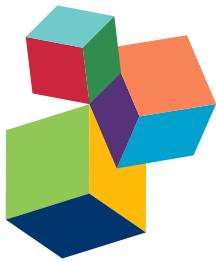


# THE PROTEINS OF PLASTID NUCLEOIDS – STRUCTURE, FUNCTION AND REGULATION

EDITED BY: Thomas Pfannschmidt and Jeannette Pfalz

PUBLISHED IN: Frontiers in Plant Science



# frontiers

## **Frontiers Copyright Statement**

© Copyright 2007-2016 Frontiers Media SA. All rights reserved.

All content included on this site, such as text, graphics, logos, button icons, images, video/audio clips, downloads, data compilations and software, is the property of or is licensed to Frontiers Media SA ("Frontiers") or its licensees and/or subcontractors. The copyright in the text of individual articles is the property of their respective authors, subject to a license granted to Frontiers.

The compilation of articles constituting this e-book, wherever published, as well as the compilation of all other content on this site, is the exclusive property of Frontiers. For the conditions for downloading and copying of e-books from Frontiers' website, please see the Terms for Website Use. If purchasing Frontiers e-books from other websites or sources, the conditions of the website concerned apply.

Images and graphics not forming part of user-contributed materials may not be downloaded or copied without permission.

Individual articles may be downloaded and reproduced in accordance with the principles of the CC-BY licence subject to any copyright or other notices. They may not be re-sold as an e-book.

As author or other contributor you grant a CC-BY licence to others to reproduce your articles, including any graphics and third-party materials supplied by you, in accordance with the Conditions for Website Use and subject to any copyright notices which you include in connection with your articles and materials.

All copyright, and all rights therein, are protected by national and international copyright laws.

The above represents a summary only. For the full conditions see the Conditions for Authors and the Conditions for Website Use.

**ISSN 1664-8714**

**ISBN 978-2-88919-927-3**

**DOI 10.3389/978-2-88919-927-3**

## **About Frontiers**

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

## **Frontiers Journal Series**

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

## **Dedication to quality**

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.

Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.

By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

## **What are Frontiers Research Topics?**

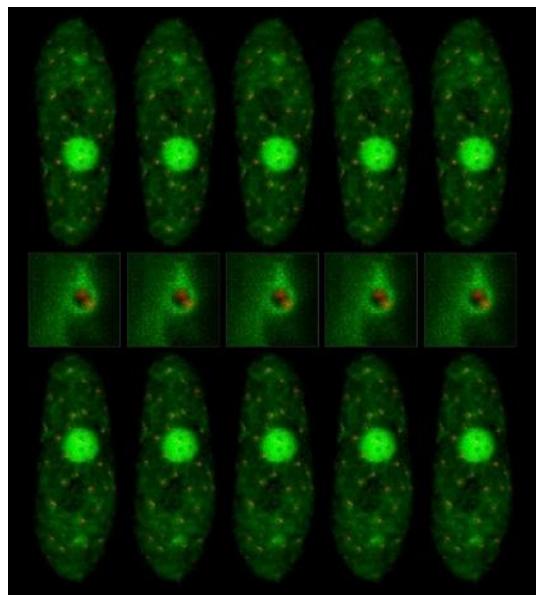
Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: [researchtopics@frontiersin.org](mailto:researchtopics@frontiersin.org)

# THE PROTEINS OF PLASTID NUCLEOIDS – STRUCTURE, FUNCTION AND REGULATION

Topic Editors:

**Thomas Pfannschmidt**, University Grenoble-Alpes, France

**Jeannette Pfalz**, Friedrich-Schiller-University Jena, Germany



Cover image artwork uses pictures taken by confocal imaging microscopy of an onion epidermal cell transformed by gold particle bombardment. Top and bottom row display an onion cell transiently expressing the green-fluorescent protein (GFP) as a cytosolic and nucleoplasmic marker. The bright-green circle in the center represents the nucleus. The cell was co-bombarded with a PAP10-DsRed construct marking the nucleoids in non-green plastids of the onion cell as red dots. The DsRed signal of a magnified plastid is shown in the middle row. The PAP10 protein is a subunit of the plastid RNA polymerase complex that localizes to two nucleoids of that plastid. The organelle itself is not visible but can be recognized as negative image within the surrounding GFP signal.

Photo credits: Monique Liebers and Robert Blanvillain

Plastids are plant cell-specific organelles of endosymbiotic origin that contain their own genome, the so-called plastome. Its proper expression is essential for faithful chloroplast biogenesis during seedling development and for the establishment of photosynthetic and other biosynthetic functions in the organelle. The structural organisation, replication and expression of this plastid genome, thus, has been studied for many years, but many essential steps are still not

understood. Especially, the structural and functional involvement of various regulatory proteins in these processes is still a matter of research. Studies from the last two decades demonstrated that a plethora of proteins act as specific regulators during replication, transcription, post-transcription, translation and post-translation accommodating a proper inheritance and expression of the plastome. Their number exceeds by far the number of the genes encoded by the plastome suggesting that a strong evolutionary pressure is maintaining the plastome in its present stage. The plastome gene organisation in vascular plants was found to be highly conserved, while algae exhibit a certain flexibility in gene number and organisation. These regulatory proteins are, therefore, an important determinant for the high degree of conservation in plant plastomes. A deeper understanding of individual roles and functions of such proteins would improve largely our understanding of plastid biogenesis and function, a knowledge that will be essential in the development of more efficient and productive plants for agriculture. The latter represents a major socio-economic need of fast growing mankind that asks for increased supply of food, fibres and biofuels in the coming decades despite the threats exerted by global change and fast spreading urbanisation.

**Citation:** Pfannschmidt, T., Pfalz, J., eds. (2016). *The Proteins of Plastid Nucleoids – Structure, Function and Regulation*. Lausanne: Frontiers Media. doi: 10.3389/978-2-88919-927-3

# Table of Contents

- 05 *Plastid nucleoids: evolutionary reconstruction of a DNA/protein structure with prokaryotic ancestry***  
Jeannette Pfalz and Thomas Pfannschmidt
- Section 1: Organisation and replication of nucleoids**
- 08 *Dynamic composition, shaping and organization of plastid nucleoids***  
Marta Powikowska, Svenja Oetke, Poul E. Jensen and Karin Krupinska
- 21 *WHIRLY1 is a major organizer of chloroplast nucleoids***  
Karin Krupinska, Svenja Oetke, Christine Desel, Maria Mulisch, Anke Schäfer, Julien Hollmann, Jochen Kumlehn and Götz Hensel
- 32 *Enzymes involved in organellar DNA replication in photosynthetic eukaryotes***  
Takashi Moriyama and Naoki Sato
- Section 2: Evolution and function of regulatory proteins in transcription**
- 44 *Nuclear-encoded factors associated with the chloroplast transcription machinery of higher plants***  
Qing-Bo Yu, Chao Huang and Zhong-Nan Yang
- 54 *Recent advances in the study of chloroplast gene expression and its evolution***  
Yusuke Yagi and Takashi Shiina
- 61 *AtSIG6 and other members of the sigma gene family jointly but differentially determine plastid target gene expression in *Arabidopsis thaliana****  
Sylvia Bock, Jennifer Ortelt and Gerhard Link
- 72 *Plastid encoded RNA polymerase activity and expression of photosynthesis genes required for embryo and seed development in *Arabidopsis****  
Dmitry Kremnev and Åsa Strand
- 84 *Establishment of the chloroplast genetic system in rice during early leaf development and at low temperatures***  
Kensuke Kusumi and Koh Iba
- Section 3: Proteins involved in post-transcriptional processes**
- 90 *The nucleoid as a site of rRNA processing and ribosome assembly***  
Alexandra-Viola Bohne
- 95 *Complex(iti)es of the ubiquitous RNA-binding CSP41 proteins***  
Dario Leister
- 99 *A purification strategy for analysis of the DNA/RNA-associated sub-proteome from chloroplasts of mustard cotyledons***  
Yvonne Schröter, Sebastian Steiner, Wolfram Weisheit, Maria Mittag and Thomas Pfannschmidt

# Plastid nucleoids: evolutionary reconstruction of a DNA/protein structure with prokaryotic ancestry

Jeannette Pfalz<sup>1</sup> and Thomas Pfannschmidt<sup>2,3,4,5\*</sup>

<sup>1</sup> Department of Plant Physiology, Institute of General Botany and Plant Physiology, Friedrich-Schiller-University Jena, Jena, Germany, <sup>2</sup> UMR5168, University Grenoble-Alpes, Grenoble, France, <sup>3</sup> Centre National de la Recherche Scientifique, UMR5168, Grenoble, France, <sup>4</sup> Commissariat à l'Energie Atomique et aux Energies Alternatives, iRTSV, Laboratoire de Physiologie Cellulaire and Végétale, Grenoble, France, <sup>5</sup> Institut National de la Recherche Agronomique, USC1359, Grenoble, France

**Keywords:** plastids, nucleoids, endosymbiosis, replication, transcription, post-transcriptional events

Understanding the evolutionary establishment of plastids within eukaryotic cells and the principles that govern the process of endosymbiosis have been integral to research in plant sciences during the past three decades. Determination of the primary DNA sequence of the plastome from many plants and algae represented a milestone in this field, making it possible to deduce evolutionary lineages via bioinformatic approaches. These have greatly improved our understanding of endosymbiosis, the evolution of plastids and the reshaping of the eukaryotic host genome following massive horizontal gene transfer from the ancient cyanobacterial progenitor toward the host nucleus. Astonishingly, much less is known about the current structure and organization of plastid DNA and its association with different kinds of proteins that are involved in its stabilization, replication and expression.

## OPEN ACCESS

### Edited and reviewed by:

Steven Carl Huber,  
United States Department of  
Agriculture, USA

### \*Correspondence:

Thomas Pfannschmidt,  
[thomas.pfannschmidt@ujf-grenoble.fr](mailto:thomas.pfannschmidt@ujf-grenoble.fr)

### Specialty section:

This article was submitted to Plant  
Physiology, a section of the journal  
*Frontiers in Plant Science*

**Received:** 23 February 2015

**Accepted:** 20 March 2015

**Published:** 08 April 2015

### Citation:

Pfalz J and Pfannschmidt T (2015)

Plastid nucleoids: evolutionary  
reconstruction of a DNA/protein  
structure with prokaryotic ancestry.

*Front. Plant Sci.* 6:220.

doi: 10.3389/fpls.2015.00220

As in bacteria, the DNA in plant and algal plastids appears to be organized in nucleoids that can be easily visualized by fluorescence microscopy using DNA-specific dyes. This approach identifies nucleoids as dots of distinctive shape that are located close to the thylakoid or envelope membrane depending on the developmental stage of the plastid. However, at the molecular level nucleoids represent a less well defined structure as they have been found to be a highly dynamic protein/DNA/RNA structure. In particular, its protein subunit composition is highly variable depending on the developmental stage of the plastid and the tissue context in which it resides, as well as on the environmental condition of the organism. In addition, the structure and organization of the DNA itself is still under debate. A definition of what precisely is a nucleoid in terms of protein sub-unit composition and structure, therefore, appears to be difficult on the basis of current knowledge. This research topic gives a snapshot of the current state-of-the-art on nucleoids focussing on their structure and composition. It zooms through the different levels of proteins involved in processes that are prerequisite for proper nucleoid structure and faithful gene expression.

The primary topic of the articles in this research topic is the various proteins found in nucleoids or likely associated with them based on their functional contribution to gene expression. Current knowledge and open questions about the organization of nucleoids are summarized in an initial review by Powikrowska et al. (2014). This article discusses the various appearances of nucleoids in different microscopy techniques, focussing heavily on the structural organization of DNA and the proteins that mediate it. It summarizes the characteristics of known plastid nucleoid associated proteins (ptNAPs) proposed to be involved in shaping and organization of nucleoids in plants. It also compares nucleoid morphology and organization in bacteria with that found in plants and extensively discusses the dynamics of nucleoid re-organization during the different phases of chloroplast development. This review is complemented by a research article that analyses the role of the protein Whirly1 in barley (Krupinska et al., 2014). Down-regulation of Whirly1 via

RNAi results in the occurrence of larger and more irregularly formed patches of DNA than are normally found in nucleoids. The data suggest an important role for Whirly1 in compacting nucleoid DNA and thereby affecting DNA replication.

These two articles set the scene for a detailed review about the enzymes involved in organellar replication contributed by Moriyama and Sato (2014), who describe the history of studies on organellar DNA polymerases and their enzymatic characteristics, including sensitivity to inhibitors or exonuclease activity. The article furthermore highlights other enzymes involved in replication such as helicases, DNA primase and topoisomerase as well as single-stranded DNA binding proteins. The review also covers the evolution of all these enzymes and their phylogenetic origins and relationships, and ends with an interesting model for the exchange of organellar replication enzymes during the evolution of photosynthetic eukaryotes.

The first level of gene expression is the transcription of the genetic information encoded by DNA. In chloroplasts, RNA is synthesized by two different types of RNA polymerases, the plastid-encoded RNA polymerase (PEP) and nuclear-encoded RNA polymerase (NEP). The PEP enzyme constitutes a genetically chimeric multi-protein complex with plastid-encoded core subunits structurally related to the bacterial *E. coli* RNA polymerase. One new feature of the PEP in higher plants, however, is its assembly with numerous nucleus-encoded eukaryotic components (PEP-associated proteins), which are reviewed in two articles (Yu et al., 2014; Yagi and Shiina, 2014). During the past decade, several approaches have established an im-portant role for such PEP-associated proteins (PAPs) in a variety of biological processes. These include transcriptional regulation, DNA/RNA metabolism, posttranslational modification and detoxification. More recently, it has been proposed that these proteins serve also as building blocks in the PEP assembly, but how exactly these proteins contribute to transcription and gene regulation awaits further investigation.

One important characteristic of plastid gene expression is the observation that PEP activity changes both in a developmentally regulated fashion and in response to environmental variables. Key proteins that mediate these changes in transcription are the different members of the sigma family (e.g., six in *Arabidopsis*) which initiate transcription in a complementary and flexible manner. Their concerted action allow greater flexibility

in developmental- and tissue-specific cellular responses (Bock et al., 2014). Other proteins that appear to influence developmental changes of plastid transcription are PRIN2 in *Arabidopsis* (Kremnev and Strand, 2014) or NUS1 in rice (Kusumi and Iba, 2014). PRIN2 was found to generate complexes with another protein called CSP41b (see also below). This complex appears to possess DNA binding activity *in vitro*, suggesting a regulatory role in plastid gene expression (Kremnev and Strand, 2014). NUS1 appears to be a regulator of plastid 16S rRNA expression that is responsible for the establishment of the plastid gene expression machinery in early stages of chloroplast development of rice exposed to low-temperature conditions. It works in conjunction with regulators of organellar and cytosolic nucleotide metabolism, indicating that nucleotide metabolism is essential for chloroplast development (Kusumi and Iba, 2014).

Post-transcriptional regulation is a further important level of control in plastids, and is high-lighted by two opinion articles in this issue (Bohne, 2014; Leister, 2014). The first discusses the roles of rRNA processing and maturation in nucleoids (Bohne, 2014). Based on experimental observations in bacteria, plastids and mitochondria, a new model was developed in which, in organelles, rRNA processing and ribosome assembly most likely take place in nucleoids (Bohne, 2014). The second article focusses on the roles of the CSP41 proteins (e.g., CSP41a and CSP41b) (Leister, 2014). These are multifunctional proteins of high abundance which have been found in several stromal protein complexes in different contexts, including RNA cleavage, RNA stabilization, transcription and carbon metabolism. Considering the abundance, CSP41 may have a key role in RNA stabilization.

The issue closes with a research article which describes an effective biochemical purification strategy that helps to isolate many of the aforementioned proteins from chloroplast nucleoids (Schröter et al., 2014). This strategy might be helpful in future in order to study native properties of nucleoid proteins isolated from plants in different developmental or environmental conditions. In summary, this research topic covers the full breadth of structural and functional implications of plastid nucleoids as currently known. It provides a comprehensive overview to the interested newcomer to the field and demonstrates open questions and topics which promise fundamental new discoveries in the years to come.

## References

- Bock, S., Ortelt, J., and Link, G. (2014). AtSIG6 and other members of the sigma gene family jointly but differentially determine plastid target gene expression in *Arabidopsis thaliana*. *Front. Plant Sci.* 5:667. doi: 10.3389/fpls.2014.00667
- Bohne, A. V. (2014). The nucleoid as a site of rRNA processing and ribosome assembly. *Front. Plant Sci.* 5:257. doi: 10.3389/fpls.2014.00257
- Kremnev, D., and Strand, A. (2014). Plastid encoded RNA polymerase activity and expression of photosynthesis genes required for embryo and seed development in *Arabidopsis*. *Front. Plant Sci.* 5:385. doi: 10.3389/fpls.2014.00385
- Krupinska, K., Oetke, S., Desel, C., Mulisch, M., Schäfer, A., Hollmann, J., et al. (2014). WHIRLY1 is a major organizer of chloroplast nucleoids. *Front. Plant Sci.* 5:432. doi: 10.3389/fpls.2014.00432
- Kusumi, K., and Iba, K. (2014). Establishment of the chloroplast genetic system in rice during early leaf development and at low temperatures. *Front. Plant Sci.* 5:386. doi: 10.3389/fpls.2014.00386
- Leister, D. (2014). Complex(iti)es of the ubiquitous RNA-binding CSP41 proteins. *Front. Plant Sci.* 5:255. doi: 10.3389/fpls.2014.00255
- Moriyama, T., and Sato, N. (2014). Enzymes involved in organellar DNA replication in photosynthetic eukaryotes. *Front. Plant Sci.* 5:480. doi: 10.3389/fpls.2014.00480
- Powikowska, M., Oetke, S., Jensen, P. E., and Krupinska, K. (2014). Dynamic composition, shaping and organization of plastid nucleoids. *Front. Plant Sci.* 5:424. doi: 10.3389/fpls.2014.00424

- Schröter, Y., Steiner, S., Weisheit, W., Mittag, M., and Pfannschmidt, T. (2014). A purification strategy for analysis of the DNA/RNA-associated sub-proteome from chloroplasts of mustard cotyledons. *Front. Plant Sci.* 5:557. doi: 10.3389/fpls.2014.00557
- Yagi, Y., and Shiina, T. (2014). Recent advances in the study of chloroplast gene expression and its evolution. *Front. Plant Sci.* 5:61. doi: 10.3389/fpls.2014.00061
- Yu, Q. B., Huang, C., and Yang, Z. N. (2014). Nuclear-encoded factors associated with the chloroplast transcription machinery of higher plants. *Front. Plant Sci.* 5:316. doi: 10.3389/fpls.2014.00316

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Pfalz and Pfannschmidt. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Dynamic composition, shaping and organization of plastid nucleoids

**Marta Powikrowska<sup>1</sup>, Svenja Oetke<sup>2</sup>, Poul E. Jensen<sup>1</sup> and Karin Krupinska<sup>2\*</sup>**

<sup>1</sup> Department of Plant and Environmental Sciences, VILLUM Research Centre for Plant Plasticity and Copenhagen Plant Science Centre, University of Copenhagen, Copenhagen, Denmark

<sup>2</sup> Plant Cell Biology, Institute of Botany, Christian-Albrechts-University of Kiel, Kiel, Germany

**Edited by:**

Jeannette Pfalz,  
Friedrich-Schiller-University Jena,  
Germany

**Reviewed by:**

Alice Barkan, University of Oregon,  
USA  
Wataru Sakamoto, Okayama  
University, Japan

**\*Correspondence:**

Karin Krupinska, Plant Cell Biology,  
Institute of Botany,  
Christian-Albrechts-University of  
Kiel, Olshausenstrasse 40,  
24098 Kiel, Germany  
e-mail: kkrupinska@bot.uni-kiel.de

In this article recent progress on the elucidation of the dynamic composition and structure of plastid nucleoids is reviewed from a structural perspective. Plastid nucleoids are compact structures of multiple copies of different forms of ptDNA, RNA, enzymes for replication and gene expression as well as DNA binding proteins. Although early electron microscopy suggested that plastid DNA is almost free of proteins, it is now well established that the DNA in nucleoids similarly as in the nuclear chromatin is associated with basic proteins playing key roles in organization of the DNA architecture and in regulation of DNA associated enzymatic activities involved in transcription, replication, and recombination. This group of DNA binding proteins has been named plastid nucleoid associated proteins (ptNAPs). Plastid nucleoids are unique with respect to their variable number, genome copy content and dynamic distribution within different types of plastids. The mechanisms underlying the shaping and reorganization of plastid nucleoids during chloroplast development and in response to environmental conditions involve posttranslational modifications of ptNAPs, similarly to those changes known for histones in the eukaryotic chromatin, as well as changes in the repertoire of ptNAPs, as known for nucleoids of bacteria. Attachment of plastid nucleoids to membranes is proposed to be important not only for regulation of DNA availability for replication and transcription, but also for the coordination of photosynthesis and plastid gene expression.

**Keywords:** chromatin, nucleoid, plastid DNA, ptNAP, thylakoids

## INTRODUCTION

Plastids are the characteristic organelles of photosynthetic eukaryotes. They are the sites of photosynthesis, and their biosynthetic pathways supply the plant cell with many essential compounds. Chloroplasts evolved from a cyanobacterial ancestor after a single endosymbiotic event, that was followed by an extensive reduction of the plastid genome size (Timmis et al., 2004; Bock and Timmis, 2008; Green, 2011). Among the genes still present in the 100–200 kbp plastid genomes are the ribosomal RNA genes, 27–31 genes encoding tRNAs, and a variable number of other genes, that in higher plants include about 85 encoding proteins of the photosynthetic apparatus (Green, 2011).

Within the chloroplast, multiple copies of the plastid DNA (ptDNA) together with RNA and proteins are organized in structures that are similar to bacterial nucleoids. The compact structure of DNA in such nucleoids has been compared with the chromatin in the nucleus of eukaryotic cells (Sakai et al., 2004). The fundamental difference between genome organization in plastids vs. that in bacteria is, that plastids have multiple nucleoids with a varying number of genome copies, whereas bacteria only have a single nucleoid containing a variable number of DNA molecules. Nucleoids contain all enzymes necessary for transcription, replication and segregation of the plastid genome (Sakai

et al., 2004). Moreover, posttranscriptional processes including RNA splicing and editing, as well as ribosome assembly, take place in association with the nucleoid, suggesting that these processes occur co-transcriptionally (Majeran et al., 2012). However, among the many proteins found in the nucleoid and identified by proteomic analyses (Phinney and Thelen, 2005; Majeran et al., 2012; Melonek et al., 2012) only a few have been functionally characterized so far. In Table 1, proteins, that were proposed to play roles in nucleoid architecture, and which in analogy to the architectural proteins of bacterial nucleoids have been named plastid nucleoid associated proteins (ptNAPs) (Krupinska et al., 2013), are listed.

