H share same promoter elements partially involved in light-induced expression (Hudson and Quail, 2003).

Hierarchical clustering using Affymetrix data defined 13 groups of transcriptomes, which largely correspond with the respective biotic and abiotic treatments (Figure 12A). The identified expression signatures are therefore indicative for the functional state of the plant

and can be utilised to estimate stress responses. A strong light induction is the most prominent criteria for genes present in cluster H. In accordance with that, about 37.2% of known genes present in this cluster encode photosynthetic components whereas the function of about 23.21 % of all genes is still unknown suggesting that about one third of them encode novel photosynthetic components. Furthermore, clustering of transcriptomes provides a framework for classification of nuclear promoters related to photosynthetic functions. On the other hand, the tight co-regulation of genes within the eight clearly discernable clusters and the high proportion of yet unknown gene functions will offer a means to elucidate their roles especially under the conditions chosen.

This observation largely escapes from the previous described 'master switch' defining three transcriptome groups, two major groups showed a preferential up and down regulation of almost all genes, and a smaller group containing only eight trancriptomes displayed a mixed response (Richly et al., 2003). Surprisingly, only 79 out of 168 genes present in the photosynthetic cluster H, correspond with genes previous identified in two photosynthetic regulons containing 188 genes (Biehl et al., 2005). Many of the genes (22.91%) present in cluster F co-related with genes found in both photosynthetic regulons. Basically cluster H and F are very similar except the strong light induction found in cluster H and a relative strong UV-B light induction found in cluster F. Furthermore, the distribution of gene members of other identified regulons (Biehl et al., 2005) within the clusters defined in this work was rather random and a bias of regulon members in even one of the clusters was not found (data not shown). In addition, this transcriptomic approach identified seven discernable gene clusters as compared to the previous identified 23 regulons using the same gene set. Apparently, the discrepancy between the data of Biehl et al., (2005) and this work is very likely the result of different array systems, normalisation methods, statistical and clustering analysis, which results in the identification of different co-regulated gene clusters and groups of transcriptomes. Therefore, groups of nuclear genes for chloroplast proteins respond similarly to related treatments and show a rather complex response.

4.5 Phytochrome Signalling Suppresses Plastid Gene Expression during Early Seedling De-Etiolation before Nuclear Genes Start to Respond

Compared to extensive genome wide expression analysis of nuclear genes, expression of plastid genes in the framework of light conditions was mostly restricted to individual genes, notably *rbcL*, *psbA* and *psbD/C* (Sexton et al., 1990; Baumgartner et al., 1993; Kim et al.,

1993; DuBell et al., 1995). Therefore, the role of Phy-mediated response of nuclear gene expression is well studied but little is known about the role of photoreceptors in global regulation of plastid gene expression. Significant down regulation of most plastid genes in response to green light has been reported (Dhingra et al., 2006) although the photoreceptor specific for green light is still unknown.

Here, it has been shown that light-dependent development of four days dark-grown seedlings occurs concomitantly with a general down-regulation of plastid gene expression regardless of the light quality. In contrast, plastid gene expression was highly up-regulated in *phyA*, *phyB,* and *phyA/B* double mutants after the onset of one hour illumination. This data strongly indicates that PhyA and PhyB are involved in the light mediating signaling pathway to down-regulate plastid gene expression in de-etiolated seedlings within the first hour. Recent studies revealed that levels of cpDNA in maize decreased dramatically when dark-grown plants were transferred to light due to cpDNA degradation which is triggered by light (Oldenburg et al., 2006). Although, a transient drop in cpDNA levels have been shown not to affect levels of proteins under normal conditions (Zoschke et al., 2007), changes in the amount of cpDNA could explain reduced transcript levels after the onset of dark-light switches. Interestingly, the expression response of plastid genes observed immediate after the onset of light in young seedlings mostly consisting of non-photosynthetic tissue, like roots and hypocotyls, is different from the response in photosynthetic tissue showing a general light-induced up-regulation (Monde et al., 2000). This shows that regulation of mRNA stability and/or transcriptional activation are not synchronized in both types of tissues indicating a rapid and tissue-specific response of plastid gene expression to phytochrome A and B signaling.

Generally, genetic system genes did not follow the strong light induction characteristic for most photosynthetic genes (Figure 14) with the exception of *sig5, sig1, csp41, hcf107,* and *hcf152,* which are also highly light induced in dark-grown seedlings. However, *sig5* does not respond to red light in mature leaves and to heterochromatic light in light-adapted young seedlings grown on sucrose-containing medium (Tsunoyama et al. 2002; Ishizaki et al., 2005) indicating that integration of light signals depend on the developmental stage and energy input. Expression of *sig2, sig3, sig4,* and *sig6* is preferentially under developmental control whereas that of *sig1* and *sig5* primarily responds to light signals (Shiina et al., 2005). This is consistent with the circadian expression of *sig1* and light-induced activation of *psbD* and *psbA* gene expression mediated by *sig5*. Although sigma factors can be distinguished due to their specific roles, none of the mutants identified is lethal indicating overlapping functions as well

(Liere and Börner, 2006). The data show that expression of genetic systems genes are integrated into different gene regulatory networks.

Although several DNA binding proteins, peptides associated with sigma factors as well as nucleoid-associated proteins have been described (Shiina et al., 2005; Liere and Börner, 2006), their regulatory role in plastid transcription remained largely unknown. In addition, three unidentified factors bind to *psbD* promoter elements indicating an important role in transcriptional regulation (Kim and Mullet, 1995; Baba et al., 2001). Except the sigma factor mutants identified in reverse genetic screens, all other nuclear mutants with altered levels of plastid transcripts mostly isolated in forward genetic screens are deficient in gene specific or general posttranscriptional processes (Table 14) (reviewed in Bollenbach et al., 2004). Since alone 14 genes are involved in processing of only *psaA* in *Chlamydomonas* (Merendino et al., 2006) and the genetic screens are far from being saturated, several hundred nuclear genes might be expected to play crucial roles in regulation of posttranscriptional processes.

Posttranscriptional RNA modifications are relevant not only for the control of transcript abundance but also for generation of spliced, edited, endo- and exonucleolytically cleaved plastid transcripts in order to generate translational-competent mRNAs (Barkan et al., 1994; Hirose and Sugiura, 1997; Felder et al., 2001). Processing of plastid primary transcripts seems to be especially relevant to regulate both transcript abundance and translation of individual gene segments of polycistronic mRNAs to fine tune regulation of gene expression independent from transcriptional control. This is further supported by the finding that clustering of plastid genes under various biological conditions is less prominent under global transcriptional control.

In summary, the present report provides a rich source of information to investigate the involvement of the chloroplast and the role of yet unknown nuclear genes in the management and gene expression of this organelle especially under abiotic and biotic stress conditions. On the other hand, expression signatures can be used to estimate the functional state of the plant and to identify mRNA metabolism mutants, which escape from the general response. The acquisition of a large number of nuclear genes significantly increased the complexity of the plastid mRNA metabolism during endosymbiosis. The frequent occurrence of plant specific genes important for chloroplast mRNA homeostasis demonstrates that transcript regulation represents a fast evolving process during evolution resulting in plant specific expression systems.

SUMMARY

Chloroplast mRNA metabolism, defined by transcriptional as well as numerous posttranscriptonal processes is regulated by tissue-specificity, environmental signals as well as developmental programs and depends on numerous largely unknown proteins encoded by nuclear genes. Here, chloroplast macroarrays using 78 mutants affected in various chloroplast functions as well as different biological conditions and tissues in *Arabidopsis* have been generated and reliably applied. These data have been complemented by microarray data including 79 plastid and 1,430 nuclear genes likely to encode chloroplast proteins. Based on clustering analysis applied to 323 transcriptomes eight major conclusions can be drawn. (1) The plastid transcriptome reacts rather sensitive upon adaptation to environmental changes and to mutations in various plastid functions. (2) Phytochrome (PHY) A and B mediated responses of plastid gene expression are much faster and more prominent than that of nuclear genes during early stages of de-etiolation. (3) Hierachical clustering revealed that plastid and nuclear genes encoding photosynthetic components are generally reversely regulated under various stress conditions. (4) The obtained transcript profiles are suitable to identify nuclear mutants primarily affected in chloroplast RNA metabolism. Either levels of one or few transcripts or the general expression pattern were affected in the mutants. (5) According to the expression signatures mutants could be classified into two major groups; one group was indicative for albino mutants, supposed to be affected in the general plastid mRNA metabolism and the other group was mainly impaired in other chloroplast functions or with specific deficiencies in plastid mRNA metabolism. (6) Cluster analysis classified plastid genes into two reversely co-expressed clusters A and B containing mostly non-photosynthetic (genetic system) and photosynthetic genes, respectively. Genes of cluster A and B contain promoters for the nuclear and plastid encoded RNA polymerase, respectively, indicating that global clustering occurs via transcriptional control. (7) 10 sub-clusters of co-expressed plastid genes have been defined. They mostly contain co-transcribed but also independently transcribed and often functionally tightly associated genes suggesting additional clustering via posttranscriptional control. (8) The *Arabidopsis* macroarray can be reliably applied heterologously in a wide range of plant species due to the high conservation of plastid genes. In conclusion, the data presented in this work demonstrate that integration of chloroplast functions into the ontogenetic program of the plant cell was established via controlling and clustering of plastid and nuclear gene expression mediated by nuclear-encoded factors.

REFERENCES

Alexeyenko, A., Millar, A.H., Whelan, J., and Sonnhammer, E.L. (2006). Chromosomal clustering of nuclear genes encoding mitochondrial and chloroplast proteins in *Arabidopsis*. Trends Genet. **22:** 589-593.

Allison, L.A. (2000). The role of sigma factors in plastid transcription. Biochimie **82:** 537-548.

Archibald, J.M. (2005). Jumping genes and shrinking genomes-probing the evolution of eukaryotic photosynthesis using genomics. IUBMB Life **57:** 539-547.

Aseeva, E., Ossenbühl, F., Sippel, C., Cho, W.K., Stein, B., Eichacker, L.A., Meurer, J., Wanner, G., Westhoff, P., Soll, J., and Vothknecht, U.C. (2007). Vipp1 is essential for basic thylakoid membrane formation but not for the assembly of thylakoid protein complexes. Plant Physiol. Biochem. **45:** 119-128.

Baba, K., Nakano, T., Yamagishi, K., and Yoshida, S. (2001). Involvement of a nuclear-encoded basic helixloop-helix protein in transcription of the light-responsive promoter of *psbD*. Plant Physiol. **125:** 595-603.

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

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

Barneche, F., Winter, V., Crevecoeur, M., and Rochaix, J.D. (2006). ATAB2 is a novel factor in the signalling pathway of light-controlled synthesis of photosystem proteins. EMBO J. **25:** 5907-5918.

Barrett, T., and Edgar, R. (2006). Mining microarray data at NCBI's gene expression omnibus (GEO). Methods Mol. Biol. **338:** 175-190.

Baumgartner, B.J., Rapp, J.C., and Mullet, J.E. (1993). Plastid genes encoding the transcription/translation apparatus are differentially transcribed early in barley (*hordeum vulgare*) chloroplast development (evidence for selective stabilization of *psba* mRNA). Plant Physiol. **101:** 781-791.

Biehl, A., Richly, E., Noutsos, C., Salamini, F., and 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.