The dynamic shaping of nuclear chromatin and bacterial nucleoids is known to have profound effects on gene expression. Whereas the mechanisms underlying chromatin remodeling in the nucleus of plant cells have been investigated intensively, research on the mechanisms underlying the dynamics of the structure and organization of plastid nucleoids is still in its infancy. This is in sharp contrast with the enormous importance of chloroplast metabolism for growth and productivity of plants. Expression of plastid genes needs to be continuously coordinated with the activity of the nuclear genome. Structural changes are likely to be involved in the crosstalk between plastid and nuclear genomes.

**Table 1 | Characteristics of plastid nucleoid associated proteins proposed to be involved in shaping and organization of nucleoids in plants.**

| Name  | MW<br>theoretical | pl               | Proposed function                                    | Occurrence/<br>Species characterized  | References   |
|---|-------------------|------------------|--|---|--|
| PEND, plastid envelope DNA binding protein                | 130 kDa           | 4.6#<br>(10.3*)# | anchoring of nucleoids to the envelope membrane      | dicots/ <i>Pisum sativum</i> , <i>Brassica napus</i>  | Sato et al., 1993, 1998; Sato and Ohta, 2001; Terasawa and Sato, 2005; Wycliffe et al., 2005                       |
| MFP1, MAR-binding filament-like protein                   | 90 kDa            | 8.5              | anchoring of nucleoids to thylakoids                 | angiosperms/ <i>Lycopersicum esculentum</i>   | Meier et al., 1996; Jeong et al., 2003, 2004   |
| TCP34, (tetratricopeptide-containing chloroplast protein) | 38 kDa            | 5.4#             | candidate nucleoid anchoring protein                 | higher plants/ <i>Spinacia oleracea</i>   | Weber et al., 2006   |
| SWIB-4, domain of SWI/SNF complex B                       | 12 kDa            | 10               | packaging of DNA                                     | angiosperms/ <i>Spinacia oleracea</i> , <i>Arabidopsis thaliana</i>                               | Melonek et al., 2012   |
| pTAC3**   | 68 kDa#           | 4.6#<br>(9.6)#   | candidate DNA packaging protein                      | land plants except gymnosperms/ <i>Zea mays</i>   | Majeran et al., 2012   |
| SiR (DCP68), sulfite reductase (DNA compacting protein)   | 68 kDa            | 9.1#             | bifunctional: DNA compaction and sulfur assimilation | cyanobacteria, algae and land plants/ <i>Glycine max</i> , <i>Pisum sativum</i> , <i>Zea mays</i> | Cannon et al., 1999; Sekine et al., 2002, 2007, 2009; Sato et al., 2003; Kang et al., 2010; Wiedemann et al., 2010 |
| YlmG  | 23 kDa            | 10.9             | nucleoid partitioning                                | cyanobacteria and plastid containing eucaryotes, <i>Arabidopsis thaliana</i>                      | Kabeya et al., 2010  |
| SVR4/-like (MRL7/-like), suppressor of variegation 4      | 28 kDa            | 5.2#             | putative chaperones for NAPs                         | mosses, clubferns and angiosperms, <i>Arabidopsis thaliana</i> , <i>Hordeum vulgare</i>           | Qiao et al., 2011; Yu et al., 2011; Yua et al., 2013; Powikowska et al., 2014                                      |
| pTAC16**  | 54 kDa            | 8.9#             | putative membrane-anchor                             | angiosperms/ <i>Arabidopsis thaliana</i>  | Ingelsson and Vener, 2012  |
| WHIRLY1, 3 (pTAC1, pTAC11)**                              | 24-26 kDa         | 9.3#             | condensation of DNA of a subgroup of nucleoids       | angiosperms/ <i>Arabidopsis thaliana</i> , <i>Zea mays</i> , <i>Hordeum vulgare</i>               | Pfalz et al., 2006; Prikryl et al., 2008; Maréchal and Brisson, 2010; Krupinska et al., 2014                       |

MW, molecular weight; pl, isoelectric point.

#pl of the proteins basic region.

\*#pl or protein molecular weight was determined with ExPASy ProtParam (<http://web.expasy.org/cgi-bin/protparam/protparam>).

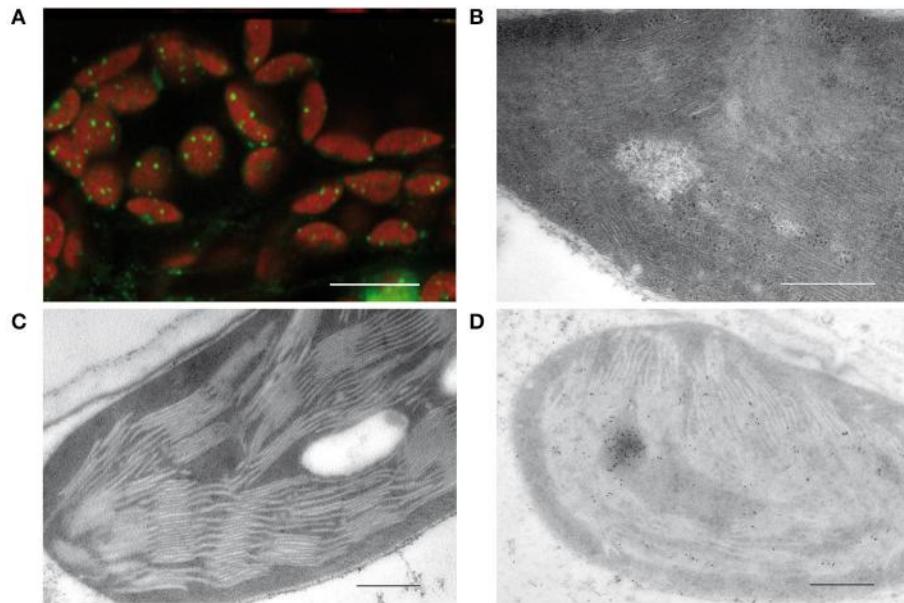
\*\*pTAC, protein detected in the transcriptional active chromosomes of chloroplasts from *Arabidopsis thaliana* (Pfalz et al., 2006).

In this article recent progress in the elucidation of the composition of plastid nucleoids is reviewed in the context of the complex DNA-protein architecture. The unique characteristics of plastid nucleoids will be highlighted by comparison with bacterial nucleoids and nuclear chromosomes. The involvement of plastid specific NAPs in regulation of DNA availability for replication and transcription and the functional significance of nucleoid association with the thylakoid membranes in chloroplasts will be discussed.

## MICROSCOPIC ANALYSES OF PLASTID NUCLEOID MORPHOLOGY

In 1962, Ris and Plaut discovered irregularly shaped bodies containing DNA in the chloroplast of Chlamydomonas by staining with acridine orange. Electron micrographs revealed microfibrils in areas of low density corresponding to DNA macromolecules

similar to those that were shown before in bacteria (Robinow and Kellenberger, 1994). These microfibrils suggested that at least part of the plastid DNA is “naked” in contrast to the nuclear DNA that together with basic proteins, histones, is organized in highly compact structures known as chromatin (Kuroiwa, 1991). Images obtained by staining with 4',6-diamidino-2-phenylindole (DAPI) or other DNA dyes such as SYBR Green revealed a quite different organization of plastid DNA. In chloroplasts, tiny compact structures associated with the thylakoids are detectable (Figure 1A). Protease treatment and reconstitution assays on such isolated structures indicated that the packaging degree of DNA is higher than in the metaphase chromosomes of animals (Nemoto et al., 1988; Kuroiwa, 1991). From these results it was concluded that ptDNA is not “naked,” but tightly packed in nucleoids by interactions with basic proteins as it is also known for the nuclear chromatin.



**FIGURE 1 | Visualization of plastid nucleoids by using different microscopic techniques. (A)** Nucleoids visualized by fluorescence microscopy of SYBR Green in leaf sections, bar: 10  $\mu\text{m}$ . **(B)** Conventional electron micrographs showing nucleoids with DNA filaments in mesophyll

chloroplasts. **(C)** Specimen prepared by high pressure freezing and freeze substitution (HPF-FS). **(D)** Immunogold labeling of nucleoids in leaf sections obtained from specimen prepared by high pressure freezing and freeze substitution (HPF-FS) using a DNA specific antibody, bar: 500 nm.

Indeed, the concept of “naked” DNA in plastids and bacteria was based only on conventional electron microscopy employing chemical fixation and dehydration of the tissue, known to lead to denaturing and loss of proteins. As a result, DNA filaments devoid of proteins get visualized in electron-lucent areas from which proteinous material was lost during dehydration (**Figure 1B**). When instead of chemical fixation, physical fixation by high pressure freezing and freeze substitution (HPF-FS) is employed, no DNA filaments are detectable (**Figure 1C**). Specimens prepared by HPF-FS were used for immunogold labeling with an antibody specific for single- and double-stranded DNA. Thereby regions of intensive labeling could be detected that have about the size of nucleoids as detected by epifluorescence or confocal microscopy (**Figure 1D**).

## DNA ORGANIZATION AND GENE EXPRESSION IN THE NUCLEUS

Genomic DNA in most eukaryotic cells is hierarchically organized within the chromatin (Campos and Reinberg, 2009; Fudenberg and Mirny, 2012). The basic unit of chromatin is the nucleosome that consists of double stranded DNA wrapped around a histone octamer. The nucleosomes organize into 11 nm fibers that resemble beads on strings. This structure is thought to further fold into so-called 30 nm fibers stabilized by the H1 linker histone. Although very little is known about the organization of chromatin beyond this stage, it is assumed that organization of the higher order chromatin structure involves formation of interacting fibers, chromatin loops and positioning to generate a distinctive spatial arrangement of the genome within the three-dimensional space of the nucleus (for a review see Li and Reinberg, 2011).

In general, the higher-order structures of nuclear chromatin inhibit DNA transaction processes, i.e., replication, repair, recombination and transcription of the DNA (Li and Reinberg, 2011). These DNA transaction processes require chromatin remodeling by mechanisms such as: (i) posttranslational modifications (acetylation and methylation) of N- and C-terminal tails of histones, (ii) exchanging histones variants, (iii) DNA methylation, (iv) non-histone architectural proteins, (v) ATP-dependent nucleosome remodelers, as well as (vi) the action of negatively charged histone chaperones.

Most eukaryotic genes are transcribed by RNA polymerase II (RNAP II). Interestingly, transcription by RNAP II requires dynamic changes in the chromatin structures of the templates (Orphanides and Reinberg, 2000; Studitsky, 2005). During high rates of transcription, nucleosomes are completely disassembled and reassembled with the assistance of ATP-dependent nucleosome remodelers and histone chaperones altering contacts between DNA and histones. These remodelers are specific for certain genes in different cell types and contexts of cell differentiation (de la Serna et al., 2006). ATP-dependent nucleosome remodelers allow the DNA to “inch-worm” around the histone octamer. Acidic histone chaperones, on the other hand, “collect” the basic histones after the histone-DNA interactions have been broken by the ATP-dependent nucleosome remodelers.

Non-histone architectural proteins, such as high mobility group (HMG) proteins (Grasser, 1995) also play a role in chromatin structural dynamics, since they decrease the compactness of the chromatin fiber and enhance the accessibility of DNA to regulatory factors. Members of the HMGN family contain a functional nucleosome-binding domain (NBD) and a negatively charged C-terminus of varying length. It has been shown that the

negatively charged C-terminal domain of HMGN5 interacts with the positively charged C-terminal domain of the linker histone H1 and thereby counteracts the H1-mediated compaction of a nucleosomal array. In turn, this facilitates transcriptional activation (Rochman et al., 2010).

Packaging of DNA by histones into nucleosomes is not a distinguishing feature of eukaryotes, but also occurs in some groups of archaeabacteria which might have participated in the origin of eukaryotes (Bendich and Drlica, 2000). In any case, a nucleosome based packaging of DNA results in a rather closed structure, and the access of DNA by DNA transaction enzymes involves several interconnected processes modeling the chromatin.

## DNA ORGANIZATION AND GENE EXPRESSION IN BACTERIA

Whereas the ability of histones to interfere with the nuclear chromatin structure and thereby to regulate transcription is rather well conserved among eukaryotes and understood in great detail, the situation in eubacteria seems to be more diverse and complicated. Research on the folding of bacterial DNA began in the 1970s, but the first systematic inventory of nucleoid associated proteins (NAP) (Azam and Ishihama, 1999) is still being extended (Dillon and Dorman, 2010). Many of these proteins are abundant basic proteins similar to histones and were found to influence chromatin structure and gene transcription. Accordingly, they were earlier named “histone like” proteins (Drlica and Rouviere-Yaniv, 1987; Dorman and Deighan, 2003). This group includes the highly conserved HU (heat unstable), the H-NS (histone-like nucleoid-structuring), IHF (integration host factor) and FIS (factor for inversion stimulation) (Dillon and Dorman, 2010). By the use of a bioinformatics approach it has been estimated that the bacterial nucleoid contains approximately one NAP per 100 bp (Li et al., 2009). According to their architectural mode of action toward DNA, three classes of architectural proteins are distinguished: wrappers, benders, or bridgers (Luijsterburg et al., 2008).

Importantly, there is no sequence or structural similarity between the prokaryotic histone-like proteins and eukaryotic histones (Macvanin and Adhya, 2012). The histone-like HU, H-NS, IHF and FIS proteins bind to AT-rich regions and shape the local structure of DNA upon binding (Browning et al., 2010). In contrast to histones that bind to both coding and non-coding DNA, the binding of these proteins occurs mostly in non-coding regulatory regions of the genome as shown by *in vivo* protein occupancy display (Grainger et al., 2006; Vora et al., 2009).

By electron microscopy, isolated nucleoids of *Escherichia coli* (*E. coli*) were shown to be organized as rosettes with a compact central core from which supercoiled DNA loops with an average size of 10 kbp were observed to radiate (Delius, 1974; Postow et al., 2004). The loops comprise topologically isolated domains with boundaries set by different NAPs such as H-NS and FIS, that can cross-link either different genomic loci or one locus with a membrane (Postow et al., 2004; Travers and Muskhelishvili, 2005; Luijsterburg et al., 2008). At a higher organizational level, the *E. coli* genome is folded into a structure containing four so-called macro-domains with specific NAPs and two less structured regions (Espeli et al., 2008). In *Caulobacter crescentus*, some domain specific NAPs are involved in control of replication and distribution of nucleoids (Dame et al.,

2011), while others were shown to regulate the position of chromosomes and the initiation of cytokinesis (Mohl et al., 2001).

NAPs have both structural and regulatory roles. They shape the overall organization of nucleoids depending on the external conditions and growth phase (Rimsky and Travers, 2011). The composition of the NAPs is known to change during the cell cycle, in response to growth phase and external conditions such as nutrient supply and stress factors. For example, FIS is a bending NAP with high levels in growing cells, but it is absent under conditions of slow growth and in cells of the stationary phase (Dillon and Dorman, 2010). In contrast, Dps (DNA protection from starvation), whose expression is regulated by FIS and other NAPs, accumulates at the end of the stationary phase mediating the formation of stable and highly ordered nucleoprotein complexes, also termed biocrystals, that are important for the protection of DNA during stress (Wolf et al., 1999).

In addition to their dynamic functions as structural proteins most NAPs serve dual or multiple purposes and also have specific functions (Dillon and Dorman, 2010; Dame et al., 2011). The HU protein was shown to form transcription foci that are spatially confined aggregations of RNA polymerases (Berger et al., 2010). Other NAPs such as CRP (cyclic AMP regulatory protein) act as transcription factors of specific genes (Nasser et al., 2001; Rimsky and Travers, 2011). The NAP repertoire has considerable impact on global gene expression and in many cases NAPs regulate gene expression by mutually antagonistic activities (Dillon and Dorman, 2010).

Taken together, in contrast to the eukaryotic chromatin, the composition of bacterial nucleoids is more diverse and dynamic. The composition of the NAP fraction is regulated mainly at the level of NAP gene expression whereby NAPs can regulate both the transcription of genes encoding other NAPs and/or their own genes (Travers and Muskhelishvili, 2005).

## DNA ORGANIZATION IN PLASTIDS

### THE PLASTID GENOME—SIZE, COPY NUMBER, AND TOPOLOGY

The size of the plastid genome of photosynthetically active algae and higher plants ranges from 120 to 190 kbp depending on the species (Wicke et al., 2011), e.g., in *Arabidopsis thaliana* it is 154 kbp (Sato et al., 1999). The percentage of coding sequence ranges from 50% in the green alga *Chlamydomonas reinhardtii* (Maul et al., 2002) to 93.5% in the red alga *Cyanidioschyzon merolae* (Misumi et al., 2005). Each plastid contains multiple copies of the genome which are distributed among a variable number of nucleoids. Despite the growing number of proteins shown to play roles in DNA replication and maintenance (Maréchal and Brisson, 2010), the mechanism of ptDNA replication is yet not well understood and might depend on the developmental stage of plastids (Nielsen et al., 2010). In fact, several mechanisms of DNA replication were proposed and one involves a chloroplast-targeted RecA protein (Rowan et al., 2010). Of particular importance for ptDNA levels is the activity of an organelle targeted DNA polymerase sharing homology with bacterial DNA polymerase I (Moriyama et al., 2011). In some maize mutants with mutations in the gene encoding the organelle targeted DNA polymerase ptDNA accumulation was observed to be approximately 100-fold reduced (Udy et al., 2012).

The number and positions of nucleoids were shown to depend on the developmental stage of the plastids (Boffey et al., 1979; Kuroiwa et al., 1981). In a recent study on *Beta vulgaris* 12–330 plastid chromosomes per organelle with about 4–7 copies per nucleoid were determined (Rauwolf et al., 2010). It had been suggested long ago that nucleoids even within one plastid contain varying amounts of DNA (Kowallik and Herrmann, 1972). The number of genome copies per plastid changes during chloroplast development (Boffey et al., 1979; Baumgartner et al., 1989), in *Arabidopsis* ranging from more than 100 in rapidly dividing cells to 20 or fewer in mature cells (Zoschke et al., 2007). Detailed information on plastid DNA copies per cell and per plastid in different plants and in different tissues and stages of development are presented in a recent review (Liere and Börner, 2013). There is controversial information on the DNA content of mature and senescing chloroplasts. Oldenburg and Bendich (2004) reported that mature chloroplasts do not contain DNA, being in contradiction with many other reports (Liere and Börner, 2013). In a recent article a reappraisal of this issue is presented using a combination of high resolution fluorescence microscopy, transmission electron microscopy and real-time quantitative PCR. Thereby the authors demonstrated that considerable levels of DNA and nucleoids are even detectable in plastids of ageing and senescent leaves in different species (Golczyk et al., 2014). The discrepancies between these studies and the former studies of Bendich and co-workers (Rowan et al., 2004) were proposed to be due to methodological insufficiencies of the experimental approaches. Indeed, it is rather unlikely that chloroplasts before entering the degradative phase of late senescence lack DNA, because the D1 protein of the photosynthetic apparatus is known to have a high turnover requiring a continuous re-synthesis (Melis, 1999). The high demand for new synthesis cannot be met by an extremely high stability of plastid mRNAs as claimed by Oldenburg et al. (2014) in their response to the article of Golczyk et al. (2014). In fact, plastid genes are actively transcribed in senescing barley leaves as shown by run-on assays (Krause et al., 1998; Krupinska and Humberg, 2004). When dark-induced senescence is reverted by light, in particular the transcriptional activities of photosynthesis associated plastid genes were shown to increase again (Krause et al., 1998).

The plastid genome can be divided into four major regions: (1) The large single copy region (LSC) which in *Arabidopsis* comprises as much as 54% of the genome, (2) the small single copy region (SSC) making up 12% of the plastid genome in *Arabidopsis*, and (3) the two inverted repeats, IR<sub>A</sub> and IR<sub>B</sub>, which contain the same genetic information in inverse orientation. Hence the genes contained in these repeats have two copies in the genome. In most plant species the repeats contain three or four ribosomal RNA genes and a number of other genes (Green, 2011). This domain based organization resembles the macrodomain organization of the bacterial genome. However, it is unknown whether, as in the case of bacteria, the different regions of the plastid genome comprise topological and functional units that are associated with specific NAPs as reported for the domains of the bacterial genome.

In contrast to the genomes of the eukaryotic nucleus and of bacteria, the organelle genomes are considered to be highly variable in structure (Bendich, 2004; Oldenburg and Bendich,

2004). Studies employing *in situ* hybridization showed that besides circular chromosomes, linear forms occur in plastids that were proposed to be the major forms in chloroplasts where many small nucleoids are attached to thylakoids (Bendich, 2004). Moreover, the majority of plastid DNA molecules are arranged in multimeric (concatemeric) structures (Deng et al., 1989; Oldenburg and Bendich, 2004; Maréchal and Brisson, 2010). So far, the mechanisms of concatemer formation, linkage and breakage of DNA in plastids are largely unknown (Wicke et al., 2011).

As in bacteria, the DNA in plastids is supercoiled, and plastid DNA topoisomerases play important roles in replication, repair and recombination of DNA (Day and Madesis, 2007). Changes in the DNA topology which especially happen during chloroplast development were proposed to have also dramatic consequences for gene expression (Lam and Chua, 1987; Zaitlin et al., 1989; Salvador et al., 1998).

## NUCLEOID ASSOCIATED PROTEINS IN PLASTIDS

Although several experiments have confirmed that the compact organization of plastid nucleoids is retained by electrostatic interactions between ptDNA and proteins, only a few structural proteins interacting with the ptDNA have been identified so far (Sakai et al., 2004; see Krupinska et al., 2013 for a detailed description of ptNAPs). Most of them have high isoelectric points in accordance with their DNA binding properties (**Table 1**). Homologs of bacterial HU proteins, namely HU-like proteins, which are known as basic non-specific DNA binding proteins, have been found instead of histones in the nucleus of most dinoflagellates (Sala-Rovira et al., 1991; Wong et al., 2003) and some algae (Bendich and Drlica, 2000). HU-like proteins (HLP) were found to be encoded by the chloroplast genomes of the primitive red alga *Cyanidioschyzon merolae* (Kobayashi et al., 2002) and the green algae *Chlamydomonas reinhardtii* (Karcher et al., 2009). These and nuclear encoded HU-like plastid proteins of algae were shown to be functional equivalents of the HU protein by complementation of bacterial mutants lacking HU (Kobayashi et al., 2002). However, in land plants, genes for HU-like proteins have neither been found in any of the sequenced plastid genomes nor in any of the sequenced nuclear genomes (Sato, 2001; Yagi and Shiina, 2014). Novel DNA binding proteins residing in plastids could have evolved from eukaryotic proteins involved in DNA transaction processes in the nucleus (Kodama, 2007; Kodama and Sano, 2007). An intensively studied ptNAP is the plastid envelope DNA binding protein (PEND) having a basic region and a leucine zipper (bZIP) domain. PEND was originally discovered in developing pea chloroplasts (Sato et al., 1993) and shown to tether nucleoids to the inner envelope membrane where replication takes place (Sato et al., 1993, 1998). Interestingly, a PEND:GFP fusion protein was shown to be targeted to the nucleus when the plastid targeting sequence was deleted (Terasawa and Sato, 2009).