Birnbaum, K., Shasha, D.E., Wang, J.Y., Jung, J.W., Lambert, G.M., Galbraith, D.W., and Benfey, P.N. (2003). A Gene expression map of the *Arabidopsis* Root. Science **302:** 1956-1960.

Bisaz, C., Begot, L., Carol, P., Perez, P., Bligny, M., Pesey, H., Gallois, J.L., Lerbs-Mache, S., and Mache, R. (2003). The Arabidopsis nuclear DAL gene encodes a chloroplast protein which is required for the maturation of the plastid ribosomal RNAs and is essential for chloroplast differentiation. Plant Mol. Biol. **51:** 651-663.

Bock, R. (2006). Plastid biotechnology: prospects for herbicide and insect resistance, metabolic engineering and molecular farming. Curr. Opin. Biotechnol. **17:** 1-7.

Bollenbach, T.J., Schuster, G., and Stern, D.B. (2004). Cooperation of endo- and exoribonucleases in chloroplast mRNA turnover. Prog. Nucleic Acid Res. Mol. Biol. **78:** 305-337.

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

Canton, F.R., and Quail, P.H. (1999). Both phyA and phyB mediate light-imposed repression of *PHYA* gene expression in *Arabidopsis*. Plant Physiol. **121:** 1207-1216.

Casal, J.J. (1996). Phytochrome A enhances the promotion of hypocotyl growth caused by reductions in levels of phytochrome B in its far-red-light-absorbing form in light-grown *Arabidopsis thaliana*. Plant Physiol. **112:** 965-973.

Chory, J., Peto, C., Feinbaum, R., Pratt, L., and Ausubel, F. (1989). *Arabidopsis thaliana* mutant that develops as a light-grown plant in the absence of light. Cell **58:** 991-999.

Cleveland, W.S., and Devlin, S.J. (1988). Locally weighted regression: an approach to regression analysis by local fitting. Journal of the American Statistical Association **83:** 596-610.

Danon, A., and Mayfield, S.P. (1991). Light regulated translational activators: Identification of chloroplast gene specific mRNA binding proteins. EMBO J. **10:** 3993-4001.

Demarsy, E., Courtois, F., Azevedo, J., Buhot, L., Lerbs-Mache, S. (2006). Building up of the plastid transcriptional machinery during germination and early plant development. Plant Physiol. **142:** 993-1003.

Dhingra, A., Bies, D.H., Lehner, K.R., and Folta, K.M. (2006). Green light adjusts the plastid transcriptome during early photomorphogenic development. Plant Physiol. **142:** 1256-1266.

DuBell, A.N., and Mullet, J.E. (1995). Differential transcription of pea chloroplast genes during light-induced leaf development (Continuous far-red light activates chloroplast transcription). Plant Physiol. **109:** 105-112.

Duggan, D.J., Bittner, M., Chen, Y., Meltzer, P. and Trent, J.M. (1999) Expression profiling using cDNA microarrays. Nat. Genet. **21:** 10 -14.

Eisen, M.B., Spellman, P.T., Brown, P.O., and Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. Proc. Natl. Acad. Sci. USA **95:** 14863-14868.

Emanuel, C., Weihe, A., Graner, A., Hess, W.R., and Börner, T. (2004). Chloroplast development affects expression of phage-type RNA polymerase in barley leaves. Plant J. **38:** 460-472.

Erickson, B., Stern, D.B., and Higgs, D.C. (2005). Microarray analysis confirms the specificity of a *Chlamydomonas reinhardtii* chloroplast RNA stability mutant. Plant Physiol. **137:** 534-544.

Favory, J.J., Kobayshi, M., Tanaka, K., Peltier, G., Kreis, M., Valay, J.G., and Lerbs-Mache, S. (2005). Specific function of a plastid sigma factor for *ndhF* gene transcription. Nucleic Acids Res. **33:** 5991-5999.

Felder, S., Meierhoff, K., Sane, A.P., Meurer, J., Driemel, C., Plücken, H., Klaff, P., Stein, B., Bechtold, N., and Westhoff, P. (2001). The nucleus-encoded *HCF107* gene of *Arabidopsis* provides a link between intercistronic RNA processing and the accumulation of translation-competent *psbH* transcripts in chloroplasts. Plant Cell **13:** 2127-2141.

Gibson, G. (2003). Microarray Analysis. Genome-scale hypothesis scanning. PLos Biol. **1:** 27-28.

Glanz, S., Bunse, A., Wimbert, A., Balczun, C., and Kück, U. (2006). A nucleosome assembly protein-like polypeptide binds to chloroplast group II intron RNA in chlamydomonas reinhardtii. Nucleic Acid. Res. **34:** 5337-5351.

Gray, J.C., Sullivan, J.A., Wang, J.H., Jerome, C.A., and MacLean, D. (2002). Coordination of plastid and nuclear gene expression. Phil. Trans. R. Soc. Lond. B. **358:** 135-145.

Grebe, M. (2005). Growth by auxin: When a weed needs acid. Science **310:** 60-61.

Gruissem, W., and Tonkyn, J. (1993). Control mechanisms of plastid gene expression. CRC Critical Rev. Plant Biol. **12:** 19-55.

Harr, B., and Schlötterer C. (2006). Comparison of algorithms for the analysis of Affymetrix microarray data as evaluated by co-expression of genes in known operons. Nucleic Acids Res. **34:** e8.

Hashimoto, M., Endo, T., Peltier, G., Tasaka, M., and Shikanai, T. (2003). A nucleus-encoded factor, CRR2, is essential for the expression of chloroplast ndhB in Arabidopsis. Plant J. **36:** 541-549.

Hedtke, B., Legen, J., Weihe, A., Herrmann, R.G., and Börner, T. (2002). Six active phage-type RNA polymerase genes in *Nicotiana tabacum*. Plant J. **30:** 625-637.

Hennig, L., Poppe, C., Sweere, U., Martin, A., and Schafer, E. (2001). Negative interference of endogenous phytochrome B with phytochrome A function in *Arabidopsis*. Plant Physiol. **125:** 1036-1044.

Herrmann, R.G. (1997). Eukaryotism, towards a new intrepretation. In Schenk, H.E.A., Herrmann, R.G., Jeon, K.W., and Schwemmler, W. eds, Eukaryotism and Symbiosis. Springer, Heidelberg, Germany, pp 73–118.

Herrmann, R.G., and Westhoff, P. (2001). Thylakoid Biogenesis and Dynamics: The Result of a Complex Phylogenetic Puzzle. Springer, Heidelberg, Germany, pp 1-24.

Hihara, Y., Kamei, A., Kanehisa, M., Kaplan, A., and Ikeuchi, M. (2001). DNA Microarray analysis of cyanobacterial gene expression during acclimation to high light. Plant Cell **13:** 793-806.

Hirose, T., and Sugiura, M. (1997). Both RNA editing and RNA cleavage are required for translation of tobacco chloroplast *ndhD* mRNA: a possible regulatory mechanism for the expression of a chloroplast operon consisting of functionally unrelated genes. EMBO J. **16:** 6804-6811.

Hudson, M.E., and Quail, P.H. (2003). Identification of promoter motifs involved in the network of phytochrome A-regulated gene expression by combined analysis of genomic sequence and microarray data. Plant Physiol. **133:** 1605-1616.

Ichikawa, K., Sugita, M., Imaizumi, T., Wada, M., and Aoki, S. (2004). Differential expression on a daily basis of plastid sigma factor genes from the moss *Physcomitrella patens*. Regulatory interactions among *PpSig5*, the circadian clock, and blue light signaling mediated by cryptochromes. Plant Physiol. **136:** 4285-4298.

Ishizaki, Y., Tsunoyama, Y., Hatano, K., Ando, K., Kato, K., Shinmyo, A., Kobori, M., Takeda, G., Nakahira, Y., and Shiina, T. (2005). A nuclear-encoded sigma factor,

Arabidopsis SIG6, recognize sigma-70 type chloroplast promoters and regulates early chloroplast development in cotyledons. Plant J. **42:** 133-144.

Jenkins, B.D., Kulhanek, D.J., and Barkan, A. (1997). Nuclear mutations that block group II RNA splicing in maize chloroplasts reveal several intron classes with distinct requirements for splicing factors. Plant Cell **9:** 283-296.

Kanamaru, K., and Tanaka, K. (2004). Roles of chloroplast RNA polymerase sigma factors in chloroplast development and stress response in higher plants. Biosci. Biotechnol. Biochem. **68:** 2215-2223.

Kim, J., and Mayfield, S.P. (1997). Protein disulfide isomerase as a regulator of chloroplast translational activation. Science **278:** 1954-1957.

Kim, J., and Mayfield, S.P. (2002). The active site of the thioredoxin-like domain of chloroplast protein disulfide isomerase, RB60, catalyzes the redox-regulated binding of chloroplast poly(A)-binding protein, RB47, to the 59 untranslated region of *psbA* mRNA.

Kim, M., Christopher, D.A., and Mullet, J.E. (1993). Direct evidence for selective modulation of *psbA, rpoA, rbcL* and *16S* RNA stability during barley chloroplast development. Plant Mol. Biol. **22:** 447-463.

Kim, M., and Mullet, J.E. (1995). Identification of a sequence-specific DNA binding factor required for transcription of the barley chloroplast blue light-responsive *psbD-psbC* promoter. Plant Cell **7:** 1445-1457.

Koussevitzky, S., Nott, A., Mockler, T.C., Hong, F., Sachetto-Martins, G., Surpin, M., Lim, J., Mittler, R., and Chory, J. (2007). Signals from chloroplasts converge to regulate nuclear gene expression. Science **4:** 715-719.

Kuhn, K., Bohnem A,V., Liere, K., Weihe, A., and Börner, T. (2007). *Arabidopsis* phage-type RNA polymerases: accurate in vitro transcription of organellar genes. Plant Cell **19:** 959-971.

Legen, J., Kemp, S., Krause, K., Profanter, B., Herrmann, R.G., and Maier R.M. (2002). Comparative analysis of plastid transcription profiles of entire plastid chromosomes from tobacco attributed to wild-type and PEP-deficient transcription machineries. Plant J. **31:** 171-188.

Legen, J., Misera, S., Herrmann, R.G., and Meurer, J. (2001). Map positions of 69 *Arabidopsis thaliana* genes of all known nuclear encoded constituent polypeptides and various regulatory factors of the photosynthetic membrane: a case study. DNA Res. **8:** 53-60.

Lennartz, K., Bossmann, S., Westhoff, P., Bechtold, N., and Meierhoff, K. (2006). HCF153, a novel nuclear-encoded factor necessary during a. post-translational step in biogenesis of the cytochrome b_6f. complex. **45:** 101-112.

Lezhneva, L., and Meurer, J. (2004). The nuclear factor HCF145 affects chloroplast *psaA-psaB-rps14* transcript abundance in *Arabidopsis thaliana*. Plant J. **38:** 740-753.

Liere, K., and Börner, T. (2006). Transcription of plastid genes. In: K.D. Grasser (ed.) Regulation of Transcription in plants. Blackwell Publishing, Oxford, pp. 184-223.

Liere, K., and Maliga, P. (1999). In vitro characterization of the tobacco *rpoB* promoter reveals a core sequence motif conserved between phage-type plastid and plant mitochondrial promoters. EMBO J. **4:** 249-257.

Liere, K., Kaden, D., Maliga, P., and Börner, T. (2004). Overexpression of phage-type RNA polymerase *RpoTp* in tobacco demonstrates its role in chloroplast transcription by recognizing a distinct promoter type. Nucleic Acids Res. **32:** 1159-1165.