Several ptNAPs are multifunctional (Krupinska et al., 2013). One of the most abundant proteins in nucleoids is DCP68 (Cannon et al., 1999) which is identical with sulfite reductase (SiR), an enzyme catalyzing the reduction of sulfite to sulfide (Sato, 2001). SiR was found to bind and compact ptDNA, thereby having a negative effect on *in vitro* replication (Cannon et al.,

1999) and transcription (Sekine et al., 2002, 2007; Sato et al., 2003) as well as on chloroplast development (Kang et al., 2010). However, its compacting effect on ptDNA differs from the mode of action of HU-like proteins. In contrast to the DNA packed by HU, DNA tightly packed by SiR, is in an inactive state and is not available for DNA transacting enzymes. SiR was suggested to repress transcriptional activity in non-photosynthetic plastids of spores and seeds (Sato et al., 2003). In some aspects SiR might rather play a similar role as Dps, the bacterial DNA binding protein abundant in starved cells (Dillon and Dorman, 2010). As mentioned above, SiR has the ability to tightly compact DNA, but the impact of this condensation on DNA protection has not been studied so far. On the other hand, considering the association of SiR with nucleoids in mature chloroplasts, SiR may be important beyond the seed stage, putatively playing a role in selective silencing of chloroplast encoded genes.

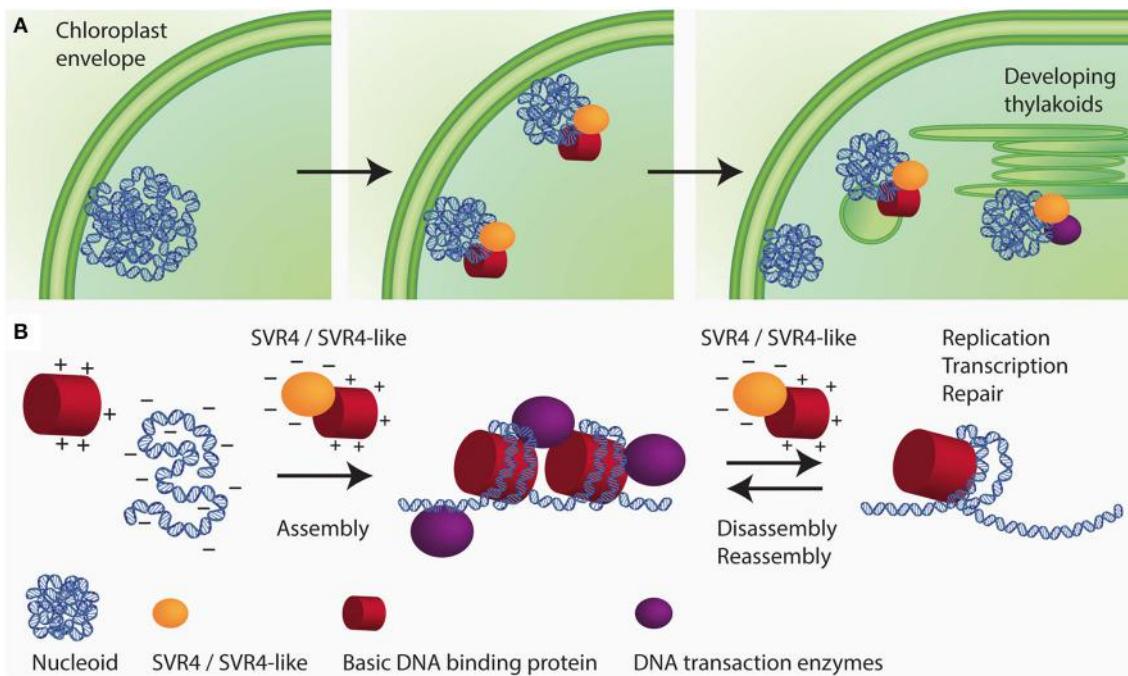
Novel candidates for architectural ptNAP proteins were identified in a recent study by Melonek and coworkers (2012). A group of six organelle targeted, low molecular weight proteins have a SWIB (switch/sucrose nonfermentable complex B) domain that is typically found in ATP-dependent chromatin remodelers of the nucleus. One of them, SWIB-4, has a histone H1-motif next to the SWIB domain and was shown to bind to DNA. The recombinant SWIB-4 protein was shown to induce compaction and condensation of nucleoids and to functionally complement a mutant of *E. coli* lacking the histone-like nucleoid structuring protein H-NS (Melonek et al., 2012). Interestingly, SWIB domain proteins are also found in *Chlamydophila felis*. This species has a histone 1 like protein (Hc1) and a stand-alone SWIB domain protein, the only type of SWIB proteins found in bacteria. Chlamydiae are a group of bacteria living as endosymbionts and parasites in other bacteria or in eukaryotic cells. Phylogenetic analyses suggested that an ancestral member of the group of Chlamydiae facilitated the establishment of the primary endosymbiosis between cyanobacteria and an early eukaryote (Huang and Gogarten, 2007), and that Chlamydiae have contributed at least 55 genes to plant genomes. Genes encoding members of this subgroup of the SWIB domain proteins (Melonek et al., 2012) are found in the sequenced genomes of all land plants, but not in those of algae. The homology is very high among the sequences found in angiosperms, gymnosperms, mosses and clubmosses (Lycopodiaceae).

Other highly abundant proteins of nucleoids are WHIRLY1 (pTAC1) and WHIRLY3 (pTAC11) that have been found in the proteome of transcriptionally active chromosomes (TAC) isolated from Arabidopsis chloroplasts (Pfalz et al., 2006), and that belong to a small family of single-stranded DNA binding proteins specifically found in higher plants. While in most plants one WHIRLY protein is targeted to chloroplasts and one to mitochondria, in Arabidopsis two are targeted to chloroplasts (WHIRLY1, WHIRLY3) (Krause et al., 2005). In other plants such as barley and maize, plastids contain only one WHIRLY protein which is associated to nucleoids (Prikryl et al., 2008; Melonek et al., 2010; Majeran et al., 2012). It has been suggested that WHIRLY1 of barley chloroplasts is located at the periphery of nucleoids, because it is lost during purification of TAC (Melonek et al., 2010). In chloroplasts of transgenic barley plants with an RNAi mediated

knockdown of the *WHIRLY1* gene, only few tiny nucleoids are found besides unpacked DNA covering large areas in the organelle (Krupinska et al., 2014). This indicates that WHIRLY1 plays an important role in condensation of plastid DNA of a subset of nucleoids.

Additional nucleoid associated proteins specifically found in higher plants are the SVR4 (suppressor of variegation) and SVR4-like proteins, which were originally identified as important proteins for chloroplast development in *Arabidopsis* (Yu et al., 2011) and were named MRL7 and MRL7-like in another study (Qiao et al., 2011). In the lower land plants, *Physcomitrella patens* and *Selaginella moellendorffii*, only one protein with sequence similarities to both *Arabidopsis* proteins, SVR4 (MRL7) and SVR4-like (MRL7-like), was found (Qiao et al., 2011). In *Arabidopsis*, the knockout mutants of either SVR4 or SVR4-like are seedling lethal and can only be grown on media supplemented with sucrose giving rise to pigment deficient plants that are, however, unable to complete their life cycle. SVR4 and SVR4-like are already present in plastids at early stages of chloroplast development. In the absence of either SVR4 or SVR4-like, the nucleoid organization was found to be disturbed. Fewer and larger nucleoids with the tendency to form ring-like structures were detected in the mutants (Powikrowska et al., 2014). In the primary amino acid sequence SVR4 and SVR4-like contain 20% negatively charged glutamic or aspartic acid residues which is a characteristic feature for chaperone proteins, that might assist in assembly and maintenance of DNA/RNA-protein complexes (Powikrowska et al., 2014). During the assembly and dynamic functioning of DNA/RNA-protein complexes there is a high risk of random aggregation due to the fact that very strong interactions occur between the negatively charged nucleic acids and basic proteins such as histones and ribosomal subunits (Jäkel et al., 2002; Frehlick et al., 2007; Lindström, 2011). Negatively charged proteins have been reported to act as chaperones for exposed basic domains most probably by mimicking the interaction with nucleic acids (Jäkel et al., 2002; Koch et al., 2012). It has been proposed that SVR4 and SVR4-like are putative functional homologs of negatively charged molecular chaperones involved in establishing proper ptDNA-protein interaction in developing chloroplasts (Figure 2). The expression of the genes encoding SVR4 and SVR4-like was reported to be high in growing tissues, i.e., young leaves, flowers and stems (Qiao et al., 2011). Interestingly, the level of the SVR4-like is high in the meristematic tissue at the base of a barley leaf, whereas the level of SVR4 increases with chloroplast development (Powikrowska et al., 2014) indicating that the two proteins might have similar functions, but at different stages of chloroplast development.

In conclusion, it seems that most ptNAPs identified so far are unique to land plants (Table 1). Surprisingly, the maize homologs of SiR and PEND were not detected in the extensive nucleoid proteome of maize plastids, although they were found in unfractionated maize plastids (Majeran et al., 2012). It remains to be determined whether the altered distribution of the proteins reflects differences between the different groups of plants or whether it is caused by the method used for preparation of nucleoids. Pfalz and Pfannschmidt (2013) reported that also most of the nucleoid proteins found to be associated with PEP do



**FIGURE 2 | Schematic drawing of SVR4/SVR4-like functioning as putative chaperones for ptNAPs during chloroplast development. (A)**

The spatial arrangement of plastid nucleoids dynamically changes in close relationship with the development of the inner membrane system of the plastids. During chloroplast development the nucleoids decrease in size but increase in number. The segregation process is thought to take place on the envelope membrane and eventually distribute to the thylakoids. **(B)** During

the assembly of DNA-protein complexes in the developing chloroplast, there is a high risk of random aggregation due to the fact that very strong interaction occurs between oppositely charged molecular species, i.e., negatively charged DNA and positively charged structural proteins. The negatively charged proteins SVR4 and SVR4-like, transiently interact with positively charged DNA binding proteins, supporting essential DNA transaction processes in chloroplasts.

not have orthologous proteins in the green alga Chlamydomonas indicating that also the prokaryotic transcription machinery has been altered during evolution of land plants. A striking feature of some plastid DNA binding proteins, such as SIR and also CND41, is their multifunctionality (Murakami et al., 2000; Krupinska et al., 2013). A unique example for a multifunctional ptNAP is WHIRLY1. Besides its impact on compactness of a subset of chloroplast nucleoids (Krupinska et al., 2014), WHIRLY1 (pTAC1) has been reported to affect RNA splicing in plastids (Prikryl et al., 2008; Melonek et al., 2010), to be important for DNA stability (Maréchal and Brisson, 2010) and to act furthermore as a transcription factor in the nucleus (Desveaux et al., 2000; Grabowski et al., 2008; Xiong et al., 2009; Krupinska et al., 2014). It remains to be investigated whether the architectural role of WHIRLY1 is connected to its other functions.

## DYNAMICS OF NUCLEOID ORGANIZATION DURING CHLOROPLAST DEVELOPMENT

The number and positions of nucleoids were shown to depend on the developmental stage of the plastids (Kuroiwa et al., 1981; Miyamura et al., 1986). Intensive remodeling of nucleoids occurs during the development of proplastids to photosynthetically competent chloroplasts and during interconversions between different plastid types (Hashimoto, 1985; Kuroiwa, 1991; Chi-Ham et al., 2002). Proplastids contain a cluster of nucleoids located in the

center of the plastids. At the beginning of seed germination, these nucleoids are considered to move to the envelope, where extensive DNA amplification takes place, and eventually the enlarged nucleoids form a spherical ring (Figure 3A). Upon illumination, during transition from proplastids to chloroplasts, small nucleoids are distributed along developing thylakoid membranes (for review see Sakai et al., 2004). Sections from barley primary foliage leaves were stained with SYBR Green to show typical stages of nucleoid organization. At the border between white and green stripes of a heterozygous leaf of the mutant *albosstrians* small undifferentiated and photosynthetically inactive plastids were found besides chloroplasts. As observed in proplastids of leaf primordial of imbibed wheat seeds (Miyamura et al., 1986), the nucleoid in the plastids of white *albosstrians* leaves and leaf stripes appears to be ring-shaped. The ring-shaped nucleoid is typical for proplastids developing in darkness. In a basal segment from a primary foliage leaf of seedlings grown for 5 days in the light, developing chloroplasts were found to be organized as a necklace of pearls in the peripheries of the organelles indicating a light-dependent disintegration of the nucleoid ring as proposed previously (Miyamura et al., 1986). In the upper part of a leaf from 7 days old seedlings, mature chloroplasts with many tiny nucleoids attached to thylakoids were found. During development of barley seedlings in darkness, proplastids differentiate into etioplasts, where a few large nucleoids are found

that might be distributed at the periphery of prolamellar bodies (**Figure 3B**). Temporal changes of nucleoid structure have also been studied intensively in variegated leaves of *Arabidopsis* mutants, e.g., *var2* (Sakamoto et al., 2009). Plastids in white leaf sectors were observed to contain few large nucleoids. During chloroplast development nucleoids were observed to become smaller in size, more dense and more abundant (Sakamoto et al., 2009). The only protein so far identified to be involved in the distribution of nucleoids, YlmG1 (**Table 1**), is of prokaryotic origin. Overexpression or knockdown of the gene was shown to impair nucleoid partitioning (Kabeya et al., 2010).

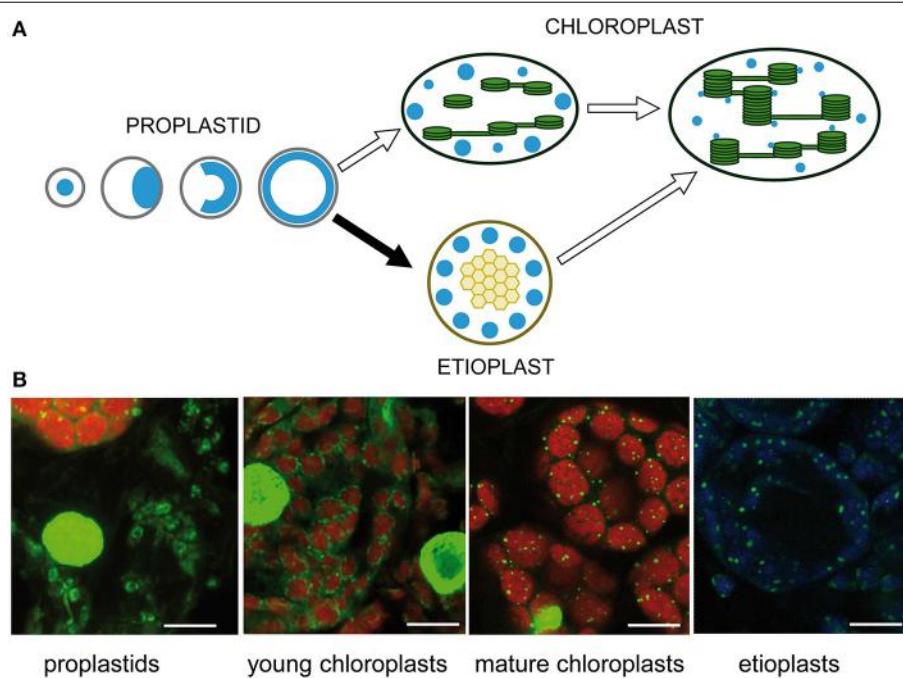
Attachment of nucleoids to membranes was proposed to be important for organization, replication and transcription of ptDNA (Sato et al., 1993; Sato, 2001). In chloroplasts, formation of thylakoids seems to be tightly linked with nucleoid morphology and distribution (Kobayashi et al., 2012). Nucleoid structure and transcriptional activity are not affected in mutants developing residual thylakoids with altered lipid composition and impaired photosynthetic machinery (Kobayashi et al., 2012). These studies suggest that the formation of the thylakoid system and the attachment of nucleoids to these membranes precede the assembly of the photosynthetic machinery. On the other hand, ptDNA displays reduced compaction in plastids of the yellow leaf tissue of mutants with silencing of the *CHLD* (Mg chelatase subunit D) and *CHLI* (Mg chelatase subunit I) genes. These mutants possess thylakoids but lack grana stacks and are devoid of the photosynthetic complexes resulting in compromised photosynthesis (Luo et al., 2013). Taken together, these studies

demonstrate that thylakoid formation during chloroplast development and alterations in shape and distribution of nucleoids are interconnected processes.

## FUNCTIONAL IMPLICATIONS OF THE THYLAKOID ASSOCIATION OF NUCLEOIDS IN CHLOROPLASTS

Interestingly, architectural reorganization of nucleoids during light-dependent chloroplast differentiation is correlated with a switch in polymerase usage: transcription of PEP (plastid encoded polymerase)-dependent genes increases, whereas at the same time the expression of NEP (nuclear encoded polymerase)-dependent genes decreases (Liere et al., 2011). It was recently proposed that the structural establishment of the transcriptional subdomain within the nucleoid represents a bottleneck in chloroplast development (Pfalz and Pfannschmidt, 2013). In this context it has been proposed that the assistance of DNA/RNA-protein assembly factors SVR4 and SVR4-like is required for expression of a set of chloroplast encoded genes involved in chloroplast formation (Qiao et al., 2011; Powikowska et al., 2014).

Of particular importance for the activity of nucleoids is their association with the thylakoid membranes where the photosynthetic machinery undergoes changes in composition in response to environmental conditions. A prerequisite for remodeling of the photosynthetic apparatus is the regulation of plastid gene transcription in response to light-dependent changes in the redox state of the photosynthetic apparatus (Pfannschmidt et al., 1999). Thereby the composition of the photosynthetic apparatus can continuously be adjusted to the ever changing environmental



**FIGURE 3 | Nucleoid organization during chloroplast development.**

(A) Schematic drawing adapted from Sakai et al. (2004). (B) Detection of nucleoids by fluorescence microscopy of barley leaf sections stained with SYBR Green. The left image shows undifferentiated rudimentary plastids in white parts of a striped leaf of the mutant *albostrans*

besides a green part containing chloroplasts. Chloroplasts were analyzed in sections of primary foliage leaves of barley seedlings grown for either 5 or 7 days in the light. Etioplasts were analyzed in primary foliage leaves of seedlings after 5 days in darkness. The bar represents 10 μm.

conditions. Recent research in Chlamydomonas has clearly shown that chloroplast nucleoids are able to sense the redox state and that also the DNA replication activity can be adjusted accordingly (Kabeya and Miyagishima, 2013). SVR4 seems to be among the nucleoid proteins that are able to sense the redox state and to modulate the nucleoid architecture in response to redox changes. SVR4 was reported to possess disulfide reductase activity *in vitro* and to interact *in vivo* with thioredoxin Z (TrxZ), as well as with the two nucleoid associated superoxide dismutases FSD2 and FSD3 (Qiao et al., 2013; Yua et al., 2013). TrxZ regulates the redox state of proteins in response to light and has been shown to be associated with PEP (Steiner et al., 2011; Pfalz and Pfannschmidt, 2013) and to be required for transcriptional activity (Arsova et al., 2010; Schröter et al., 2010). FSD2 and FSD3 are two iron superoxide dismutases found in the chloroplast nucleoid associated with PEP. Both proteins were shown to act as ROS scavengers within the nucleoids (Myouga et al., 2008). It is likely that the redox state of nucleoid associated proteins is regulated by electrons provided from the photosynthetic machinery. It is interesting to note that the enzymatic activity of SiR was shown to be regulated by photoreduced ferredoxin. The DNA binding of SiR did not affect the enzymatic activity suggesting that both ferredoxin and sulfites are accessible to SiR within the nucleoids (Sekine et al., 2007). It remains, however, unknown whether the redox activity of SiR has an impact on DNA binding.

In this context, proteins found to be located at the interface between nucleoids and the thylakoid membrane are of particular interest. MFP1 was proposed to anchor nucleoids to thylakoids in chloroplasts (Jeong et al., 2003). WHIRLY1/pTAC1 is another nucleoid associated protein (Pfalz et al., 2006) associated with thylakoid membranes. Prikryl et al. (2008) showed that the attachment to thylakoid membranes is disrupted by DNaseI. WHIRLY1 was shown to form 24-mer complexes (Cappadocia et al., 2010, 2012) and was proposed to function analogously as the oligomeric NONEXPRESSOR OF PR1 (NPR1) in the cytoplasm (Foyer et al., 2014). Upon changes in the redox state of the photosynthetic machinery the complexes might get monomerized and the monomer might change gene expression in the nucleus. In accordance with this model the WHIRLY3 protein in Arabidopsis chloroplasts was identified among the redox-sensitive proteins (Ströher and Dietz, 2008). Another protein found to be distributed between thylakoids and the nucleoid is pTAC16. Its phosphorylation in response to redox-changes of the photosynthetic apparatus was suggested to regulate membrane-anchoring functions of the nucleoid (Ingelsson and Vener, 2012). It seems that the nucleoid containing besides WHIRLY1 further central proteins shown to be involved in plastid-to-nucleus signaling such as GUN1 (Koussevitzky et al., 2007) and PRIN2 (Kindgren et al., 2012; Barajas-López Jde et al., 2013), is the place where redox signals known to induce changes in nuclear gene expression are integrated.

Transgenic plants with different levels of nucleoid/thylakoid associated proteins might help to elucidate the roles of these proteins in linking the activity of the photosynthetic machinery to organization and expression of plastid genes as well as the expression of nuclear genes.

## MECHANISMS UNDERLYING THE RESTRUCTURING OF PLASTID NUCLEOIDS

In the nucleus the availability of DNA for transcription is regulated mainly by posttranslational modifications, whereas in bacteria regulation of transcription involves the exchange of DNA binding proteins (Luijsterburg et al., 2008) and changes in the compaction of the nucleoid (Berger et al., 2010). It seems, that in plastids the architecture of the nucleoids is regulated by both kinds of mechanisms. Both in the nucleus and in plastid nucleoids, posttranslational modifications are important for nucleoid associated processes. For MFP, SiR and SWIB-4 it was shown that the binding to DNA is regulated by phosphorylation. The three proteins seem not to bind to DNA when they are phosphorylated (Chi-Ham et al., 2002; Jeong et al., 2004; Melonek et al., 2012). Several kinases were found to be associated with nucleoids, e.g., the fructokinase like protein FLN1/2, the casein kinase CK-II, as well as two atypical ABCK1 type kinases (Lundquist et al., 2012). Probably, other posttranslational modifications are likely to play important roles as well. Counterparts of enzymes known to be involved in histone modifications in the nucleus, such as the Arabidopsis SET-domain proteins ATXR5 and ATXR6 involved in methylation (Raynaud et al., 2006; Jacob et al., 2009) and de-acetylases were found in plastids (Chung et al., 2009).

It remains to be shown whether the different packaging of DNA in different regions of the nucleoid changes during development and in response to environmental cues. Whether, however, the central body of nucleoids with dense packaging (Sakai et al., 2004) can be compared with eukaryotic heterochromatin (Sato, 2001) remains questionable. It rather seems that DNA packaging is beneficial for a high transcriptional activity as it was described for bacteria as well (Dillon and Dorman, 2010; Krupinska et al., 2013).