Liere, K., and Link, G. (1997). Chloroplast endoribounclease p54 invovled in RNA 3' end processing is regulated by phosphorylation and redox state. Nucleic Acids Res. **12:** 2403-2408.

Ma, L., Zhao, H., and Deng, X.W. (2003). Analysis of the mutational effects of the COP/DET/FUS loci on genome expression profiles reveals their overlapping yet not identical roles in regulating *Arabidopsis* seedling development. Development **130:** 969-981.

Mao, J., Zhang, Y.C., Sang, Y., Li, Q,H., and Yang, H.Q. (2005). A role for *Arabidopsis* crytochromes and COP1 in the regulation of stomatal opening. Proc. Natl. Acad. Sci. USA **102:** 12270-12275.

Mayfield, S.P., Cohen, A., Danon, A., and Yohn, C.B. (1994). Translation of the *psbA* mRNA of Chlamydomonas reinhardtii requires a structured RNA element contained within the 5' untranslated region. J. Cell Biol. **127:** 1537-1545.

Martin, W., and Herrmann, R.G. (1998). Gene transfer from organelles to the nucleus: how much, what happens, and Why? Plant Physiol. **118:** 9-17.

Meierhoff, K., Felder, S., Nakamura, T., Bechtold, N., and Schuster, G. (2003). HCF152, an *Arabidopsis* RNA binding pentatricopeptide repeat protein involved in the processing of chloroplast *psbB-psbT-psbH-petB-petD* RNAs. Plant Cell **15:** 1480-1495.

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

Meurer, J., Berger, A., and Westhoff, P. (1996a). A nuclear mutant of *Arabidopsis* with impaired stability on distinct transcripts of the plastid *psbB, psbD/C, ndhH*, and *ndhC* operons. Plant Cell **8:** 1193-1207.

Meurer, J., Grevelding, C., Westhoff, P., and Reiss, B. (1998). The PAC protein affects the maturation of specific chloroplast mRNAs in *Arabidopsis thaliana*. Mol. Gen. Genet. **258:** 342-351.

Meurer, J., Lezhneva, L., Amann, K., Gödel, M., Bezhani, S., Sherameti, I., and Oelmüller, R. (2002). A peptide chain release factor 2 affects the stability of UGA-containing transcripts in *Arabidopsis* chloroplasts. Plant Cell **14:** 3255-3269.

Meurer, J., Meierhoff, K., and Westhoff, P. (1996b). Isolation of high-chlorophyll-fluorescence mutants of *Arabidopsis thaliana* and their characterisation by spectroscopy, immunoblotting and northern hybridisation. Planta. **198:** 385-396.

Monde, R.A., Schuster, G., and Stern, D.B. (2000). Processing and degradation of chloroplast mRNA. Biochimie **82:** 573-582.

Murashige, T., and Skoog. F. (1962). A revised medium for rapid growth and bioassays with tobacco tissue culture. Physiology Plantarum **15:** 473-497.

Nakamura, T., Furuhashi, Y., Hasegawa, K., Hashimoto, H., Watanabe, K., Obokata, J., Sugita, M., and Sugiura, M. (2003). Array-based analysis on tobacco plastid transcripts: preparation of a genomic microarray containing all genes and all intergenic regions. Plant Cell Physiol. **44:** 861-867.

Nagashima, A., Hanaoka, M., Motohashi, R., Seki, M., Shinozaki, K., Kanamaru, K., Takahashi, H., and Tanaka, K. (2004). DNA microarray analysis of plastid gene expression in an *Arabidopsis* mutant deficient in a plastid transcription factor sigma, SIG2. Biosci. Biotechnol. Biochem. **68:** 694-704.

Ohgishi, M., Saji, K., Okada, K., and 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.

Okuda, K., Nakamura, T., Sugita, M., Shimizu, T., and Shikanai, T. (2006). A pentatricopeptide repeat protein is a site recognition factor in chloroplast RNA editing. J. Bio. Chem. **281:** 37661-37667.

Ossenbühl, F., Gohre, V., Meurer, J., Krieger-Liszkay, A., Rochaix, J.D., and Eichacker, L.A. (2004). Efficient assembly of photosystem II in *Chlamydomonas reinhardtii* requires Alb3.1p, a homolog of *Arabidopsis* ALBINO3. Plant Cell **16:** 1790-1800.

Ossenbühl, F., and Nickelsen, J. (2000). cis-and trans-acting determinants for translation of psbD mRNA in *Chlamydomonas reinhardtii*. Mol. Cell. Biol. **20:** 8134-8142.

Ostheimer, G.J., Williams-Carrier, R., Belcher, S., Osborne, E., Gierke, J., and Barkan, A. (2003). Group II intron splicing factors derived by diversification of an ancient RNA-binding domain. EMBO J. **22:** 3919-3899.

Perron, K., Goldschmidt-Clermont, M., and Rochaix, J.D. (2004). A multiprotein complex involved in chloroplast group II intron splicing. RNA **10:** 704-711.

Pfalz, J., Liere, K., Kandlbinder, A., Dietz, K.J., and Oelmüller, R. (2006). pTAC2,-6 and -12 are components of the transcriptionally active plastid chromosome that are required for plastid gene expression. Plant Cell **18:** 176-197.

Privat, I., Hakimi, M.A., Buhot, L., Favory, JJ., and Lerbs-Mache, S. (2003). Characterization of *Arabidopsis* plastid sigma-like transcription factors SIG1, SIG2 and SIG3. Plant Mol. Biol. **51:** 385-399.

Raab, S., Toth, Z., de Groot, C., Stamminger, T., and Hoth, S. (2006). ABA-reponsive RNA-binding proteins are involved in chloroplast and stromule function in Arabidopsis seedlings. Planta **224:** 900-914.

Race, H.L., Herrmann, R.G., and Martin, W. (1999). Why have organelles retained genomes? Trends Genet. **15:** 364-370.

Reinke, V., and White, K.P. (2002). Developmental genomic approaches in model organisms. Annu. Rev. Genomics Hum. Genet. **3:** 153-178.

Rensink, W.A., and Buell, C.R. (2005). Microarray expression profiling resources for plant genomics. Trends Plant Sci. **10:** 603-609.

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

Rochaix, J.D. (2001). Posttranscriptional control of chloroplast gene expression. From RNA to photosynthetic complex. Plant Physiol. **125:** 142-144.

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

Rymarquis, L., Higgs, D.C., and Stern, D.B. (2006). Nuclear suppressors define three factors that participate in both 5' and 3' end processing of mRNAs in chlamydomonas chloroplasts. Plant J. **46:** 448-461.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). Molecular cloning: a Laboratory Manual. 2nd.

Sato, S., Nakamura, Y., Kaneko, T., Asamizu, E., and Tabata, S. (1999). Complete structure of the chloroplast genome of *Arabidopsis thaliana*. DNA Res. **6:** 283-290.

Schena, M., Shalon, D., Davis, R.W., and Brown, P.O. (1995). Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science **270:** 467-470.

Schmid, M., Davison, T.S., Henz, S.R., Pape, U.J., Demar, M., Vingron, M., Schölkopf, B., Weigel, D., and Lohmann, J.U. (2005). A gene expression map of *Arabidopsis thaliana* development. Nat. Genet. **37:** 501-506.

Schmitz-Linneweber, C., Williams-Carrier, R.E., Williams-Voelker, P.M., Kroeger, T.S., Vichas, A., and Barkan, A. (2006). A pentatricopeptide repeat protein facilitates the trans-splicing of the maize chloroplast *rps12* pre-mRNA. Plant Cell **18:** 2650-2663.

Schreiber, U. (1986). Detection of rapid induction kinetics with a new type of high-frequency modulated chlorophyll fluorometer. Photosynth. Res. **9:** 261-272.

Schult, K., Meierhoff, K., Paradies, S., Toller, T., Wolff, P., and Westhoff, P. (2007). The Nuclear-Encoded Factor HCF173 Is Involved in the Initiation of Translation of the *psbA* mRNA in *Arabidopsis thaliana*. Plant Cell **19:** 1329-1346.

Sexton, T.B., Christopher, D.A., and Mullet, J.E. (1990). Light-induced switch in barley *psbD-psbC* promoter utilization: a novel mechanism regulating chloroplast gene expression. EMBO J. **9:** 4485-94.

Shedden, K., Chen, W., Kuick, R., Ghosh, D., Macdonald, J., Cho, K.R., Giordano, T.J., Gruber, S.B., Fearon, E.R., and Taylor, J.M. (2005). Comparison of seven methods for producing Affymetrix expression scores based on False Discovery Rates in disease profiling data. BMC Bioinformatics **6:** 26.

Shiina, T., Tsunoyama, Y., Nakahira, Y., and Khan, M.S. (2005). Plastid RNA polymerases, promoters, and transcription regulators in higher plants. Int. Rev. Cytol. **244:** 1-68.

Sturn, A., Quackenbush, J., and Trajanoski, Z. (2002). Genesis: Cluster analysis of microarray data. Bioinformatics **18:** 207-208.

Soukas, A., Cohen, P., Socci, N.D., and Friedman, J.M. (2000). Leptin-specific patterns of gene expression in white adipose tissue. Genes Dev. **15:** 963-980.

Tamayo, P., Slonim, D., Mesirov, J., Zhu, Q., Kitareewan, S., Dmitrovsky, E., Lander, E.S., and Golub, T.R. (1999). Interpreting patterns of gene expression with self-

organizing maps: Methods and application to hematopoietic differentiation. Proc. Natl. Acad. Sci. USA **96:** 2907-2912.

Tanaka, K., Oikawa, K., Ohta, N., Kuroiwa, H., Kuroiwa, T. and Takahashi, H. (1996). Nuclear encoding of a chloroplast RNA polymerase sigma subunit in a red alga. Science **272:** 1932-1935.

Tepperman, J.M., Hudson, M.E., Khanna, R., Zhou, T., Chang, S.H., Wang, X., and Quail, P.H. (2004). Expression profiling of *phyB* mutant demonstrates substantial contribution of other phytochromes to red-light-regulated gene expression during seedling de-etiolation. Plant J. **38:** 725-739.

Tepperman, J.M., Hudson, M.E., Khanna, R., Zhou, T., Chang, S.H., Wang, X., and Quail, P.H. (2004). Expression profiling of *phyB* mutant demonstrates substantial contribution of other phytochromes to red-light-regulated gene expression during seedling de-etiolation. Plant J. **38:** 725-739.

Tepperman, J.M., Zhu, T., Chang, H.S., Wang, X., and Quail, P.H. (2001). Multiple transcription-factor genes are early targets of phytochrome A signalling. Proc. Natl. Acad. Sci. USA **98:** 9437-9442.

Tepperman, J.M., Hwang, Y.S., and Quail, P.H. (2006). PhyA dominates in transduction of red-light signals to rapidly responding genes at the initiation of *Arabidopsis* seedling de-etiolation. Plant J. **48:** 728-742.

The Arabidopsis genome initiative. (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature **408:** 796-815.

O'Connor, T.R., Dyreson, C., and Wyrick, J.J. (2005). Athena: a resource for rapid visualization and systematic analysis of *Arabidopsis* promoter sequences. Bioinformatics **21:** 4411-4413.