## CONCLUDING REMARKS

The comparison of DNA organization in plastids, nucleus and bacteria shows that the shaping and organization of plastid nucleoids involves novel organelle specific mechanisms resembling those acting on eukaryotic chromatin besides mechanisms described for eubacterial nucleoids. During evolution of plants, the architectural proteins of bacterial nucleoids have been lost and replaced by new proteins. Some of these are enzymes that have acquired an additional function as DNA binding proteins. Others might have been contributed by Chlamydiae which facilitated establishment of the primary endosymbiosis between an early eukaryote and the cyanobacterial ancestor of plastids. These proteins do not exhibit sequence or structural conservation with the eukaryotic histones, but similar to the histones they might be regulated by posttranslational modifications.

In comparison to eukaryotic chromatin, nucleoids of plastids have as those of bacteria a more open structure, that allows easy access for DNA transaction enzymes. The enrichment of enzymes involved in RNA processing and translation in the nucleoid fraction suggests that transcription, RNA processing and translation are tightly connected with each other.

Similarly to bacteria, also in plastids, membranes seem to play a key role in the organization and maintenance of nucleoids. In chloroplasts, the proximity of nucleoids and photosynthetic machinery as well as the presence of several redox active proteins in nucleoids, allows for a tight coordination of photosynthesis and nucleoid function, i.e., replication and gene expression. It is striking that not only particular enzymes involved in gene expression but also architectural proteins are controlled by redox signals. Thereby these proteins might have a tremendous impact on the different enzymatic activities associated with nucleoids; in particular replication, transcription and DNA repair.

The architectural organization of the plastid genetic machinery is not well understood. Since principles underlying the dynamic shaping of genomes are uniform in all forms of life, the knowledge about DNA organization in bacteria and eukaryotes can be used in future studies on the dynamic architecture of chloroplast nucleoids.

## ACKNOWLEDGMENTS

We thank Maria Mulisch and Christine Desel for providing images for **Figures 1** and **3B**, and the Central Microscopy of the University of Kiel for providing electron microscopy and confocal microscopy facilities. Lærke Marie Münter Lassen is thanked for help with **Figure 2**. The authors gratefully acknowledge financial support from the VILLUM Center of Excellence “Plant Plasticity” and from the “Center of Synthetic Biology” funded by the UNIK research initiative of the Danish Ministry of Science, Technology and Innovation.

## REFERENCES

- Arsova, B., Hoja, U., Wimmelbacher, M., Greiner, E., Ustün, S., Melzer, M., et al. (2010). Plastidial thioredoxin z interacts with two fructokinase-like proteins in a thiol-dependent manner: evidence for an essential role in chloroplast development in *Arabidopsis* and *Nicotiana benthamiana*. *Plant Cell* 22, 1498–1515. doi: 10.1105/tpc.109.071001
- Azam, T. A., and Ishihama, A. (1999). Twelve species of the nucleoid-associated protein from *Escherichia coli*. Sequence recognition specificity and DNA binding affinity. *J. Biol. Chem.* 274, 33105–33113. doi: 10.1074/jbc.274.46.33105
- Baumgartner, B. J., Rapp, J. C., and Mullet, J. E. (1989). Plastid transcription activity and DNA copy number increase early in barley chloroplast development. *Plant Physiol.* 89, 1011–1018. doi: 10.1104/pp.89.3.1011
- Bendich, A. J. (2004). Circular chloroplast chromosomes: the grand illusion. *Plant Cell* 16, 1661–1666. doi: 10.1105/tpc.160771
- Bendich, A. J., and Drlica, K. (2000). Prokaryotic and eukaryotic chromosomes: what's the difference? *Bioessays* 22, 481–486. doi: 10.1002/(SICI)1521-1878(200005)22:5&lt;481::AID-BIES10&gt;3.0.CO;2-T
- Berger, M., Farcas, A., Geertz, M., Zhelyazkova, P., Brix, K., Travers, A., et al. (2010). Coordination of genomic structure and transcription by the main bacterial nucleoid-associated protein HU. *EMBO Rep.* 11, 59–64. doi: 10.1038/embor.2009.232
- Bock, R., and Timmis, J. N. (2008). Reconstructing evolution: gene transfer from plastids to the nucleus. *Bioessays* 30, 556–566. doi: 10.1002/bies.20761
- Boffey, S. A., Ellis, J. R., Sellén, G., and Leech, R. M. (1979). Chloroplast division and DNA synthesis in light-grown wheat leaves. *Plant Physiol.* 64, 502–505. doi: 10.1104/pp.64.3.502
- Browning, D. F., Grainger, D. C., and Busby, S. J. (2010). Effects of nucleoid-associated proteins on bacterial chromosome structure and gene expression. *Curr. Opin. Microbiol.* 13, 773–780. doi: 10.1016/j.mib.2010.09.013
- Campos, E. I., and Reinberg, D. (2009). Histones: annotating chromatin. *Annu. Rev. Genet.* 43, 559–599. doi: 10.1146/annurev.genet.032608.103928
- Cannon, G. C., Ward, L. N., Case, C. I., and Heinhorst, S. (1999). The 68 kDa DNA compacting nucleoid protein from soybean chloroplasts inhibits DNA synthesis *in vitro*. *Plant Mol. Biol.* 39, 835–845. doi: 10.1023/A:1006135615924
- Cappadocia, L., Maréchal, A., Parent, J. S., Lepage, E., Sygusch, J., and Brisson, N. (2010). Crystal structures of DNA-Whirly complexes and their role in *Arabidopsis* organelle genome repair. *Plant Cell* 22, 1849–1867. doi: 10.1105/tpc.109.071399
- Cappadocia, L., Parent, J. S., Zampini, E., Lepage, E., Sygusch, J., and Brisson, N. (2012). A conserved lysine residue of plant Whirly proteins is necessary for higher order protein assembly and protection against DNA damage. *Nucleic Acids Res.* 40, 258–269. doi: 10.1093/nar/gkr740
- Chi-Ham, C. L., Keaton, M. A., Cannon, G. C., and Heinhorst, S. (2002). The DNA-compactting protein DCP68 from soybean chloroplasts is ferredoxin:sulfite reductase and co-localizes with the organellar nucleoid. *Plant Mol. Biol.* 49, 621–631. doi: 10.1023/A:1015500431421
- Chung, P. J., Kim, Y. S., Park, S. H., Nahm, B. H., and Kim, J. K. (2009). Subcellular localization of rice histone deacetylases in organelles. *FEBS Lett.* 583, 2249–2254. doi: 10.1016/j.febslet.2009.06.003
- Dame, R. T., Tark-Dame, M., and Schiessl, H. (2011). A physical approach to segregation and folding of the *Caulobacter crescentus* genome. *Mol. Microbiol.* 82, 1311–1315. doi: 10.1111/j.1365-2958.2011.07898.x
- Day, A., and Madesis, P. (2007). “DNA replication, recombination, and repair in plastids,” in *Cell and Molecular Biology of Plastids*, ed R. Bock (Heidelberg: Springer-Verlag), 65–119. doi: 10.1007/4735\_2007\_0231
- de la Serna, I. L., Ohkawa, Y., and Imbalzano, A. N. (2006). Chromatin remodelling in mammalian differentiation: lessons from ATP-dependent remodellers. *Nat. Rev. Genet.* 7, 461–473. doi: 10.1038/nrg1882
- Delius, H. (1974). Electron microscopic visualization of the folded chromosome of *Escherichia coli*. *J. Mol. Biol.* 82, 107–109. doi: 10.1016/0022-2836(74)90577-4
- Deng, X. W., Wing, R. A., and Gruissem, W. (1989). The chloroplast genome exists in multimeric forms. *Proc. Natl. Acad. Sci. U.S.A.* 86, 4156–4160. doi: 10.1073/pnas.86.11.4156
- Desveaux, D., Després, C., Joyeux, A., Subramaniam, R., and Brisson, N. (2000). PBF-2 is a novel single-stranded DNA binding factor implicated in PR-10a gene activation in potato. *Plant Cell* 12, 1477–1489. doi: 10.1105/tpc.12.8.1477
- Dillon, S. C., and Dorman, C. J. (2010). Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. *Nat. Rev. Microbiol.* 8, 185–195. doi: 10.1038/nrmicro2261
- Dorman, C. J., and Deighan, P. (2003). Regulation of gene expression by histone-like proteins in bacteria. *Curr. Opin. Genet. Dev.* 13, 179–184. doi: 10.1016/S0959-473X(03)00025-X
- Drlica, K., and Rouviere-Yaniv, J. (1987). Histonelike proteins of bacteria. *Microbiol. Rev.* 51, 301–319.
- Espeli, O., Mercier, R., and Boccard, F. (2008). DNA dynamics vary according to macrodomain topography in the *E. coli* chromosome. *Mol. Microbiol.* 68, 1418–1427. doi: 10.1111/j.1365-2958.2008.06239.x
- Foyer, C. H., Karpinska, B., and Krupinska, K. (2014). The functions of WHIRLY1 and REDOX-RESPONSIVE TRANSCRIPTION FACTOR 1 in cross tolerance responses in plants: a hypothesis. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 369:20130226. doi: 10.1098/rstb.2013.0226
- Frehlick, L. J., Eirín-López, J. M., and Ausió, J. (2007). New insights into the nucleophosmin/nucleoplasmmin family of nuclear chaperones. *Bioessays* 29, 49–59. doi: 10.1002/bies.20512
- Fudenberg, G., and Mirny, L. A. (2012). Higher-order chromatin structure: bridging physics and biology. *Curr. Opin. Genet. Dev.* 22, 115–124. doi: 10.1016/j.gde.2012.01.006
- Golczyk, H., Greiner, S., Wanner, G., Weihe, A., Bock, R., Börner, T., et al. (2014). Chloroplast DNA in mature and senescing leaves: a reappraisal. *Plant Cell* 26, 847–854. doi: 10.1105/tpc.113.117465
- Grabowski, E., Miao, Y., Mulisch, M., and Krupinska, K. (2008). Single-stranded DNA-binding protein Whirly1 in barley leaves is located in plastids and the nucleus of the same cell. *Plant Physiol.* 147, 1800–1804. doi: 10.1104/pp.108.122796
- Grainger, D. C., Hurd, D., Goldberg, M. D., and Busby, S. J. (2006). Association of nucleoid proteins with coding and non-coding segments of the *Escherichia coli* genome. *Nucleic Acids Res.* 34, 4642–4652. doi: 10.1093/nar/gkl542

- Grasser, K. D. (1995). Plant chromosomal high-mobility group (HMG) proteins. *Plant J.* 7, 185–192. doi: 10.1046/j.1365-313X.1995.7020185.x
- Green, B. R. (2011). Chloroplast genomes of photosynthetic eukaryotes. *Plant J.* 66, 34–44. doi: 10.1111/j.1365-313X.2011.04541.x
- Hashimoto, H. (1985). Changes in distribution of nucleoids in developing and dividing chloroplast and etioplasts of *Avena sativa*. *Protoplasma* 127, 119–127. doi: 10.1007/BF01273708
- Huang, J., and Gogarten, J. P. (2007). Did an ancient chlamydial endosymbiosis facilitate the establishment of primary plastids? *Genome Biol.* 8:R99. doi: 10.1186/gb-2007-8-6-r99
- Ingelson, B., and Vener, A. V. (2012). Phosphoproteomics of *Arabidopsis* chloroplasts reveals involvement of the STN7 kinase in phosphorylation of nucleoid protein pTAC16. *FEBS Lett.* 586, 1265–1271. doi: 10.1016/j.febslet.2012.03.061
- Jacob, Y., Feng, S., LeBlanc, C. A., Bernatavichute, Y. V., Stroud, H., Cokus, S., et al. (2009). ATXR5 and ATXR6 are H3K27 monomethyltransferases required for chromatin structure and gene silencing. *Nat. Struct. Mol. Biol.* 16, 763–768. doi: 10.1038/nsmb.1611
- Jäkel, S., Mingot, J. M., Schwarzmaier, P., Hartmann, E., and Görlich, D. (2002). Importins fulfil a dual function as nuclear import receptors and cytoplasmic chaperones for exposed basic domains. *EMBO J.* 21, 377–386. doi: 10.1093/emboj/21.3.377
- Jeong, S. Y., Peffer, N., and Meier, I. (2004). Phosphorylation by protein kinase CKII modulates the DNA-binding activity of a chloroplast nucleoid-associated protein. *Planta* 219, 298–302. doi: 10.1007/s00425-004-1215-8
- Jeong, S. Y., Rose, A., and Meier, I. (2003). MFP1 is a thylakoid-associated, nucleoid-binding protein with a coiled-coil structure. *Nucleic Acids Res.* 31, 5175–5185. doi: 10.1093/nar/gkg693
- Kabeya, Y., and Miyagishima, S. Y. (2013). Chloroplast DNA replication is regulated by the redox state independently of chloroplast division in *Chlamydomonas reinhardtii*. *Plant Physiol.* 161, 2102–2112. doi: 10.1104/pp.113.216291
- Kabeya, Y., Nakanishi, H., Suzuki, K., Ichikawa, T., Kondou, Y., Matsui, M., et al. (2010). The YlmG protein has a conserved function related to the distribution of nucleoids in chloroplasts and cyanobacteria. *BMC Plant Biol.* 10:57. doi: 10.1186/1471-2229-10-57
- Kang, Y. W., Lee, J. Y., Jeon, Y., Cheong, G. W., Kim, M., and Pai, H. S. (2010). In vivo effects of NbSIR silencing on chloroplast development in *Nicotiana benthamiana*. *Plant Mol. Biol.* 72, 569–583. doi: 10.1007/s11103-009-9593-8
- Karcher, D., Köster, D., Schadach, A., Klevesath, A., and Bock, R. (2009). The Chlamydomonas chloroplast HLP protein is required for nucleoid organization and genome maintenance. *Mol. Plant* 2, 1223–1232. doi: 10.1093/mp/ssp083
- Kindgren, P., Kremnev, D., Blanco, N. E., de Dios Barajas López, J., Fernández, A. P., Tellgren-Roth, C., et al. (2012). The plastid redox insensitive 2 mutant of *Arabidopsis* is impaired in PEP activity and high light-dependent plastid redox signalling to the nucleus. *Plant J.* 70, 279–291. doi: 10.1111/j.1365-313X.2011.04865.x
- Kobayashi, K., Narise, T., Sonoike, K., Hashimoto, H., Sato, N., Kondo, M., et al. (2012). Role of galactolipid biosynthesis in coordinated development of photosynthetic complexes and thylakoid membranes during chloroplast biogenesis in *Arabidopsis*. *Plant J.* 73, 250–261. doi: 10.1111/tpj.12028
- Kobayashi, T., Takahara, M., Miyagishima, S. Y., Kuroiwa, H., Sasaki, N., Ohta, N., et al. (2002). Detection and localization of a chloroplast-encoded HU-like protein that organizes chloroplast nucleoids. *Plant Cell* 14, 1579–1589. doi: 10.1105/tpc.002717
- Koch, B., Mitterer, V., Niederhauser, J., Stanborough, T., Murat, G., Rechberger, G., et al. (2012). Yar1 protects the ribosomal protein Rps3 from aggregation. *J. Biol. Chem.* 287, 21806–21815. doi: 10.1074/jbc.M112.365791
- Kodama, Y. (2007). Plastidic proteins containing motifs of nuclear transcription factors. *Plant Biotechnol.* 24, 165–170. doi: 10.5511/plantbiotechnology.24.165
- Kodama, Y., and Sano, H. (2007). Functional diversification of a basic helix-loop-helix protein due to alternative transcription during generation of amphidiploidy in tobacco plants. *Biochem. J.* 403, 493–499. doi: 10.1042/BJ20070011
- Koussevitzky, S., Nott, A., Mockler, T. C., Hong, F., Sachetto-Martins, G., Surpin, M., et al. (2007). Signals from chloroplasts converge to regulate nuclear gene expression. *Science* 316, 715–719. doi: 10.1126/science.1140516
- Kowallik, K. V., and Herrmann, R. G. (1972). Variable amounts of DNA related to size of chloroplasts. IV. Three-dimensional arrangement of DNA in fully differentiated chloroplasts of *Beta vulgaris* L. *J. Cell Sci.* 11, 357–377.
- Krause, K., Falk, J., Humbeck, K., and Krupinska, K. (1998). Responses of the transcriptional apparatus of barley chloroplasts to a prolonged dark period and to subsequent reillumination. *Physiol. Plant.* 104, 143–152. doi: 10.1034/j.1399-3054.1998.1040201.x
- Krause, K., Kilbienski, I., Mulisch, M., Rödiger, A., Schäfer, A., and Krupinska, K. (2005). DNA-binding proteins of the Whirly family in *Arabidopsis thaliana* are targeted to the organelles. *FEBS Lett.* 579, 3707–3712. doi: 10.1016/j.febslet.2005.05.059
- Krupinska, K., Dähnhardt, D., Fischer-Kilbienski, I., Kucharewicz, W., Scharrenberg, C., Trösch, M., et al. (2014). Identification of WHIRLY1 as a factor binding to the promoter of the stress- and senescence-associated gene HvS40. *J. Plant Growth Regul.* 33, 91–105. doi: 10.1007/s00344-013-9378-9
- Krupinska, K., and Humbeck, K. (2004). “Photosynthesis and chloroplast breakdown,” in *Plant Cell Death Processes*, ed L. D. Noodén (San Diego, CA: Elsevier Academic Press), 169–188.
- Krupinska, K., Melonek, J., and Krause, K. (2013). New insights into plastid nucleoid structure and functionality. *Planta* 237, 653–664. doi: 10.1007/s00425-012-1817-5
- Krupinska, K., Oetke, S., Desel, C., Mulisch, M., Schäfer, A., Hollmann, J., et al. (2014). WHIRLY1 is a major organizer of chloroplast nucleoids. *Front. Plant Physiol.* 5:432. doi: 10.3389/fpls.2014.00432
- Kuroiwa, T. (1991). The replication, differentiation, and inheritance of plastids with emphasis on the concept of organelle nuclei. *Int. Rev. Cytol.* 128, 1–62. doi: 10.1016/S0074-7696(08)60496-9
- Kuroiwa, T., Suzuki, T., Ogawa, K., and Kawano, S. (1981). The chloroplast nucleus: distribution, number, size, and shape, and a model for the multiplication of the chloroplast genome during chloroplast development. *Plant Cell Physiol.* 22, 381–396.
- Lam, E., and Chua, N. H. (1987). Chloroplast DNA gyrase and *in vitro* regulation of transcription by template topology and novobiocin. *Plant Mol. Biol.* 8, 415–424. doi: 10.1007/BF00015819
- Li, G., and Reinberg, D. (2011). Chromatin higher-order structures and gene regulation. *Curr. Opin. Genet. Dev.* 21, 175–186. doi: 10.1016/j.gde.2011.01.022
- Li, G.-W., Berg, O. G., and Elf, J. (2009). Effects of macromolecular crowding and DNA looping on gene regulation kinetics. *Nat. Phys.* 5, 294–297. doi: 10.1038/nphys1222
- Liere, K., and Börner, T. (2013). “Development-dependent changes in the amount and structural organization of plastid DNA,” in *Plastid Development in Leaves during Growth and Senescence*, eds B. Biswal, K. Krupinska and U. C. Biswal (Dordrecht: Springer), 215–238. doi: 10.1007/978-94-007-5724-0\_11
- Liere, K., Weihe, A., and Börner, T. (2011). The transcription machineries of plant mitochondria and chloroplasts: composition, function, and regulation. *J. Plant Physiol.* 168, 1345–1360. doi: 10.1016/j.jplph.2011.01.005
- Lindström, M. S. (2011). NPM1/B23: a multifunctional chaperone in ribosome biogenesis and chromatin remodeling. *Biochem. Res. Int.* 2011:195209. doi: 10.1155/2011/195209
- Luijsterburg, M. S., White, M. F., van Driel, R., and Dame, R. T. (2008). The major architects of chromatin: architectural proteins in bacteria, archaea and eukaryotes. *Crit. Rev. Biochem. Mol.* 43, 393–418. doi: 10.1080/10409230802528488
- Lundquist, P. K., Davis, J. I., and van Wijk, K. J. (2012). ABC1K atypical kinases in plants: filling the organellar kinase void. *Trends Plant Sci.* 17, 546–555. doi: 10.1016/j.tplants.2012.05.010
- Luo, T., Luo, S., Araújo, W. L., Schlicke, H., Rothbart, M., Yu, J., et al. (2013). Virus-induced gene silencing of pea CHIL and CHLD affects tetrapyrrole biosynthesis, chloroplast development and the primary metabolic network. *Plant Physiol. Biochem.* 65, 17–26. doi: 10.1016/j.plaphy.2013.01.006
- Macvanin, M., and Adhya, S. (2012). Architectural organization in *E. coli* nucleoid. *Biochim. Biophys. Acta* 1819, 830–835. doi: 10.1016/j.bbagen.2012.02.012
- Majeran, W., Friso, G., Asakura, Y., Qu, X., Huang, M., Ponnala, L., et al. (2012). Nucleoid-enriched proteomes in developing plastids and chloroplasts from maize leaves: a new conceptual framework for nucleoid functions. *Plant Physiol.* 158, 156–189. doi: 10.1104/pp.111.188474
- Maréchal, A., and Brisson, N. (2010). Recombination and the maintenance of plant organelle genome stability. *New Phytol.* 186, 299–317. doi: 10.1111/j.1469-8137.2010.03195.x
- Maul, J. E., Lilly, J. W., Cui, L., dePamphilis, C. W., Miller, W., Harris, E. H., et al. (2002). The *Chlamydomonas reinhardtii* plastid chromosome: islands of genes in a sea of repeats. *Plant Cell* 14, 2659–2679. doi: 10.1105/tpc.006155