Torres-Galea, P., Huang, L.F., Chua, N.H., and Bolle, C. (2006). The GRAS protein SCL13 is a positive regulator of phytochrome-dependent red light signaling, but can also modulate phytochrome A responses. Mol. Genet. Genomics **276:** 13-30.

Trebitsh, T., Levitan, A., Sofer, A., and Danon, A. (2000). Translation of chloroplast *psbA* mRNA is modulated in the light by counteracting oxidizing and reducing activities. Mol. Cell. Biol. **20:** 1116-1123.

Tsunoyama, Y., Ishizaki, Y., Morikawa, K., Kobori, M., Nakahira, Y., Takeba, G., Toyoshima, Y., and Shiina, T. (2004). Blue light-induced transcription of plastid-encoded *psbD* gene is mediated by a nuclear-encoded transcription initiation factor, AtSig5. Proc. Natl. Acad. Sci. USA. **2:** 3304-3309.

Tsunoyama, Y., Morikawa, K., Shiina, T., and Toyoshima, Y. (2002). Blue light specific and differential expression of a plastid sigma factor, *Sig5* in *Arabidopsis thaliana*. FEBS Lett. **516:** 225-228.

Unseld, M., Marienfeld, J.R., Brandt, P., and Brennicke, A. (1997). The mitochondrial genome of Arabidopsis thaliana contains 57 genes in 366,894 nucleotides. Nat. Genet. **15:** 57-61.

Weihe, A., and Börner, T. (1999). Transcription and the architecture of promoters in chloroplast. Trends Plant Sci. **4:** 169-170.

Wellmer, F., and Riechmann, J.L. (2005). Gene network analysis in plant development by genomic technologies. Int. J. Dev. Biol. **49:** 745-759.

Wellmer, F., Riechmann, J.L., Alves-Ferreira, M., and Meyerowitz, E.M. (2004). Genome-wide analysis of spatial gene expression in *Arabidopsis* flowers. Plant Cell **16:** 1314-1326.

Williams, P.M., and Barkan, A. (2003). A chloroplast-localized PPR protein required for plastid ribosome accumulation. **36:** 675-686.

Wilson, C.L., and Miller, C.J. (2005). Simpleaffy: a BioConductor package for Affymetrix Quality Control and data analysis. Bioinformatics **21:** 3683-3685.

Wu, H., Kerr, M.K., Cui, X.Q., and Churchill, G.A. (2003). MAANOVA: A Software Package for the Analysis of Spotted cDNA Microarray Experiments in the Analysis of Gene Expression Data: An Overview of Methods and Software, Parmigiani, G., Garret, E.S., Irizarry, R.A., Zeger, S.L., [eds], Springer, N.Y.

Yamagishi, A., and Fork, D.C. (1987). Photoreduction of Q_A, Q_B, and cytochrome b-559 in an oxygen-evolving photosystem II preparation from the thermophilic cyanobacterium *Synechococcus* sp. Arch. Biochem. Biophys. **15:** 124-130.

Yang, J., and Stern, D.B. (1997). The spinach chloroplast endoribonuclease CSP41 cleaves the 3'-untranslated region of *petD* mRNA primarily within its terminal stem-loop structure. J. Biol. Chem. **9:** 12874-12880.

Yohn, C.B., Cohen, A., Danon, A., and Mayfield, S.P. (1998a). A poly(A) binding protein functions in the chloroplast as a messagespecific translation factor. Proc. Natl. Acad. Sci. USA **95:** 2238–2243.

Yohn, C.B., Cohen, A., Danon, A., and Mayfield, S.P. (1996). Altered mRNA binding activity and decreased translational initiation in a nuclear mutant lacking translation of the chloroplast *psbA* mRNA. Mol. Cell Biol. **16:** 3560-3566.

Yohn, C.B., Cohen, A., Rosch, C., Kuchka, M.R., and Mayfield, S.P. (1998b). Translation of the chloroplast *psbA* mRNA requires the nuclear-encoded poly(A)-binding protein, RB47. J. Cell Biol. **142:** 435-442.

Zghidi, W., Merendino, L., Cottet, A., Mache, R., and Lerbs-Mache, S. (2007). Nucleus-encoded plastid sigma factor SIG3 transcribes specifically the *psbN* gene in plastids. Nucleic Acids Res. **35:** 455-464.

Zimmermann, P., Hirsch-Hoffmann, M., Hennig, L., and Gruissem, W. (2004). GENEVESTIGATOR. *Arabidopsis* Microarray Database and Analysis Toolbox. Plant Physiol. **136:** 2621-2632.

Zoschke, R., Liere, K., and Börner, T. (2007). From seedling to mature plant: *Arabidopsis* plastidial genome copy number, RNA accumulation and transcription are differentially regulated during leaf development. Plant J. **50:** 710-722.

ACKNOWLEDGMENTS

First of all, I would like to express my sincere gratitude to my supervisor PD Dr. Meurer for giving me the opportunity to perform my PhD work under pleasant working conditions and unlimited support for research. On backstage, it was always a great pleasure to enjoy his humor and to share loud laughing with him.

I am very grateful to Prof. Dr. Herrmann for the opportunity to work in his laboratory and for his support during my thesis. I also would like to thank to PD Dr. Bolle for her valuable discussions.

Arabidopsis seeds obtained from Prof. P. Westhoff, Prof. K. Apel, Prof. E. Neuhaus, and PD Dr. U. Vothknecht are greatly acknowledged.

I would like to thank my parents, aunt Mi Kyong and sister Jin Kyong for providing their love, care and support during my PhD. I also thank to my aunt Sun Young and uncle Rudi for introducing me into the German language and life style.

I would also like to thank my lab colleagues Lada, Agata, Serena, Rhea, Katrin, Uwe, Stephan, Jarda, Andy, and Simon for the nice atmosphere and lot of fun.