- Meier, I., Phelan, T., Gruissem, W., Spiker, S., and Schneider, D. (1996). MFP1, a novel plant filament-like protein with affinity for matrix attachment region DNA. *Plant Cell* 8, 2105–2115. doi: 10.1105/tpc.8.11.2105
- Melis, A. (1999). Photodamage-II damage and repair cycle in chloroplasts: what modulates the rate of photodamage *in vivo*? *Trends Plant Sci.* 4, 130–135. doi: 10.1016/S1360-1385(99)01387-4
- Melonek, J., Matros, A., Trösch, M., Mock, H. P., and Krupinska, K. (2012). The core of chloroplast nucleoids contains architectural SWIB domain proteins. *Plant Cell* 24, 3060–3073. doi: 10.1105/tpc.112.099721
- Melonek, J., Mulisch, M., Schmitz-Linneweber, C., Grabowski, E., Hensel, G., and Krupinska, K. (2010). Whirly1 in chloroplasts associates with intron containing RNAs and rarely co-localizes with nucleoids. *Planta* 232, 471–481. doi: 10.1007/s00425-010-1183-0
- Misumi, O., Matsuzaki, M., Nozaki, H., Miyagishima, S. Y., Mori, T., Nishida, K., et al. (2005). *Cyanidioschyzon merolae* genome. A tool for facilitating comparable studies on organelle biogenesis in photosynthetic eukaryotes. *Plant Physiol.* 137, 567–585. doi: 10.1104/pp.104.053991
- Miyamura, S., Nagata, T., and Kuroiwa, T. (1986). Quantitative fluorescence microscopy on dynamic changes of plastid nucleoids during wheat development. *Protoplasma* 133, 66–72. doi: 10.1007/BF01293188
- Mohl, D. A., Easter, J. Jr., and Gober, J. W. (2001). The chromosome partitioning protein, ParB, is required for cytokinesis in *Caulobacter crescentus*. *Mol. Microbiol.* 42, 741–755. doi: 10.1046/j.1365-2958.2001.02643.x
- Moriyama, T., Terasawa, K., and Sato, N. (2011). Conservation of POPs, the plant organellar DNA polymerases, in eukaryotes. *Protist* 162, 177–187. doi: 10.1016/j.protis.2010.06.001
- Murakami, S., Kondo, Y., Nakano, T., and Sato, F. (2000). Protease activity of CND41, a chloroplast nucleoid DNA-binding protein, isolated from cultured tobacco cells. *FEBS Lett.* 468, 15–18. doi: 10.1016/S0014-5793(00)01186-8
- Myouga, F., Hosoda, C., Umezawa, T., Iizumi, H., Kuromori, T., Motohashi, R., et al. (2008). A heterocomplex of iron superoxide dismutases defends chloroplast nucleoids against oxidative stress and is essential for chloroplast development in Arabidopsis. *Plant Cell* 20, 3148–3162. doi: 10.1105/tpc.108.061341
- Nasser, W., Schneider, R., Travers, A., and Muskhelishvili, G. (2001). CRP modulates fts transcription by alternate formation of activating and repressing nucleoprotein complexes. *J. Biol. Chem.* 276, 17878–17886. doi: 10.1074/jbc.M100632200
- Nemoto, Y., Kawano, S., Nakamura, S., Mita, T., Nagata, T., and Kuroiwa, T. (1988). Studies on plastid-nuclei (nucleoids) in *Nicotiana tabacum* L. I. Isolation of proplastid-nuclei from cultured cells and identification of proplastid-nuclear proteins. *Plant Cell Physiol.* 29, 167–177.
- Nielsen, B. L., Cupp, J. D., and Brammer, J. (2010). Mechanisms for maintenance, replication, and repair of the chloroplast genome in plants. *J. Exp. Bot.* 61, 2535–2537. doi: 10.1093/jxb/erq163
- Oldenburg, D. J., and Bendich, A. J. (2004). Changes in the structure of DNA molecules and the amount of DNA per plastid during chloroplast development in maize. *J. Mol. Biol.* 344, 1311–1330. doi: 10.1016/j.jmb.2004.10.001
- Oldenburg, D. J., Rowan, B. A., Kumar, R. A., and Bendich, A. J. (2014). On the fate of plastid DNA molecules during leaf development: response to the Golczyk et al. commentary. *Plant Cell* 26, 855–861. doi: 10.1105/tpc.113.121772
- Orphanides, G., and Reinberg, D. (2000). RNA polymerase II elongation through chromatin. *Nature* 407, 471–475. doi: 10.1038/35035000
- 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. doi: 10.1105/tpc.105.036392
- Pfalz, J., and Pfannschmidt, T. (2013). Essential nucleoid proteins in early chloroplast development. *Trends Plant Sci.* 18, 186–194. doi: 10.1016/j.tplants.2012.11.003
- Pfannschmidt, T., Nilsson, A., and Allen, J. F. (1999). Photosynthetic control of chloroplast gene expression. *Nature* 397, 625–628. doi: 10.1038/17624
- Phinney, B. S., and Thelen, J. J. (2005). Proteomic characterization of a triton-insoluble fraction from chloroplasts defines a novel group of proteins associated with macromolecular structures. *J. Proteome Res.* 4, 497–506. doi: 10.1021/pr049791k
- Postow, L., Hardy, C. D., Arsuaga, J., and Cozzarelli, N. R. (2004). Topological domain structure of the *Escherichia coli* chromosome. *Gene. Dev.* 18, 1766–1779. doi: 10.1101/gad.1207504
- Powikowska, M., Khrouchchova, A., Martens, H. J., Zygallo-Nielsen, A., Melonek, J., Schulz, A., et al. (2014). SVR4 (suppressor of variegation 4) and SVR4-like: two proteins with a role in proper organization of the chloroplast genetic machinery. *Physiol. Plant.* 150, 477–492. doi: 10.1111/pl.12108
- Prikryl, J., Watkins, K. P., Friso, G., van Wijk, K. J., and Barkan, A. (2008). A member of the Whirly family is a multifunctional RNA- and DNA-binding protein that is essential for chloroplast biogenesis. *Nucleic Acids Res.* 36, 5152–5165. doi: 10.1093/nar/gkn492
- Qiao, J., Li, J., Chu, W., and Luo, M. (2013). PRDA1, a novel chloroplast nucleoid protein, is required for early chloroplast development and is involved in the regulation of plastid gene expression in Arabidopsis. *Plant Cell Physiol.* 54, 2071–2084. doi: 10.1093/pcp/pct148
- Qiao, J., Ma, C., Wimmelbacher, M., Börnke, F., and Luo, M. (2011). Two novel proteins, MRL7 and its paralog MRL7-L, have essential but functionally distinct roles in chloroplast development and are involved in plastid gene expression regulation in Arabidopsis. *Plant Cell Physiol.* 52, 1017–1030. doi: 10.1093/pcp/pcr054
- Rauwolf, U., Golczyk, H., Greiner, S., and Herrmann, R. G. (2010). Variable amounts of DNA related to the size of chloroplasts III. Biochemical determinations of DNA amounts per organelle. *Mol. Genet. Genomics* 283, 35–47. doi: 10.1007/s00438-009-0491-1
- Raynaud, C., Sozzani, R., Glab, N., Domenichini, S., Perennes, C., Cella, R., et al. (2006). Two cell-cycle regulated SET-domain proteins interact with proliferating cell nuclear antigen (PCNA) in Arabidopsis. *Plant J.* 47, 395–407. doi: 10.1111/j.1365-313X.2006.02799.x
- Rimsky, S., and Travers, A. (2011). Pervasive regulation of nucleoid structure and function by nucleoid-associated proteins. *Curr. Opin. Microbiol.* 14, 136–141. doi: 10.1016/j.mib.2011.01.003
- Ris, H., and Plaut, W. (1962). Ultrastructure of DNA-containing areas in the chloroplast of Chlamydomonas. *J. Cell Biol.* 13, 383–391. doi: 10.1083/jcb.13.3.383
- Robinow, C., and Kellenberger, E. (1994). The bacterial nucleoid revisited. *Microbiol. Rev.* 58, 211–232.
- Rochman, M., Malicet, C., and Bustin, M. (2010). HMGN5/NSBP1: a new member of the HMGN protein family that affects chromatin structure and function. *Biochim. Biophys. Acta* 1799, 86–92. doi: 10.1016/j.bbagen.2009.09.012
- Rowan, B. A., Oldenburg, D. J., and Bendich, A. J. (2004). The demise of chloroplast DNA in Arabidopsis. *Curr. Genet.* 46, 176–181. doi: 10.1007/s00294-004-0515-7
- Rowan, B. A., Oldenburg, D. J., and Bendich, A. J. (2010). RecA maintains the integrity of chloroplast DNA molecules in Arabidopsis. *J. Exp. Bot.* 61, 2575–2588. doi: 10.1093/jxb/erq088
- Sakai, A., Takano, H., and Kuroiwa, T. (2004). Organelle nuclei in higher plants: structure, composition, function, and evolution. *Int. Rev. Cytol.* 238, 59–118. doi: 10.1016/S0074-7696(04)38002-2
- Sakamoto, W., Uno, Y., Zhang, Q., Miura, E., Kato, Y., and Sodmergen (2009). Arrested differentiation of proplastids into chloroplasts in variegated leaves characterized by plastid ultrastructure and nucleoid morphology. *Plant Cell Physiol.* 50, 2069–2083. doi: 10.1093/pcp/pcp127
- Sala-Rovira, M., Geraud, M. L., Caput, D., Jacques, F., Soyergobillard, M. O., Vernet, G., et al. (1991). Molecular cloning and immunolocalization of two variants of the major basic nuclear protein (HCc) from the histone-less eukaryote *Cryptothecodium cohnii* (Pyrrhophyta). *Chromosoma* 100, 510–518. doi: 10.1007/BF00352201
- Salvador, M. L., Klein, U., and Bogorad, L. (1998). Endogenous fluctuations of DNA topology in the chloroplast of *Chlamydomonas reinhardtii*. *Mol. Cell Biol.* 18, 7235–7242.
- Sato, N. (2001). Was the evolution of plastid genetic machinery discontinuous? *Trends Plant Sci.* 6, 151–155. doi: 10.1016/S1360-1385(01)01888-X
- Sato, N., Albrieux, C., Joyard, J., Douce, R., and Kuroiwa, T. (1993). Detection and characterization of a plastid envelope DNA binding protein which may anchor plastids nucleoids. *EMBO J.* 12, 555–561.
- Sato, N., Ohshima, K., Watanabe, A., Ohta, N., Nishiyama, Y., Joyard, J., et al. (1998). Molecular characterization of the PEND protein, a novel bZIP protein present in the envelope membrane that is the site of nucleoid replication in developing plastids. *Plant Cell* 10, 859–872. doi: 10.1105/tpc.10.5.859
- Sato, N., and Ohta, N. (2001). DNA-binding specificity and dimerization of the DNA-binding domain of the PEND protein in the chloroplast envelope membrane. *Nucleic Acids Res.* 29, 2244–2250. doi: 10.1093/nar/29.11.2244

- Sato, N., Terasawa, K., Miyajima, K., and Kabeya, Y. (2003). Organization, developmental dynamics, and evolution of plastid nucleoids. *Int. Rev. Cytol.* 232, 217–262. doi: 10.1016/S0074-7696(03)32006-6
- 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. doi: 10.1093/dnares/6.5.283
- Schröter, Y., Steiner, S., Matthäi, K., and Pfannschmidt, T. (2010). Analysis of oligomeric protein complexes in the chloroplast sub-proteome of nucleic acid-binding proteins from mustard reveals potential redox regulators of plastid gene expression. *Proteomics* 10, 2191–2204. doi: 10.1002/pmic.200900678
- Sekine, K., Fujiwara, M., Nakayama, M., Takao, T., Hase, T., and Sato, N. (2007). DNA binding and partial nucleoid localization of the chloroplast stromal enzyme ferredoxin:sulfite reductase. *FEBS J.* 274, 2054–2069. doi: 10.1111/j.1742-4658.2007.05748.x
- Sekine, K., Hase, T., and Sato, N. (2002). Reversible DNA compaction by sulfite reductase regulates transcriptional activity of chloroplast nucleoids. *J. Biol. Chem.* 277, 24399–24404. doi: 10.1074/jbc.M201714200
- Sekine, K., Sakakibara, Y., Hase, T., and Sato, N. (2009). A novel variant of ferredoxin-dependent sulfite reductase having preferred substrate specificity for nitrite in the unicellular red alga *Cyanidioschyzon merolae*. *Biochem. J.* 423, 91–98. doi: 10.1042/BJ20090581
- Steiner, S., Schröter, Y., Pfalz, J., and Pfannschmidt, T. (2011). Identification of essential subunits in the plastid-encoded RNA polymerase complex reveals building blocks for proper plastid development. *Plant Physiol.* 157, 1043–1055. doi: 10.1104/pp.111.184515
- Ströher, E., and Dietz, K. J. (2008). The dynamic thiol-disulphide redox proteome of the *Arabidopsis thaliana* chloroplast as revealed by differential electrophoretic mobility. *Physiol. Plant.* 133, 566–583. doi: 10.1111/j.1399-3054.2008.01103.x
- Studitsky, V. M. (2005). Chromatin remodeling by RNA-polymerase II. *Mol. Biol.* 39, 639–654. doi: 10.1007/s11008-005-0071-3
- Terasawa, K., and Sato, N. (2005). Occurrence and characterization of PEND proteins in angiosperms. *J. Plant Res.* 118, 111–119. doi: 10.1007/s10265-005-0200-z
- Terasawa, K., and Sato, N. (2009). Plastid localization of the PEND protein is mediated by a noncanonical transit peptide. *FEBS J.* 276, 1709–1719. doi: 10.1111/j.1742-4658.2009.06901.x
- Timmis, J. N., Ayliffe, M. A., Huang, C. Y., and Martin, W. (2004). Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. *Nat. Rev. Genet.* 5, 123–135. doi: 10.1038/nrg1271
- Travers, A., and Muskhelishvili, G. (2005). Bacterial chromatin. *Curr. Opin. Genet. Dev.* 5, 507–514. doi: 10.1016/j.gde.2005.08.006
- Udy, D. B., Belcher, S., Williams-Carrier, R., Gualberto, J. M., and Barkan, A. (2012). Effects of reduced chloroplast gene copy number on chloroplast gene expression in maize. *Plant Physiol.* 160, 1420–1431. doi: 10.1104/pp.112.204198
- Vora, T., Hottes, A. K., and Tavazoie, S. (2009). Protein occupancy landscape of a bacterial genome. *Mol. Cell* 35, 247–253. doi: 10.1016/j.molcel.2009.06.035
- Weber, P., Fulgosi, H., Piven, I., Müller, L., Krupinska, K., Duong, V. H., et al. (2006). TCP34, a nuclear-encoded response regulator-like TPR protein of higher plant chloroplasts. *J. Mol. Biol.* 357, 535–549. doi: 10.1016/j.jmb.2005.12.079
- Wicke, S., Schneeweiss, G. M., dePamphilis, C. W., Müller, K. F., and Quandt, D. (2011). The evolution of the plastid chromosome in land plants: gene content, gene order, gene function. *Plant Mol. Biol.* 76, 273–297. doi: 10.1007/s11103-011-9762-4
- Wiedemann, G., Hermßen, C., Melzer, M., Büttner-Mainik, A., Rennenberg, H., Reski, R., et al. (2010). Targeted knock-out of a gene encoding sulfite reductase in the moss *Physcomitrella patens* affects gametophytic and sporophytic development. *FEBS Lett.* 584, 2271–2278. doi: 10.1016/j.febslet.2010.03.034
- Wolf, S. G., Frenkel, D., Arad, T., Finkel, S. E., Kolter, R., and Minsky, A. (1999). DNA protection by stress-induced biocrystallization. *Nature* 400, 83–85. doi: 10.1038/21918
- Wong, J. T., New, D. C., Wong, J. C., and Hung, V. K. (2003). Histone-like proteins of the dinoflagellate *Cryptocodinium cohnii* have homologies to bacterial DNA-binding proteins. *Eukaryot. Cell* 2, 646–650. doi: 10.1128/EC.2.3.646-650.2003
- Wycliffe, P., Sitbon, F., Wernersson, J., Ezcurra, I., Ellerström, M., and Rask, L. (2005). Continuous expression in tobacco leaves of a *Brassica napus* PEND homologue blocks differentiation of plastids and development of palisade cells. *Plant J.* 44, 1–15. doi: 10.1111/j.1365-313X.2005.02482.x
- Xiong, J. Y., Lai, C. X., Qu, Z., Yang, X. Y., Qin, X. H., and Liu, G. Q. (2009). Recruitment of AtWHY1 and AtWHY3 by a distal element upstream of the kinesin gene AtKP1 to mediate transcriptional repression. *Plant Mol. Biol.* 71, 437–449. doi: 10.1007/s11103-009-9533-7
- Yagi, Y., and Shiina, T. (2014). Recent advances in the study of chloroplast gene expression and its evolution. *Front. Plant Sci.* 5:61. doi: 10.3389/fpls.2014.00061
- Yu, F., Park, S. S., Liu, X., Foudree, A., Fu, A., Powikowska, M., et al. (2011). SUPPRESSOR OF VARIEGATION4, a new var2 suppressor locus, encodes a pioneer protein that is required for chloroplast biogenesis. *Mol. Plant* 4, 229–240. doi: 10.1093/mp/sqz074
- Yua, Q. B., Ma, Q., Kong, M. M., Zhao, T. T., Zhang, X. L., Zhou, Q., et al. (2013). AtECB1/MRL7, a thioredoxin-like fold protein with disulfide reductase activity, regulates chloroplast gene expression and chloroplast biogenesis in *Arabidopsis thaliana*. *Mol. Plant* 7, 206–217. doi: 10.1093/mp/sqz074
- Zaitlin, D., Hu, J., and Bogorad, L. (1989). Binding and transcription of relaxed DNA templates by fractions of maize chloroplast extracts. *Proc. Natl. Acad. Sci. U.S.A.* 86, 876–880. doi: 10.1073/pnas.86.3.876
- 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. doi: 10.1111/j.1365-313X.2007.03084.x
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Received:** 15 May 2014; **accepted:** 08 August 2014; **published online:** 04 September 2014.
- Citation:** Powikowska M, Oetke S, Jensen PE and Krupinska K (2014) Dynamic composition, shaping and organization of plastid nucleoids. *Front. Plant Sci.* 5:424. doi: 10.3389/fpls.2014.00424
- This article was submitted to Plant Physiology, a section of the journal Frontiers in Plant Science.*
- Copyright © 2014 Powikowska, Oetke, Jensen and Krupinska. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*



# WHIRLY1 is a major organizer of chloroplast nucleoids

Karin Krupinska<sup>1\*</sup>, Svenja Oetke<sup>1</sup>, Christine Desel<sup>1</sup>, Maria Mulisch<sup>1,2</sup>, Anke Schäfer<sup>1</sup>,  
Julien Hollmann<sup>1</sup>, Jochen Kumlehn<sup>3</sup> and Götz Hensel<sup>3</sup>

<sup>1</sup> Institute of Botany, Christian-Albrechts-University of Kiel, Kiel, Germany

<sup>2</sup> Central Microscopy of the Center of Biology, Christian-Albrechts-University of Kiel, Kiel, Germany

<sup>3</sup> Plant Reproductive Biology, Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Stadt Seeland/OT Gatersleben, Germany

**Edited by:**

Jeannette Pfalz,  
Friedrich-Schiller-University Jena,  
Germany

**Reviewed by:**

Shin-ya Miyagishima, National  
Institute of Genetics, Japan  
Naoki Sato, University of Tokyo,  
Japan

**\*Correspondence:**

Karin Krupinska, Institute of Botany,  
Christian-Albrechts-University of  
Kiel, Olshausenstrasse 40,  
24098 Kiel, Germany  
e-mail: kkrupinska@bot.uni-kiel.de

WHIRLY1 is an abundant protein of chloroplast nucleoids, which has also been named pTAC-1 with regard to its detection in the proteome of transcriptionally active chromosomes (TAC). In barley primary foliage leaves, expression of the WHIRLY1 gene is highest at the base whereas protein accumulation is highest in the middle of the leaf where young developing chloroplasts are found. In order to elucidate the function of WHIRLY1 in chloroplast nucleoids, transgenic barley plants with an RNAi-mediated knock-down of the *HvWHIRLY1* gene (RNAi-W1) were generated. The homozygous RNAi-W1-7 plants, barely containing traces of the WHIRLY1 protein, were chosen for detailed analyses of nucleoids. Nucleic acid specific-staining with YO-PRO®-1 revealed that in comparison to wild type chloroplasts, which have multiple small nucleoids attached to thylakoids, chloroplasts of the transgenic plants contain large irregularly formed patches of DNA besides nucleoids that are similar in size and shape to those of wild type chloroplasts. In large electron lucent areas, filamentous structures were detected by conventional transmission electron microscopy. Analyses of ptDNA levels by both DNA dot-blot hybridization and quantitative PCR showed that leaves of the transgenic plants have a two- to three-fold higher level of ptDNA than the wild type. The higher ptDNA level in RNAi-W1 plants coincided with an enhanced expression of the gene encoding a putative organelle targeted DNA polymerase in the mid part of primary foliage leaves. Furthermore, overexpression of the barley WHIRLY1 gene in *E. coli* cells revealed a higher compaction of bacterial nucleoids. These results suggest that WHIRLY1 belongs to the group of plastid nucleoid associated proteins (ptNAP) having a function in compacting a subpopulation of chloroplast nucleoids thereby affecting DNA replication.

**Keywords:** DNA compaction, plastid DNA, plastid nucleoid, replication, WHIRLY1

## INTRODUCTION

WHIRLY1 belongs to a small family of single-stranded DNA (ssDNA) binding proteins, which contains two members in most plants such as barley, whereas *Arabidopsis thaliana* has three WHIRLY proteins. WHIRLY1 is a chloroplast-nucleus located protein (Grabowski et al., 2008; Maréchal et al., 2009), which was first detected as a nuclear transcriptional regulator (Desveaux et al., 2000). Intriguingly, the precursor of mature WHIRLY1 has an N-terminal transit peptide for import into chloroplasts whereas WHIRLY2 is imported into mitochondria (Krause et al., 2005). In *A. thaliana* WHIRLY1 has been found together with WHIRLY3 in the proteome of the transcriptionally active chromosome (TAC), which is the transcriptionally active fraction of the nucleoids (Pfalz et al., 2006). Nucleoids are particles consisting of multiple copies of highly condensed ptDNA, RNA, and a number of different proteins (Sakai et al., 2004; Powikrowska et al., 2014b). The association of WHIRLY1 with plastid nucleoids has been confirmed in barley and maize (Melonek et al., 2010; Majeran et al., 2012). WHIRLY1 was found to bind to ptDNA in an unspecific manner (Prikryl et al., 2008; Maréchal et al., 2009) and

also to selected plastid RNAs including the *atpF* mRNA (Prikryl et al., 2008; Melonek et al., 2010). Maize mutants with severely reduced levels of the WHIRLY1 protein are impaired in chloroplast development due to greatly diminished levels of ribosomal RNA (Prikryl et al., 2008). In contrast to the maize mutants, barley plants with an RNAi-mediated knock-down of the *WHIRLY1* gene showed no obvious phenotype under standard growth conditions (Melonek et al., 2010). The *Arabidopsis* mutant *why1why3* lacking both plastid located WHIRLY proteins was shown to have variegated green/white/yellow leaves in 5% of the progeny. In such leaves ptDNA molecules with aberrations resulting from illegitimate recombination were detected (Maréchal et al., 2009), indicating that WHIRLY proteins have a function in repair of organelle DNA (Maréchal and Brisson, 2010). Plants resulting from a cross between the *Arabidopsis* double mutant *why1why3* and a mutant impaired in organelle DNA polymerase IB (*polIB*) had a more severe phenotype and increased DNA rearrangements than the *why1why3* mutant suggesting that DNA polymerase IB and WHIRLY proteins act synergistically in maintenance of plastid genome stability (Parent et al., 2011; Lepage et al., 2013).

The diversity in phenotype between maize mutants and the *why1why3* mutant was proposed to show that WHIRLY proteins can serve different purposes depending on the conditions and/or plant species (Maréchal et al., 2009). Prikryl et al. (2008) suggested that WHIRLY1 could play a similar role in plastids as the versatile nucleoid associated HU protein in bacteria. Parent et al. (2011) suggested that WHIRLY proteins might function like the major ssDNA binding protein SSB in bacteria, which affects many nucleoid associated processes by interacting with different proteins involved in DNA transaction processes, such as DNA polymerases and gyrases (Shereda et al., 2008).

In maize *why1* mutants chloroplast development is blocked. Barley RNAi-W1 plants with reduced levels of *WHIRLY1* in contrast do not show obvious phenotypes when grown under standard conditions (Melonek et al., 2010). Making use of the basipetal developmental gradient of barley leaves, in this study expression of the *WHIRLY1* gene was shown to be highest in immature cells at the leaf base as described for the expression of the *SUPPRESSOR OF VARIEGATION 4* gene (*SVR4*). This contrasts with the increase in the accumulation of the WHIRLY1 protein, which is highest in cells containing developing chloroplasts. Microscopic analyses showed that the WHIRLY1 protein compacts the DNA of a subpopulation of plastid nucleoids. A reduced compactness of nucleoids in chloroplasts of the RNAi-W1 plants correlates with an elevated level of plastid DNA and enhanced expression of the gene encoding a putative *BARLEY ORGANELLE DNA POLYMERASE* (*HvPoII-like*). In addition, *E. coli* cells overexpressing the barley *WHIRLY1* gene showed a reduced growth and contained highly condensed nucleoids. The results of these studies indicate that WHIRLY1 is involved in compaction and organization of ptDNA having consequences for replication.

## MATERIALS AND METHODS

### PLANT MATERIAL

For generation of transgenic barley plants with an RNAi-mediated knock-down of the *HvWHIRLY1* gene, the 198 bp *HvWHIRLY1* cDNA region (nucleotide -302 to -105 upstream of TAA stop codon of *HvWHIRLY1* gene) was amplified by PCR with specific primers (Supplementary Table 1), cloned into the pENTR/TOPO gateway vector (Invitrogen, Karlsruhe, Germany) and sequenced to verify the sequences of the PCR products. The *HvWHIRLY1* cDNA-fragment of the respective entry vector was transferred to the pIPKb007 binary vector using Gateway™ LR clonase mix (Invitrogen, Karlsruhe, Germany) to generate the binary vector pgH235 essentially as described elsewhere (Himmelbach et al., 2007). The transformation of immature embryos of barley cv. "Golden Promise" by *Agrobacterium tumefaciens* was performed as described by Hensel et al. (2008). Plantlets with resistance toward hygromycin were transferred into soil and cultivated in a greenhouse. Additionally, PCR with primers (Supplementary Table 1) for the hygromycin resistance cassette was performed to verify the transgene integration. For selection of homozygous plants, barley (*Hordeum vulgare* L. cv. "Golden Promise") plants were grown in a glasshouse with additional light supply. For the microscopic and immunological studies barley seedlings were sown in multipots on soil (Einheitserde

ED73, Einheitswerk Werner Tantau, Uetersen, Germany). After 3 days in darkness and low temperature (6°C), the seedlings were transferred to a chamber with 21–25°C and continuous light of 50–100 μmol photons s<sup>-1</sup> m<sup>-2</sup>. For protein extraction, RNA isolation and chlorophyll analysis, leaf sections were taken from primary foliage leaves of 7 days old seedlings. Ten days after sowing, primary foliage leaves were used for preparation of total genomic DNA and cut into sections for microscopic analyses.

### ANALYSIS OF CHLOROPHYLL CONTENT

Defined segments (area: 0.5–0.9 cm<sup>2</sup>) were excised from the base, mid, and tip of primary foliage leaves and were immediately frozen in liquid nitrogen. Until analysis by HPLC the samples were stored in a freezer at -80°C. For extraction, the leaf segments along with five glass beads were ground in the frozen state in a Geno Grinder (Type 2000, SPEX, CertiPrep, Munich, Germany) with 0.5 ml 80% (v/v) acetone buffered with 20 mM Tris, pH 7.8. After centrifugation, the pellet was extracted twice with 200 μl 100% acetone. From the combined extracts, 50 μl were used for HPLC analysis on an Agilent 1100 system (Agilent, Waldbronn, Germany) with DAD detection. The protocol was the same as published before (Niinemets et al., 1998).

### DETERMINATION OF mRNA LEVELS BY qRT-PCR

RNA was extracted from leaf sections taken from primary foliage leaves with peqGOLD-TriFast reagent (Peqlab Biotechnologie, Erlangen, Germany), and was used for cDNA synthesis employing QuantiTect® Reverse Transcriptase Kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol. Quantitative real time PCR (qRT-PCR) analyses were performed with the QuantiFast SYBR Green PCR Kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol using gene specific primers (Supplementary Table 1). Data analysis was accomplished by the Rotor-Gene Q software (version 2.0.2.4) (Qiagen, Hilden, Germany). Relative quantification of transcript levels was performed using the "Delta-delta C<sub>T</sub> method" as presented by PE Applied Biosystems (Perkin Elmer, Foster City, CA, USA). Data were normalized to the 18S rRNA.

### IMMUNOBLOT ANALYSIS

Proteins were extracted from leaf sections with a buffer consisting of 62.5 mM Tris, pH 6.8, 10% (v/v) glycerol, 1% (w/v) SDS, and 5% (v/v) β-mercaptoethanol. Equal amounts of proteins (15 μg) were subjected to SDS-PAGE on 14% (w/v) polyacrylamide gels containing a high concentration of Tris (Fling and Gregerson, 1986). Proteins were transferred to nitrocellulose by semi-dry electroblotting and treated as described (Humbeck et al., 1996). Immunoreactions were detected by chemoluminescence using different kits (GE Healthcare, Buckinghamshire, UK; Thermo Scientific, Waltham, MA, USA; Lumigen, Southfield, MI, USA). For immunological detection of WHIRLY1, the antibody directed toward peptide 2 was used (Grabowski et al., 2008). SVR4 was detected by the antibody provided by P. E. Jensen (University of Copenhagen, Denmark) (Powikrowska et al., 2014a).

### DNA GEL BLOT ANALYSIS

DNA was extracted from homozygous leaf material according to the method of Palotta et al. (2000). At least 25 μg genomic DNA

was digested either with *Hind*III or *Eco*RV, cutting the T-DNA only once, respectively. After electrophoresis DNA was transferred onto a Hybond-N+ nylon membrane (Amersham GE Healthcare, Buckinghamshire, UK) according to the manufacturer's instructions, and hybridized with digoxigenin-dUTP (DIG-11-dUTP) labeled DNA probes, as recommended by the supplier (Roche, Mannheim, Germany). To generate the DNA hybridization probes, primers used for PCR confirmation described above were used.

#### STAINING OF NUCLEOIDS WITH YO-PRO®-1

For staining of nucleoids with YO-PRO®-1 Iodide (491/509) (Molecular Probes, Life Technologies, Carlsbad, CA, USA), cross-sections excised 2–2.5 cm below the tip of primary foliage leaves were fixed overnight in a 4% (w/v) solution of formaldehyde (freshly prepared from paraformaldehyde). After three washing steps with 2× SSC (0.3 M NaCl; 30 mM sodium citrate, pH 7.0) the sections were treated with DNase-free ribonuclease A (20 µg ml<sup>-1</sup> in 2× SSC) for 1 h at 37°C. After washing with 2× SSC sections were stained with 0.5 µM YO-PRO®-1 Iodide for 15 min at room temperature. After washing with 2× SSC the segments were embedded in a solution consisting of 50% (v/v) glycerol and 1× SSC. Microscopy was performed with a confocal laser-scanning microscope (Leica TCS SP5, Leica Microsystems, Wetzlar, Germany; with LAS AF –Software, 63× 1.2 water objective HCX PLAPO). Fluorescence was excited at 488 nm (10%) using an argon laser or at 633 nm (12%) using a HeNe laser. Sequential scanning was done at emissions of 500–550 and 650–750 nm. The diameters of fluorescence signals were measured with the quantification module of the Leica software LAS AF-TCS.

#### TRANSMISSION ELECTRON MICROSCOPY

Leaf segments from primary foliage leaves (2 × 2 mm) at a position of 2 cm below the leaf tip were fixed at room temperature in 2.5% (v/v) glutardialdehyde and 1% (w/v) formaldehyde (freshly prepared from paraformaldehyde) in 0.1 M sodium cacodylate buffer, pH 7.3. After washing in buffer, the samples were post-fixed in buffered 1% (w/v) osmium tetroxide, washed, dehydrated in a graded series of ethanol, and embedded in LR white resin. The resin was polymerized at 60°C. Ultrathin sections were cut with a diamond knife in an Ultracut UCT ultramicrotome (Leica Microsystems, Wetzlar, Germany). The sections were stained with saturated uranyl acetate in water and lead citrate (Reynolds, 1963) and observed using a Philips CM10 transmission electron microscope (FEI, Eindhoven, The Netherlands).

#### HETEROLOGOUS EXPRESSION OF THE WHIRLY1 GENE AND DNA CONDENSATION ASSAYS IN *ESCHERICHIA COLI* CELLS

The coding sequence of the barley *WHIRLY1* gene (AK365452) except the sequence encoding the plastid transit peptide was cloned into the pASK-IBA3 vector (IBA Life Science, MO, USA). For induction of overexpression anhydrotetracycline was added at OD<sub>600</sub> 0.7–1.0 to a final concentration of 200 µg l<sup>-1</sup>. Staining of cells with 4',6-diamidino-2-phenylindole (DAPI) was performed as described in Melonek et al. (2012) and cells were observed by fluorescence microscopy with a Zeiss Axiophot microscope (Carl Zeiss, Oberkochen, Germany).

#### DETERMINATION OF RELATIVE ptDNA LEVELS

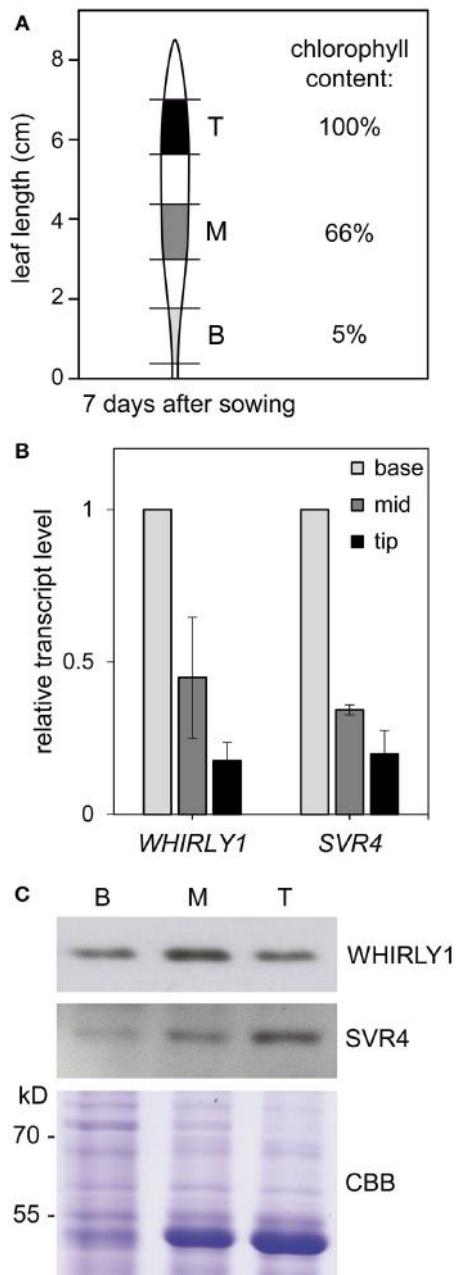
Total genomic DNA was extracted from primary foliage leaves of 10 days old barley plants and leaf sections as described (Fulton et al., 1995). For DNA dot-blot analyses, different DNA dilutions were prepared and supplied with the same volume of 4× SSC. After denaturation, DNA was transferred onto a nylon membrane (Hybond-N+, Amersham GE Healthcare, Buckinghamshire, UK) using a dot-blot device (SRC 96D Minifold I, Schleicher & Schuell, Dassel, Germany). The amplified fragments specific for either nuclear 18S rDNA or plastid *petD* were used as templates for DIG-DNA labeling (digoxigenin) using a kit (DIG High Prime DNA Labeling and Detection Starter Kit II, Roche Applied Science, Mannheim, Germany) according to the manufacturer's protocol. Primers used for amplification of templates are listed in Supplementary Table 1.

For q-PCR analyses a QuantiFast SYBR Green PCR Kit (Qiagen, Hilden, Germany) was used according to the manufacturer's protocol using gene specific primers (Supplementary Table 1). Each reaction was repeated at least three times. Data analysis and relative quantification of genomic DNA levels was performed as described in Determination of mRNA levels by qRT-PCR. Data were normalized to the 18S rDNA gene. The level of *RBCS* genes was used as reference for nuclear DNA content.

## RESULTS

#### *WHIRLY1* GENE EXPRESSION AND *WHIRLY1* PROTEIN ACCUMULATION IN BARLEY LEAVES

*WHIRLY1* gene expression was analyzed by qRT-PCR during chloroplast development using RNA extracted from three sections excised at different positions of the barley primary foliage leaf (Figure 1). Chlorophyll content of the sections from the leaf tips (T) was about 20 times higher than in the basal sections (B) (Figure 1A). The chlorophyll content of the mid-section (M) was 66% of the chlorophyll of the upper section (section T). Expression of the *WHIRLY1* gene is highest at the leaf base (section B) and decreases to a level of about 20% in sections from the leaf tips (section T) (Figure 1B). For comparison, expression of the gene encoding *SVR4* was analyzed. *SVR4* has recently been proposed to be essential for nucleoid reorganization during chloroplast development and for transcription by plastid encoded RNA polymerase (PEP) (Powikrowska et al., 2014a). The developmental changes in expression of the *HvSVR4* gene closely follow the changes in *HvWHIRLY1* expression, which is in accordance with a role of *WHIRLY1* in DNA transaction processes required for early chloroplast development. Immunological analysis with total protein extracts derived from the same leaf sections showed that the development dependent changes in protein levels of *WHIRLY1* as well as *SVR4* do not parallel the changes in mRNA levels (Figure 1C). Accumulation of the *WHIRLY1* protein is highest in section M having 66% of the chlorophyll content of the upper section containing mature chloroplasts (section T). While the level of *SVR4* steadily increased during chloroplast development, the level of *WHIRLY1* declined during maturation of chloroplasts as already observed by Grabowski et al. (2008). This discrepancy might indicate that *WHIRLY1* and *SVR4* play roles in different DNA related processes connected with chloroplast development.



**FIGURE 1 | Development dependent changes in mRNA level and protein accumulation analyzed in different sections of barley primary foliage leaves. (A)** Leaf sections designated base (B), mid (M), and tip (T) were excised from wild type barley primary foliage leaves as indicated. The chlorophyll content of the tip was set to 100%. The chlorophyll content of the base and mid is presented relative to the chlorophyll content of the tip. **(B)** *HvWHIRLY1* gene expression in the leaf sections of 7 days old primary foliage leaves was compared to expression of the *HvSVR4* gene. qRT-PCR was performed with specific primers (Supplementary Table 1). Relative quantification of transcript levels was performed using the “Delta-delta  $C_T$  method.” Data were normalized to the 18S rRNA and data for the base (B) were set to 1. The data for mid and tip are shown relative to the base. **(C)** Immunological detection of *HvWHIRLY1* and *HvSVR4* in total protein extracts isolated from leaf sections of 7 days old primary foliage leaves. Specific antibodies directed against *HvWHIRLY1* and *HvSVR4* were used. For comparison, a part of the Coomassie Brilliant Blue (CBB) stained gel is shown.

### RNAi MEDIATED KNOCK-DOWN OF THE *WHIRLY1* GENE IN BARLEY

To investigate the function of *WHIRLY1*, transgenic barley plants with a knock-down of the *HvWHIRLY1* gene were generated using an RNAi-hairpin construct (Figure 2A). Thirty hygromycin resistant RNAi-W1 plants were tested by PCR using primers specific for the two hairpin repeats (Supplementary Figure 1A). Fifteen plants carried both inverted repeats while two plants (RNAi-W1-2, RNAi-W1-10) carried the antisense repeat only (Supplementary Figure 1A). Leaf material collected from 15 T<sub>1</sub> progeny was tested for the knock-down effect at the level of the *WHIRLY1* mRNA and at the level of protein accumulation (Supplementary Figure 1B). Compared to the wild type, the *HvWHIRLY1* mRNA level was reduced in eight progeny with RNAi-W1-1, -6, -7, -8, -9, -20, and -26 showing the strongest knock-down effects (Supplementary Figure 1B). Immunoblot analysis showed that in most progeny with reduced levels of mRNA, the protein was almost undetectable (Supplementary Figure 1B).

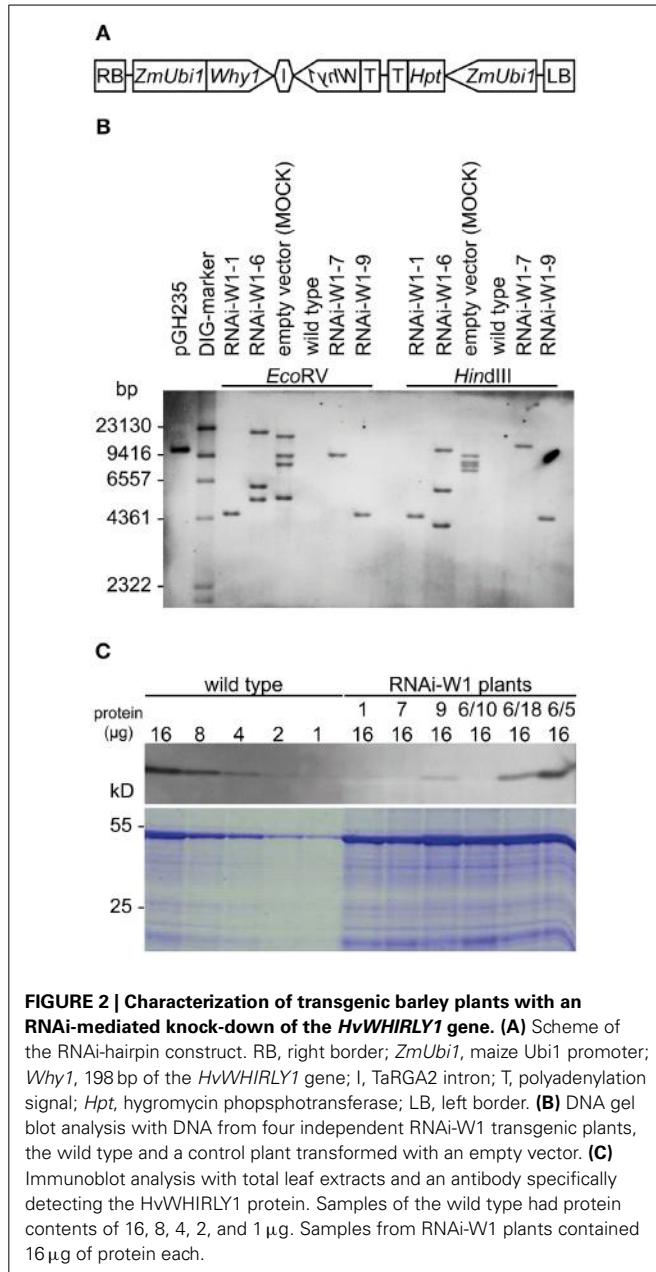
Four progeny were used for DNA gel blot analysis. Digestion with *Hind*III and *Eco*RV showed that most RNAi plants have independent insertions of the transgene (Figure 2B). Although all plants have been selected from different embryo-derived calli and were therefore considered to be independent, RNAi-W1-1 and RNAi-W1-9 show the same integration patterns (Figure 2B). Only RNAi-W1-1, -7, and -9 contain one transgene copy and were considered homozygous by resistance tests and PCR assays. The T<sub>4</sub> progeny of RNAi-W1-6 was observed to be still heterozygous.

*WHIRLY1* protein accumulation in the RNAi-W1 plants was determined with powdered material from primary foliage leaves of seedlings 10 days after sowing using an antibody specific for *HvWHIRLY1* (Grabowski et al., 2008). The signal obtained with 16 µg of total protein extracted from primary foliage leaves was compared to the signals obtained with different amounts of protein (1–16 µg) extracted from wild type leaves of the same developmental stage. The *WHIRLY1* protein was almost undetectable in RNAi-W1-7 plants and did not exceed 10% of the wild type in RNAi-W1-1 and RNAi-W1-9 plants (Figure 2C). In case of the heterozygous RNAi-W1-6 plants, protein was extracted from individual leaves. Whereas in some of these samples the abundance of *WHIRLY1* was as in the wild type, others had a reduced content of the protein (Figure 2C).

### MICROSCOPIC ANALYSES OF NUCLEOID MORPHOLOGY

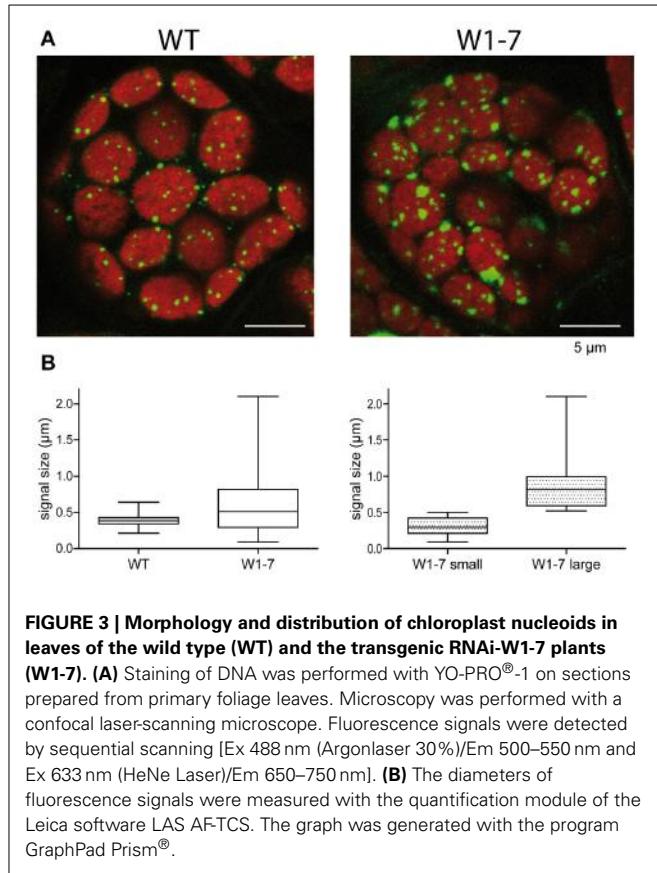
*WHIRLY1* is a major protein of chloroplast nucleoids (Pfalz et al., 2006; Melonek et al., 2010). To investigate whether a knock-down of the *WHIRLY1* gene has an impact on size, shape, and distribution of the nucleoids in chloroplasts, tip sections of primary foliage leaves of RNAi-W1 seedlings were chosen for microscopic analyses of nucleoids.