Particulary, I would like to thank to Dr. Cristina Dal Bosco, Dr. Lina Lezhneva, and Dr. Jeferson Gross for providing their exceptional advices on all scientific matters.

Especially, I really like to thank to my warm friend Dr. Pavan Umate called Guru for sharing, discussing and solving of all kinds of difficult and challenging problems in science and life. I never forget our first meeting since then he used to express his ever-valid wise saying "It's just the beginning". I would like to thank Pavan and Lee Gyan for the fun and enjoyable moments we spent together in Munich. Especially in China Garden!

I would like to thank Elli, Stefan, Yulia, and Uli for many humorous conversations.

My special thanks to Dr. Nuria Sanchez-Coll. It was a great pleasure to work and cooperate with her.

And finally, I would like to thank all people who helped me directly or indirectly during my PhD.

CURRICULUM VITAE

Personal data

Family Name : Cho
Name : Won Kyong
Birthday : June 1st, 1978
Place of Birth : Seoul, the Republic of Korea
Email : wonkyong@gmail.com

Education

10.2003 – 02.2007	Post graduate student at Department Biologie I, Botanik, LMU, München, Germany Supervisor: PD Dr. Jörg Meurer	
10.2002 – 08.2003	Exchange Student in Biology in LMU, Munich, Germany	
09.2000 – 08.2002	Master of Science in Genetic Engineering Sungkyunkwan University, Seoul, Korea Thesis : Evaluation of genetic characteristics and RAPD variations in Korean landraces of naked barley	
03.1997 – 08.2000	Bachelor of Science in Architecture Bachelor of Science in Landscape Architecture Sungkyunkwan University, Seoul, Korea	
03.1994 – 02.1997	High School (3 years)	Hoengseong, Kangwon-do, Korea
03.1991 – 02.1994	Middle School (3 years)	Hoengseong, Kangwon-do, Korea
03.1985 – 02.1991	Elementary School (6 years)	Hoengseong, Kangwon-do, Korea

PUBLICATIONS

1. Cho, W.K., Geimer, S., and Meurer, J. (2007). Cluster analysis and comparison of the chloroplast and nuclear transcriptomes in *Arabidopsis thaliana*. (submitted)

2. Sánchez, C. N., Danon, A., Meurer, J., Cho, W. K., and Apel, K. (2007) The advantage of being under stress: characterization of a novel suppressor of singlet oxygen-induced cell death in *Arabidopsis* seedlings. (in revision)

3. Aseeva, E., Ossenbühl, F., Sippel, C., Cho, W.K., Stein, B., Eichacker, L.A., Meurer, J., Wanner, G., Westhoff, P., Soll, J., and Vothknecht, U.C. (2007). Vipp1 function in thylakoid membrane biogenesis. Plant Physiol. Bioch. **45:** 119-128.

4. Mráček, J., Greiner, S., Cho, W.K., Rauwolf, U., Braun, M., Umate, P., Altstätter, J., Stoppel, R., Mlčochová, L., Silber, M.V., Volz, S.M., White, S., Selmeier, R., Rudd, S., Herrmann, R.G., and Meurer, J. (2006). Construction, database integration and applications of an *Oenothera* EST library. Genomics **88:** 372-380. (shared first authorship)

5. Gross, J., Cho, W.K., Lezhneva, L., Falk, J., Krupinska, K., Shinozaki, K., Seki, M., Herrmann, R.G., and Meurer, J. (2006). A single plant locus essential for phylloquinone (vitamin K_1) biosynthesis encodes a fusion of four genes of eubacterial origin. J. Biol. Chem. **281:** 17189-96.

6. Cho, W.K., Lee, J.M., Kwon, M.S., and Jeong, T.Y. (2002). Evaluation of morphological characteristics and RAPD analysis in Korean landraces of naked barley. Korean J. Plant Biotechnology **29:** 217-222.

Oral Presentation

Cho, W.K. (2004, 07). Phylogenie der Plastidären mRNA-Prozessierung. SFB TR 1 meeting, Freising, Germany

Poster Presentation

1. Cho, W.K., Pichler, B., Gerick, E., Lezhneva, L., Gross, J., Westhoff, P., and Meurer, J. (03, 2005). Phylogenetic development of plastid gene expression. FEBS meeting, Wildbadkreuth, Germany.

2. Cho, W.K., and Meurer, J. (04, 2006). Expression map of plastid genes in *Arabidopsis thaliana*. SFB TR1 meeting, Munich, Germany.

3. Kazmierczak, A., Stoppel, R., Schwenkert, S., Lezhneva, L., Cho, W.K., and Meurer, J. (04, 2006). Phylogenetic integration of the plastid transcript metabolism into the compartmentalized genome of the eukaryotic cell. SFB TR1 meeting, Munich, Germany.

4. Lezhneva, L., Cho, W.K., and Meurer, J. (08, 2006). Nuclear-encoded factors involved in regulation of chloroplast RNA metabolism in *Arabidopsis*. GRC on Mitochondria & Chloroplasts, Oxford, UK.

5. Gross, J., Cho, W.K., Lezhneva, L., Falk, J., and Meurer, J. (09, 2006). Phylloquinone (Vitamin K) synthesis in higher plants. 3rd-Tri-National *Arabidpsis* meeting Tübingen, Germany.

6. Stoppel, R., Cho, W.K., Lezhneva, L., Schwenkert, S., and Meurer, J. (03, 2006). Evolution and regulation of the higher plant chloroplast transcriptome. FEBS advanced lecture course. Origin and evolution of mitochondria and chloroplasts. Acquafredda di Maratea, Italy.

Eidesstattliche Erklärung

Hiermit versichere ich, dass ich die vorliegende Arbeit selbständig und nur unter Verwendung der angegebenen Hilfsmittel und Quellen angefertigt habe.

München, den 04. Juni 2007

Won Kyong Cho