To compare the nucleoids in chloroplasts of transgenic plants with those of wild type chloroplasts, sections from primary foliage leaves were fixed by formaldehyde and were stained with the fluorescent nucleic acid-specific dye YO-PRO®-1. This analysis revealed that the nucleoid population in chloroplasts of the RNAi-W1-7 plants is much more heterogeneous than the nucleoid population of control chloroplasts (Figure 3A). Whereas the fluorescence signals from nucleoids in the wild type chloroplasts have a mean diameter of 300 nm, those of the transgenic plants



**FIGURE 2 | Characterization of transgenic barley plants with an RNAi-mediated knock-down of the *HvWHIRLY1* gene. (A)** Scheme of the RNAi-hairpin construct. RB, right border; *ZmUbi1*, maize *Ubi1* promoter; *Whirly1*, 198 bp of the *HvWHIRLY1* gene; I, TaRGA2 intron; T, polyadenylation signal; Hpt, hygromycin phosphotransferase; LB, left border. **(B)** DNA gel blot analysis with DNA from four independent RNAi-W1 transgenic plants, the wild type and a control plant transformed with an empty vector. **(C)** Immunoblot analysis with total leaf extracts and an antibody specifically detecting the *HvWHIRLY1* protein. Samples of the wild type had protein contents of 16, 8, 4, 2, and 1 μg. Samples from RNAi-W1 plants contained 16 μg of protein each.

have a mean diameter of 700 nm (Figure 3B, left panel). The nucleoids in the chloroplasts of transgenic plants can be subdivided in two populations of different sizes: small round nucleoids having a signal diameter of 300 nm as those of the wild type, and large irregularly formed nucleoids with a mean signal diameter of 800 nm (Figure 3B, right panel). The sizes and shapes vary considerably in the second population with signal sizes ranging from 500 nm to 2 μm. The changes in nucleoid morphology of chloroplasts in comparable sections from RNAi-W1-1 primary foliage leaves were less pronounced than in leaves of RNAi-W1-7. This suggests that even a low amount of WHIRLY1 is sufficient for compaction of nucleoids. Transmission electron microscopy confirmed the heterogeneity in size and shape of nucleoids in RNAi-W1-7. Nucleoids of 200–300 nm in diameter are found besides



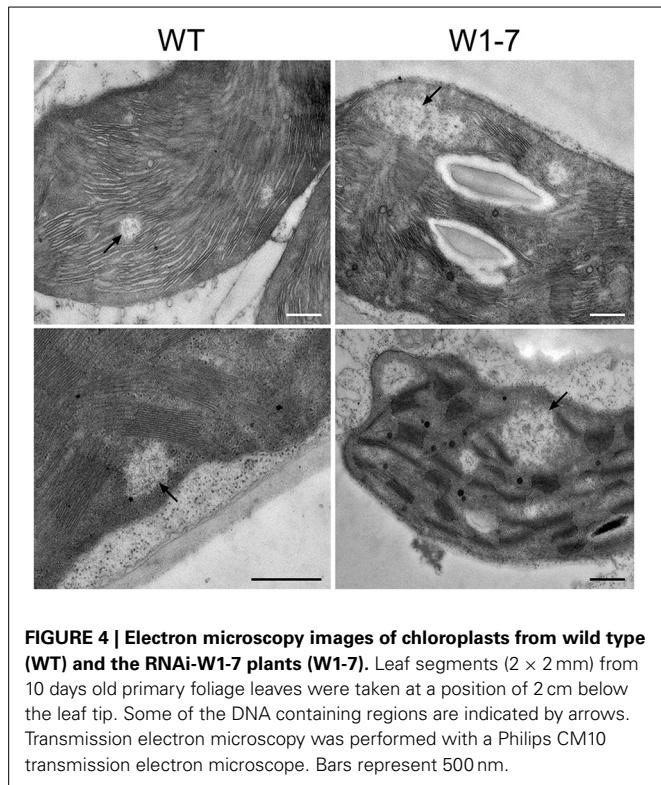
**FIGURE 3 | Morphology and distribution of chloroplast nucleoids in leaves of the wild type (WT) and the transgenic RNAi-W1-7 plants (W1-7).** **(A)** Staining of DNA was performed with YO-PRO<sup>®</sup>-1 on sections prepared from primary foliage leaves. Microscopy was performed with a confocal laser-scanning microscope. Fluorescence signals were detected by sequential scanning [Ex 488 nm (Argonlaser 30%)/Em 500–550 nm and Ex 633 nm (HeNe Laser)/Em 650–750 nm]. **(B)** The diameters of fluorescence signals were measured with the quantification module of the Leica software LAS AF-TCS. The graph was generated with the program GraphPad Prism<sup>®</sup>.

large electron lucent areas containing filamentous structures (Figure 4).

To investigate whether *HvWHIRLY1* has also an effect on the structure of bacterial nucleoids, cells of *Escherichia coli* overexpressing the *HvWHIRLY1* gene were stained with DAPI (4',6-diamidino-2-phenylindole) as described in Melonek et al. (2012). In accordance to the microscopic observation of nucleoids in the RNAi-W1-7 plants, *E. coli* cells overexpressing *HvWHIRLY1* contained more tightly condensed bacterial nucleoids compared to control cells (Figure 5A). In parallel to an enhanced compactness of bacterial DNA, *E. coli* cells showed a reduced growth after induction of *HvWHIRLY1* overexpression (Figure 5B).

#### PLASTID DNA CONTENT

Microscopic analyses indicated that nucleoids in chloroplasts of RNAi-W1-7 plants are more heterogeneous in size and shape than in wild type chloroplasts. The large sizes of a subpopulation of the nucleoids in chloroplasts of RNAi-W1-7 plants suggest that also the DNA content could be enhanced in these chloroplasts. To investigate whether the differences in shape and size of nucleoids correlate with changes in the content of ptDNA, ptDNA levels were determined by two different methods. Firstly, DNA dot-blots were hybridized with a probe specific for the repetitive 18S nuclear DNA and second with the plastid DNA specific probe *petD*. Hybridization intensities were compared among dots of different total DNA contents. Hybridization signal intensities



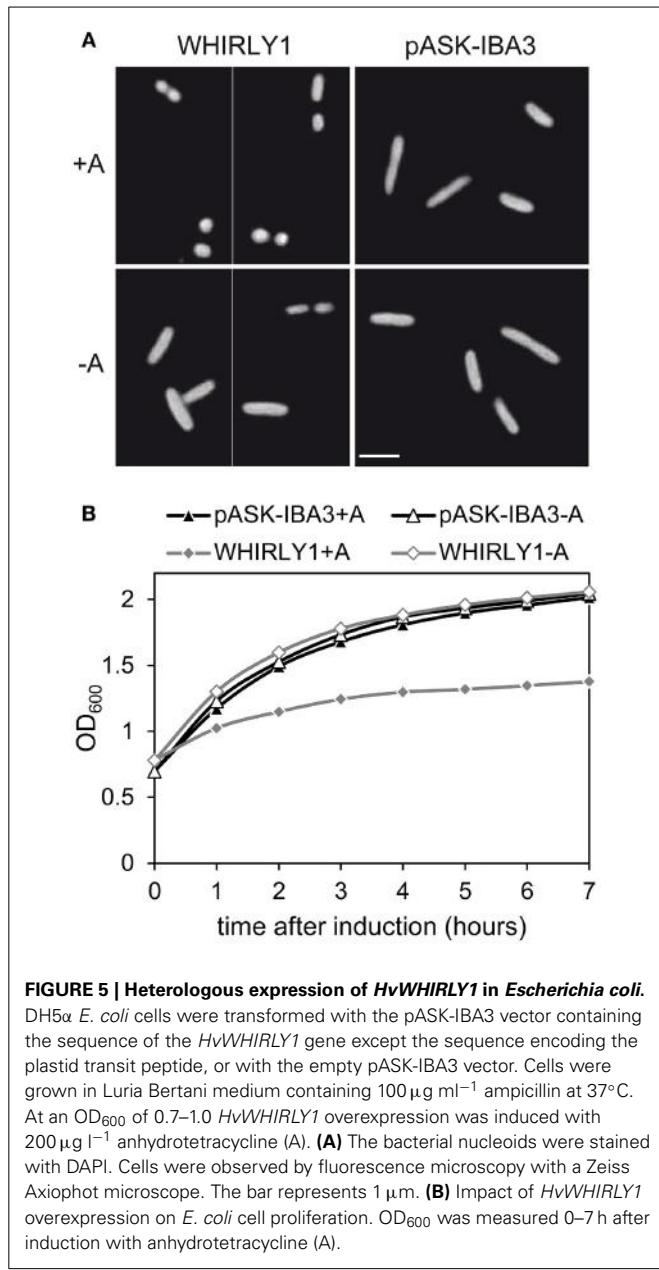
**FIGURE 4 |** Electron microscopy images of chloroplasts from wild type (WT) and the RNAi-W1-7 plants (W1-7). Leaf segments ( $2 \times 2$  mm) from 10 days old primary foliage leaves were taken at a position of 2 cm below the leaf tip. Some of the DNA containing regions are indicated by arrows. Transmission electron microscopy was performed with a Philips CM10 transmission electron microscope. Bars represent 500 nm.

obtained with the *petD* probe indicate that the level of plastid DNA is about two- to three-fold higher in the transgenic plants compared to the wild type (Figure 6A). Nuclear DNA level was similar in both cases, as shown by hybridization with the *18S* rRNA probe.

Furthermore, the relative copy number was determined by q-PCR with specific primers for two single copy plastid genes (*petD*, *psbA*) and the nuclear *RBCS* genes as internal standard. In comparison to the relative ptDNA level of the wild type, the relative level of ptDNA in leaves of transgenic RNAi-W1-7 plants was two-fold enhanced (Figure 6B). The relative level of RNAi-W1-1 plants was enhanced by about 50% in comparison to the wild type.

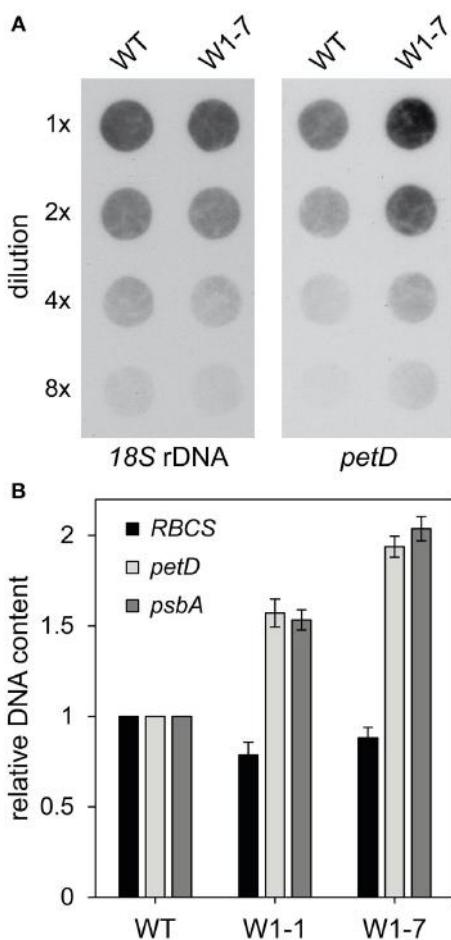
#### EXPRESSION OF A PUTATIVE BARLEY ORGANELLE DNA POLYMERASE IS REGULATED BY WHIRLY1

It has been suggested that WHIRLY proteins play roles in DNA repair together with an organelle targeted DNA polymerase (Parent et al., 2011) belonging to the family A of DNA polymerases and having sequence similarities to DNA polymerase I of *Escherichia coli* (Moriyama et al., 2011). So far, no organelle targeted DNA polymerase has been characterized for barley. To identify a sequence encoding a putative organelle DNA polymerase, barley sequence information from different sources (Consortium, 2012; Kohl et al., 2012; Thiel et al., 2012; Mascher et al., 2013) was assembled to create the full-length sequence of *HvPolI-like* (*KM236205*) using the CAP3 software (Huang and Madan, 1999). As reported for DNA polymerases from higher plants and from the primitive red alga *Cyanidioschyzon merolae* the barley sequence has an 3'-5' exonuclease domain besides a DNA



**FIGURE 5 |** Heterologous expression of *HvWHIRLY1* in *Escherichia coli*. DH5α *E. coli* cells were transformed with the pASK-IBA3 vector containing the sequence of the *HvWHIRLY1* gene except the sequence encoding the plastid transit peptide, or with the empty pASK-IBA3 vector. Cells were grown in Luria Bertani medium containing  $100 \mu\text{g ml}^{-1}$  ampicillin at  $37^\circ\text{C}$ . At an  $\text{OD}_{600}$  of 0.7–1.0 *HvWHIRLY1* overexpression was induced with  $200 \mu\text{g l}^{-1}$  anhydrotetracycline (A). (A) The bacterial nucleoids were stained with DAPI. Cells were observed by fluorescence microscopy with a Zeiss Axiohot microscope. The bar represents 1 μm. (B) Impact of *HvWHIRLY1* overexpression on *E. coli* cell proliferation.  $\text{OD}_{600}$  was measured 0–7 h after induction with anhydrotetracycline (A).

polymerase domain (Figure 7A). Whereas in *Arabidopsis* two organelle targeted DNA polymerases, also named POPs for plant organelle DNA polymerases (Moriyama et al., 2011), function redundantly in replication of both mitochondria and plastids (Parent et al., 2011), in maize a mutation of only one gene encoding a POP (*ZmPolI-like*) caused a severe decrease in plastid DNA copy number (Udy et al., 2012). The amino acid sequence of this POP (*ZmPolI-like*) has highest similarities (76.5% pairwise identity) to *HvPolI-like* in comparison to their orthologs from rice, tobacco, *Arabidopsis*, and a red alga (Figure 7A and Supplementary Figure 2). In accordance with a function in plastid DNA replication expression of the *HvPolI-like* gene is highest at the base of the leaves (Figure 7B) where also the replication activity is highest (Baumgartner et al., 1989). In the wild type,



**FIGURE 6 | DNA content of primary foliage leaves of the wild type and RNAi-W1-1 and RNAi-W1-7 plants. (A)** Total genomic DNA isolated from 10 days old primary foliage leaves was used in different dilutions (1x, 2x, 4x, and 8x) for DNA dot-blot hybridization. For detection of nuclear DNA the 18S rDNA probe was used and for detection of plastid DNA, *petD* was used. Probes were labeled with digoxigenin (DIG) (see Materials and Methods). **(B)** Analysis of the relative DNA content by q-PCR. Genomic DNA was isolated from 10 days old primary foliage leaves. For detection of nuclear DNA specific primers for *RBCS* and for detection of plastid DNA, specific primers for *petD* and *psbA* were used. Data of wild type (WT) were set to 1 and data of RNAi-W1-1 (W1-1) and RNAi-W1-7 (W1-7) are shown relative to the wild type.

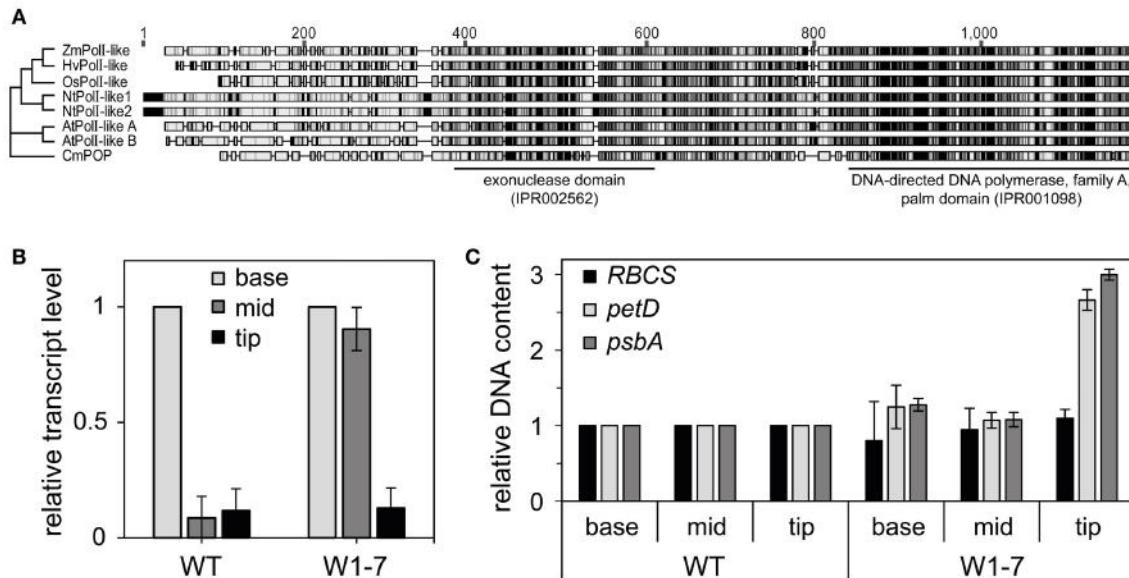
the RNA level declined rapidly during development of chloroplasts and is similar in segments from the mid and tip of the leaf (**Figure 7B**). When expression of the newly assembled *HvPOLL-like* gene was analyzed in corresponding sections from primary foliage leaves of RNAi-W1-7 plants, the level was found to be high at the base as well as in the mid part of the leaves (**Figure 7B**), suggesting that WHIRLY1 is involved in repression of *HvPOLL-like* gene expression during early chloroplast development. Interestingly, as a consequence of WHIRLY1 deficiency, ptDNA level is increased in the tip of leaves and not at the base and in the mid part of primary foliage leaves (**Figure 7C**). This might indicate that WHIRLY1 predominantly has impact on structure and functionality of nucleoids during development of mature chloroplasts.

## DISCUSSION

By using sections from different positions of barley primary foliage leaves, it has been shown that expression of the *WHIRLY1* gene is highest in immature cells at the leaf base and decreases during chloroplast development, whereas accumulation of the protein increases during early chloroplast development in parallel with that of the SVR4 protein, which was shown to be required for nucleoid organization during chloroplast development in *A. thaliana* (Powikrowska et al., 2014a). In contrast to the HvSVR4 protein, accumulation of HvWHIRLY1, however, was observed to decrease during maturation of chloroplasts in the upper part of the leaf (**Figure 1C** and Grabowski et al., 2008). This indicates that the two proteins, despite their similar patterns of gene expression, might have different functions. In Arabidopsis mutants lacking SVR4, accumulation of plastid RNAs synthesized by the plastid encoded RNA polymerase (PEP) is impaired. In contrast, transgenic barley RNAi-W1 plants were shown to have unaltered patterns of plastid transcripts when analyzed by run-on assays (Melonek et al., 2010). Considering that plastid DNA replication occurs early in leaf development and ceases during maturation of chloroplasts (Baumgartner et al., 1989), a function of WHIRLY1 in replication is likely.

Analyses of nucleoids stained with the fluorescing dye YO-PRO®-1 by confocal microscopy revealed large areas of DNA besides small punctuate nucleoids resembling those of the wild type chloroplasts. This suggests that WHIRLY1 is involved in compaction of only a subset of chloroplast nucleoids. This result is in accordance with the previous observation that a AtWHIRLY1:GFP fusion construct in tobacco protoplasts was associated with only a subset of the nucleoids (Melonek et al., 2010). The reduced compactness of nucleoids was confirmed by electron microscopy of chloroplasts in the mesophyll of the RNAi-W1-7 plants showing large electron lucent areas with filamentous structures. The compacting action of WHIRLY1 on nucleoids is not restricted to plastids, but occurs also in bacteria overexpressing the *HvWHIRLY1* gene. Compaction of the bacterial nucleoids by WHIRLY1 was accompanied by a decline in growth of the cells. Intriguingly, nucleus located WHIRLY1 is found in the heterochromatin (Grabowski et al., 2008). Whether WHIRLY1 has a function in chromatin compaction in the nucleus remains, however, to be shown. It also remains to be investigated whether compaction of plastid nucleoids by WHIRLY1 has consequences for chloroplast development and leaf growth under various conditions.

The altered organization of chloroplast nucleoids in leaves of RNAi-W1-7 plants indicates that WHIRLY1 belongs to the group of nucleoid architectural proteins (Dillon and Dorman, 2010; Krupinska et al., 2013). Architectural proteins can have different effects on nucleoids. They can organize the structure and compactness of ptDNA by forming bridges, by bending or by wrapping (Powikrowska et al., 2014b). DCP64, which is identical with sulfite reductase (SiR), was shown to bind and compact DNA (Cannon et al., 1999), thereby having negative effects on replication (Cannon et al., 1999) and transcription (Sekine et al., 2002). Another ptNAP (plastid nucleoid associated protein) shown to induce compaction of DNA is SWIB-4, which can functionally complement an *E. coli* mutant lacking the histone-like protein



**FIGURE 7 | Organelle DNA polymerases. (A)** Phylogenetic analyses of sequences predicted to encode plant organelle targeted DNA polymerases (POP). Unrooted neighbor-joining tree of plant organelle DNA polymerases (POP) generated using Geneious software (version 7.1.2) (Kearse et al., 2012). Underlying sequence alignment was generated using the ClustalW software (Larkin et al., 2007). Sequences from *Zea mays* (ZmPoll-like) (Schnable et al., 2009), *Oryza sativa* (OsPoll-like) (Kimura et al., 2002), *Nicotiana tabacum* (NtPoll-like1, NtPoll-like2) (Ono et al., 2007), *Arabidopsis thaliana* (AtPoll-like A, AtPoll-like B) (Mori et al., 2005), and *Cyanidioschyzon merolae* (CmPOP) (Moriyama et al., 2014) were compared to the predicted amino acid sequence

of a newly assembled gene (*HvPoll-like*) from *Hordeum vulgare* (KM236205). Interpro identifier (Hunter et al., 2012) were used for assignment of domains. **(B)** Expression of the *HvPoll-like* gene in sections of barley primary foliage leaves from the wild type (WT) and the RNAi-W1-7 (W1-7) line. Data of base were set to 1, the data for mid and tip are shown in relative to base. **(C)** Analysis of relative ptDNA content of the leaf sections of WT and W1-7 by q-PCR using specific primers for the *petD* and *psbA* gene. For comparison the nuclear encoded *RBCS* is shown. Data were normalized to the 18S rRNA gene. Data of WT were set to 1 and data of W1-7 are shown in relation to the wild type.

H-NS (Melonek et al., 2012). Other ptNAPs were shown to be involved in the tethering of DNA to membranes as described for the PEND protein (Sato et al., 1993) and for MFP1 (Jeong et al., 2003). Previously, it has been proposed that WHIRLY1 binding unspecifically to DNA, might have a similar function in chloroplasts as the HU protein or another abundant NAP in bacteria (Prikryl et al., 2008). Complementation assays with *E. coli* mutants lacking either HU or H-NS, another abundant NAP, however, failed, because expression of the *WHIRLY1* gene in *E. coli* has a general negative effect on cell growth (data not shown).

Fluorescence images of stained DNA in mesophyll chloroplasts of RNAi-W1-7 plants showed large irregular patches of DNA besides small punctuate nucleoids. The images suggest that the chloroplasts might contain more DNA. DNA dot-blot hybridization and q-PCR revealed that compared to wild type plants, in leaves of transgenic plants the level of ptDNA is enhanced two- to three-fold. Barley mesophyll cells were reported to contain 8000–12,000 copies of ptDNA, which are distributed among 60 chloroplasts. During mesophyll cell development in wheat leaves, an increase in plastid copy number per cell is due to an increase in plastid number and ptDNA copy number per plastid (Miyamura et al., 1986, 1990). It has been determined that ptDNA copy number per plastid increases more than two-fold during chloroplast development in the barley primary foliage leaf (Baumgartner et al., 1989), although it was observed to be already quite high in the leaf basal meristem (130 vs. the maximal number 210)

(Baumgartner et al., 1989). The authors concluded that a significant increase in DNA copy number occurs already during formation of the leaf basal meristem from cells of the grain leaf primordia, which in wheat contain 30-fold less plastid DNA than a mature leaf (Miyamura et al., 1986).

The enhanced level of plastid DNA in RNAi-W1 plants suggests that WHIRLY1 is involved in repression of replication during chloroplast development. Based on the available information on plastid located WHIRLY1, Pfalz and Pfannschmidt (2013) have assigned the protein to a replication/DNA inheritance sub-domain of the nucleoid. Localization of WHIRLY1 to a subpopulation of nucleoids only (Melonek et al., 2010) is in accordance with the observation that in a subset, and not in all nucleoids, packaging of DNA is affected. Perhaps, only a subpopulation of nucleoids is active in replication as also demonstrated for mitochondrial nucleoids (Meeusen and Nunnari, 2003). Functional and structural variance among the nucleoids of chloroplasts has already been suggested early (Kowallik and Herrmann, 1972). The association of WHIRLY1 with other proteins of the replication subdomain remains, however, to be demonstrated by co-localization studies with e.g. DNA polymerases, topoisomerases, and gyrases. Indeed, several proteins predicted to be involved in replication have been identified in nucleoid preparations (Pfalz et al., 2006; Olinares et al., 2010; Majeran et al., 2012; Melonek et al., 2012). Two DNA polymerases homologous to bacterial DNA polymerase I were shown to be targeted to both organelles

(Elo et al., 2003; Christensen et al., 2005). Divergent roles were proposed for the two PolI-like organelle polymerases Pol IA and Pol IB by Parent et al. (2011). Although both polymerases are involved in replication in both organelles, only Pol IB was shown to be in addition involved in repair of double strand breaks induced by ciprofloxacin (Parent et al., 2011). So far, barley proteins involved in plastid DNA replication were unknown. To get access to the sequence of a putative DNA polymerase, barley sequences from different sources were screened with sequence information of organelle DNA polymerases from maize (Udy et al., 2012), rice (Kimura et al., 2002), and dicots (Mori et al., 2005; Ono et al., 2007). Expression of the newly identified gene encoding a putative organelle targeted DNA polymerase of barley (*HvPolI-like*) was highest at the base of the leaves and declined dramatically during chloroplast development. This pattern of expression is in accordance with a function in replication of plastid DNA. When expression of the *HvPolI-like* gene was analyzed in RNA-W1 plants, a higher mRNA level was found only in the mid of the leaves, where in wild type leaves accumulation of WHIRLY1 is highest. This indicates that the genetic disruption of WHIRLY1 has a positive impact on expression of the *HvPolI-like* gene.

Besides DNA polymerase IB, also WHIRLY proteins have been proposed to assist the repair of double strand breaks induced by ciprofloxacin (Maréchal et al., 2009). Plastids of the *Arabidopsis why1why3* double mutant were shown to accumulate aberrant DNA molecules caused by deletions, duplication and circularization events resulting from illegitimate recombination between microhomologous repeat sequences (Maréchal et al., 2009). In about 5% of the progeny, variegated leaves containing dysfunctional plastids were observed. A triple mutant resulting from a cross between the double mutant *why1why3* and the *polIB* mutant showed a more severe phenotype, suggesting that WHIRLY proteins and DNA Pol IB act synergistically in preventing aberrant recombinations of ptDNA (Parent et al., 2011; Lepage et al., 2013). Preliminary investigations on recombination of ptDNA in chloroplasts of primary foliage leaves of the barley RNAi-W1-7 plants did not show differences between wild type and transgenic plants. Similar investigations under stress conditions and/or after addition of ciprofloxacin remain, however, to be done.

So far, it is not known which factors regulate the different activities of DNA polymerase IB. It is, however, likely that its replication activity in plastids declines during chloroplast development. Indeed, its expression is highest in tissues with high cell density where cell expansion occurs (Cupp and Nielsen, 2013). Accordingly, an *Arabidopsis polIB* mutant has a delay in cell elongation. However, so far no information is available on the accumulation of organelle targeted DNA polymerases in plastids of different developmental stages. WHIRLY1 deficiency interestingly alters the ptDNA level in mature chloroplasts, but not in younger stages. Perhaps WHIRLY proteins just change the activity of DNA polymerase at specific stages of development by structural changes in the replication subdomain of nucleoids, although a negative regulation of *HvPolI-like* gene expression might contribute to the repression of replication during chloroplast development. It had already been proposed that with regard to their multifunctionality WHIRLY proteins

resemble the bacterial SSB proteins (Maréchal and Brisson, 2010), which are dynamic centers playing key roles in choreographing diverse processes surrounding DNA replication, recombination and repair (Shereda et al., 2008). As in the case of SSB, the functional consequences of a reduced level of WHIRLY1 might differ depending on the developing stage of plastids and the environmental context. It is expected that the level of WHIRLY1 under certain conditions can have tremendous impact on growth and on productivity of crop plants.

## ACKNOWLEDGMENTS

The expert technical assistance of Cornelia Marthe, IPK Gatersleben, during preparation of transgenic plants is gratefully acknowledged. We further acknowledge expert technical assistance of Susanne Braun, Jens Herrmann and Ulrike Voigt (University of Kiel). We thank Marita Beese from the Central Microscopy, University of Kiel, for excellent technical support during preparation of specimen for electron microscopy. Microscopy facilities have been provided by the Central Microscopy of the University of Kiel. Rena Isemer and Uwe Bertsch (University of Kiel) are thanked for discussion and critical reading of the manuscript.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://www.frontiersin.org/journal/10.3389/fpls.2014.00432/abstract>

## REFERENCES

- Baumgartner, B. J., Rapp, J. C., and Mullet, J. E. (1989). Plastid transcription activity and DNA copy number increase early in barley chloroplast development. *Plant Physiol.* 89, 1011–1018. doi: 10.1104/pp.89.3.1011
- Cannon, G. C., Ward, L. N., Case, C. I., and Heinhorst, S. (1999). The 68 kDa DNA compacting nucleoid protein from soybean chloroplasts inhibits DNA synthesis *in vitro*. *Plant Mol. Biol.* 39, 835–845. doi: 10.1023/A:1006135615924
- Christensen, A. C., Lyznik, A., Mohammed, S., Elowsky, C. G., Elo, A., Yule, R., et al. (2005). Dual-domain, dual-targeting organelar protein presequences in *Arabidopsis* can use non-AUG start codons. *Plant Cell* 17, 2805–2816. doi: 10.1105/tpc.105.035287
- Consortium, T. I. B. G. S. (2012). A physical, genetic and functional sequence assembly of the barley genome. *Nature* 491, 711–716. doi: 10.1038/nature11543
- Cupp, J. D., and Nielsen, B. L. (2013). *Arabidopsis thaliana* organellar DNA polymerase IB mutants exhibit reduced mtDNA levels with a decrease in mitochondrial area density. *Physiol. Plant.* 149, 91–103. doi: 10.1111/ppl.12009
- Desveaux, D., Després, C., Joyeux, A., Subramaniam, R., and Brisson, N. (2000). PBF-2 is a novel single-stranded DNA binding factor implicated in PR-10a gene activation in potato. *Plant Cell* 12, 1477–1489. doi: 10.1105/tpc.12.8.1477
- Dillon, S. C., and Dorman, C. J. (2010). Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. *Nat. Rev. Microbiol.* 8, 185–195. doi: 10.1038/nrmicro2261
- Elo, A., Lyznik, A., Gonzalez, D. O., Kachman, S. D., and Mackenzie, S. A. (2003). Nuclear genes that encode mitochondrial proteins for DNA and RNA metabolism are clustered in the *Arabidopsis* genome. *Plant Cell* 15, 1619–1631. doi: 10.1105/tpc.010009
- Fling, S. P. S., and Gregerson, D. S. D. (1986). Peptide and protein molecular weight determination by electrophoresis using a high-molarity tris buffer system without urea. *Anal. Biochem.* 155, 83–88. doi: 10.1016/0003-2697(86)90228-9
- Fulton, T. M., Chunwongse, J., and Tanksley, S. D. (1995). Microprep protocol for extraction of DNA from tomato and other herbaceous plants. *Plant Mol. Biol. Rep.* 13, 207–209. doi: 10.1007/BF02670897

- Grabowski, E., Miao, Y., Mulisch, M., and Krupinska, K. (2008). Single-stranded DNA-binding protein Whirly1 in barley leaves is located in plastids and the nucleus of the same cell. *Plant Physiol.* 147, 1800–1804. doi: 10.1104/pp.108.122796
- Hensel, G., Valkov, V., Middlefell-Williams, J., and Kumlehn, J. (2008). Efficient generation of transgenic barley: the way forward to modulate plant–microbe interactions. *J. Plant Physiol.* 165, 71–82. doi: 10.1016/j.jplph.2007.06.015
- Himmelbach, A., Zierold, U., Hensel, G., Riechen, J., Douchkov, D., Schweizer, P., et al. (2007). A set of modular binary vectors for transformation of cereals. *Plant Physiol.* 145, 1192–1200. doi: 10.1104/pp.107.111575
- Huang, X., and Madan, A. (1999). CAP3: a DNA sequence assembly program. *Genome Res.* 9, 868–877. doi: 10.1101/gr.9.9.868
- Humbeck, K., Quast, S., and Krupinska, K. (1996). Functional and molecular changes in the photosynthetic apparatus during senescence of flag leaves from field-grown barley plants. *Plant Cell Environ.* 19, 337–344. doi: 10.1111/j.1365-3040.1996.tb00256.x
- Hunter, S., Jones, P., Mitchell, A., Apweiler, R., Attwood, T. K., Bateman, A., et al. (2012). InterPro in 2011: new developments in the family and domain prediction database. *Nucleic Acids Res.* 40, D306–D312. doi: 10.1093/nar/gkr948
- Jeong, S. Y., Rose, A., and Meier, I. (2003). MFP1 is a thylakoid-associated, nucleoid-binding protein with a coiled-coil structure. *Nucleic Acids Res.* 31, 5175–5185. doi: 10.1093/nar/gkg693
- Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., et al. (2012). Geneious basic: an integrated and extendable desktop software platform for the organisation and analysis of sequence data. *Bioinformatics* 28, 1647–1649. doi: 10.1093/bioinformatics/bts199
- Kimura, S., Uchiyama, Y., Kasai, N., Namekawa, S., Saotome, A., Ueda, T., et al. (2002). A novel DNA polymerase homologous to *Escherichia coli* DNA polymerase I from a higher plant, rice (*Oryza sativa* L.). *Nucleic Acids Res.* 30, 1585–1592. doi: 10.1093/nar/30.7.1585
- Kohl, S., Hollmann, J., Blattner, F. R., Radchuk, V., Andersch, F., Steuernagel, B., et al. (2012). A putative role for amino acid permeases in sink-source communication of barley tissues uncovered by RNA-seq. *BMC Plant Biol.* 12:154. doi: 10.1186/1471-2229-12-154
- Kowallik, K. V., and Herrmann, R. G. (1972). Variable amounts of DNA related to the size of chloroplasts. IV. Three-dimensional arrangement of DNA in fully differentiated chloroplasts of *Beta vulgaris* L. *J. Cell Sci.* 11, 357–377.
- Krause, K., Kilbienski, I., Mulisch, M., Rödiger, A., Schäfer, A., and Krupinska, K. (2005). DNA-binding proteins of the Whirly family in *Arabidopsis thaliana* are targeted to the organelles. *FEBS Lett.* 579, 3707–3712. doi: 10.1016/j.febslet.2005.05.059
- Krupinska, K., Melonek, J., and Krause, K. (2013). New insights into plastid nucleoid structure and functionality. *Planta* 237, 653–664. doi: 10.1007/s00425-012-1817-5
- Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., et al. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* 23, 2947–2948. doi: 10.1093/bioinformatics/btm404
- Lepage, É., Zampini, É., and Brisson, N. (2013). Plastid genome instability leads to reactive oxygen species production and plastid-to-nucleus retrograde signaling in *Arabidopsis*. *Plant Physiol.* 163, 867–881. doi: 10.1104/pp.113.223560
- Majeran, W., Friso, G., Asakura, Y., Qu, X., Huang, M., Ponnala, L., et al. (2012). Nucleoid-enriched proteomes in developing plastids and chloroplasts from maize leaves: a new conceptual framework for nucleoid functions. *Plant Physiol.* 158, 156–189. doi: 10.1104/pp.111.188474
- Maréchal, A., and Brisson, N. (2010). Recombination and the maintenance of plant organelle genome stability. *New Phytol.* 186, 299–317. doi: 10.1111/j.1469-8137.2010.03195.x
- Maréchal, A., Parent, J. S., Veronneau-Lafortune, F., Joyeux, A., Lang, B. F., and Brisson, N. (2009). Whirly proteins maintain plastid genome stability in *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A.* 106, 14693–14698. doi: 10.1073/pnas.0901710106
- Mascher, M., Richmond, T. A., Gerhardt, D. J., Himmelbach, A., Clissold, L., Sampath, D., et al. (2013). Barley whole exome capture: a tool for genomic research in the genus *Hordeum* and beyond. *Plant J.* 76, 494–505. doi: 10.1111/tpj.12294
- Meeusen, S., and Nunnari, J. (2003). Evidence for a two membrane-spanning autonomous mitochondrial DNA replisome. *J. Cell Biol.* 163, 503–510. doi: 10.1083/jcb.200304040
- Melonek, J., Matros, A., Trösch, M., Mock, H. P., and Krupinska, K. (2012). The core of chloroplast nucleoids contains architectural SWIB domain proteins. *Plant Cell* 24, 3060–3073. doi: 10.1105/tpc.112.099721
- Melonek, J., Mulisch, M., Schmitz-Linneweber, C., Grabowski, E., Hensel, G., and Krupinska, K. (2010). Whirly1 in chloroplasts associates with intron containing RNAs and rarely co-localizes with nucleoids. *Planta* 232, 471–481. doi: 10.1007/s00425-010-1183-0
- Miyamura, S., Kuroiwa, T., and Nagata, T. (1990). Multiplication and differentiation of plastid nucleoids during development of chloroplasts and etioplasts from proplastids in *Triticum aestivum*. *Plant Cell Physiol.* 31, 597–602.
- Miyamura, S., Nagata, T., and Kuroiwa, T. (1986). Quantitative fluorescence microscopy on dynamic changes of plastid nucleoids during wheat development. *Protoplasma* 133, 66–72. doi: 10.1007/BF01293188
- Mori, Y., Kimura, S., Saotome, A., Kasai, N., Sakaguchi, N., Uchiyama, Y., et al. (2005). Plastid DNA polymerases from higher plants, *Arabidopsis thaliana*. *Biochem. Biophys. Res. Commun.* 334, 43–50. doi: 10.1016/j.bbrc.2005.06.052
- Moriyama, T., Tajima, N., Sekine, K., and Sato, N. (2014). Localization and phylogenetic analysis of enzymes related to organellar genome replication in the unicellular rhodophyte *Cyanidioschyzon merolae*. *Genome Biol. Evol.* 6, 228–237. doi: 10.1093/gbe/evu009
- Moriyama, T., Terasawa, K., and Sato, N. (2011). Conservation of POPs, the plant organellar DNA polymerases, in eukaryotes. *Protist* 162, 177–187. doi: 10.1016/j.protis.2010.06.001
- Niinemets, Ü., Bilger, W., Kull, O., and Tenhunen, J. D. (1998). Acclimation to high irradiance in temperate deciduous trees in the field: changes in xanthophyll cycle pool size and in photosynthetic capacity along a canopy light gradient. *Plant Cell Environ.* 21, 1205–1218. doi: 10.1046/j.1365-3040.1998.00364.x
- Olinares, P. D. B., Ponnala, L., and Van Wijk, K. J. (2010). Megadalton complexes in the chloroplast stroma of *Arabidopsis thaliana* characterized by size exclusion chromatography, mass spectrometry, and hierarchical clustering. *Mol. Cell. Proteomics* 9, 1594–1615. doi: 10.1074/mcp.M000038-MCP201
- Ono, Y., Sakai, A., Takechi, K., Takio, S., Takusagawa, M., and Takano, H. (2007). NtPoli-like1 and NtPoli-like2, bacterial DNA polymerase I homologs isolated from BY-2 cultured tobacco cells, encode DNA polymerases engaged in DNA replication in both plastids and mitochondria. *Plant Cell Physiol.* 48, 1679–1692. doi: 10.1093/pcp/pcm140
- Palotta, M. A., Graham, R. D., Langridge, P., Sparrow, D. H. B., and Barker, S. J. (2000). RFLP mapping of manganese efficiency in barley. *Theor. Appl. Genet.* 101, 1100–1108. doi: 10.1007/s001220051585
- Parent, J., Lepage, É., and Brisson, N. (2011). Divergent roles for the two PolI-like organelle DNA polymerases of *Arabidopsis*. *Plant Physiol.* 156, 254–262. doi: 10.1104/pp.111.173849
- 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. doi: 10.1105/tpc.105.036392
- Pfalz, J., and Pfannschmidt, T. (2013). Essential nucleoid proteins in early chloroplast development. *Trends Plant Sci.* 18, 186–194. doi: 10.1016/j.tplants.2012.11.003
- Powikowska, M., Khrouchtchova, A., Martens, H. J., Zygałdo-Nielsen, A., Melonek, J., Schulz, A., et al. (2014a). SVR4 (suppressor of variegation 4) and SVR4-like: two proteins with a role in proper organisation of the chloroplast genetic machinery. *Physiol. Plant* 150, 477–492. doi: 10.1111/plp.12108
- Powikowska, M., Oetke, S., Jensen, P. E., and Krupinska, K. (2014b). Dynamic composition, shaping and organisation of plastid nucleoids – a structural perspective. *Front. Plant Sci.* 5:424. doi: 10.3389/fpls.2014.00424
- Prikryl, J., Watkins, K. P., Friso, G., Van Wijk, K. J., and Barkan, A. (2008). A member of the Whirly family is a multifunctional RNA- and DNA-binding protein that is essential for chloroplast biogenesis. *Nucleic Acids Res.* 36, 5152–5165. doi: 10.1093/nar/gkn492
- Reynolds, E. S. (1963). The use of lead citrate at high pH as an electron-opaque stain in electron microscopy. *J. Cell Biol.* 17, 208–212. doi: 10.1083/jcb.17.1.208
- Sakai, A., Takano, H., and Kuroiwa, T. (2004). Organelle nuclei in higher plants: structure, composition, function, and evolution. *Int. Rev. Cytol.* 238, 59–118. doi: 10.1016/S0074-7696(04)38002-2

- Sato, N., Albrieux, C., Joyard, J., Douce, R., and Kuroiwa, T. (1993). Detection and characterization of a plastid envelope DNA-binding protein which may anchor plastid nucleoids. *EMBO J.* 12, 555–561.
- Schnable, P. S., Ware, D., Fulton, R. S., Stein, J. C., Wei, F., Pasternak, S., et al. (2009). The B73 maize genome: complexity, diversity, and dynamics. *Science* 326, 1112–1115. doi: 10.1126/science.1178534
- Sekine, K., Hase, T., and Sato, N. (2002). Reversible DNA compaction by sulfite reductase regulates transcriptional activity of chloroplast nucleoids. *J. Biol. Chem.* 277, 24399–24404. doi: 10.1074/jbc.M201714200
- Shereda, R. D., Kozlov, A. G., Lohman, T. M., Cox, M. M., and Keck, J. L. (2008). SSB as an organizer/mobilizer of genome maintenance complexes. *Crit. Rev. Biochem. Mol. Biol.* 43, 289–318. doi: 10.1080/10409230802341296
- Thiel, J., Hollmann, J., Rutten, T., Weber, H., Scholz, U., and Weschke, W. (2012). 454 Transcriptome sequencing suggests a role for two-component signalling in cellularization and differentiation of barley endosperm transfer cells. *PLoS ONE* 7:e41867. doi: 10.1371/journal.pone.0041867.t003
- Udy, D. B., Belcher, S., Williams-Carrier, R., Gualberto, J. M., and Barkan, A. (2012). Effects of reduced chloroplast gene copy number on chloroplast gene expression in maize. *Plant Physiol.* 160, 1420–1431. doi: 10.1104/pp.112.204198

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Received: 27 June 2014; accepted: 13 August 2014; published online: 04 September 2014.*

*Citation: Krupinska K, Oetke S, Desel C, Mulisch M, Schäfer A, Hollmann J, Kumlehn J and Hensel G (2014) WHIRLY1 is a major organizer of chloroplast nucleoids. *Front. Plant Sci.* 5:432. doi: 10.3389/fpls.2014.00432*

*This article was submitted to Plant Physiology, a section of the journal Frontiers in Plant Science.*

*Copyright © 2014 Krupinska, Oetke, Desel, Mulisch, Schäfer, Hollmann, Kumlehn and Hensel. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*



# Enzymes involved in organellar DNA replication in photosynthetic eukaryotes

Takashi Moriyama<sup>1,2</sup> and Naoki Sato<sup>1,2 \*</sup>

<sup>1</sup> Department of Life Sciences, Graduate School of Arts and Sciences, The University of Tokyo, Tokyo, Japan

<sup>2</sup> Japan Science and Technology Agency – Core Research for Evolutional Science and Technology, Tokyo, Japan

**Edited by:**

Thomas Pfannschmidt, Université Joseph Fourier Grenoble, France

**Reviewed by:**

Ján A. Miernyk, University of Missouri, USA

Mee-Len Chee, The University of Hong Kong, China

**\*Correspondence:**

Naoki Sato, Department of Life Sciences, Graduate School of Arts and Sciences, The University of Tokyo, Komaba 3-8-1, Meguro-ku, Tokyo 153-8902, Japan  
e-mail: naokisat@bio.c.u-tokyo.ac.jp

Plastids and mitochondria possess their own genomes. Although the replication mechanisms of these organellar genomes remain unclear in photosynthetic eukaryotes, several organelle-localized enzymes related to genome replication, including DNA polymerase, DNA primase, DNA helicase, DNA topoisomerase, single-stranded DNA maintenance protein, DNA ligase, primer removal enzyme, and several DNA recombination-related enzymes, have been identified. In the reference Eudicot plant *Arabidopsis thaliana*, the replication-related enzymes of plastids and mitochondria are similar because many of them are dual targeted to both organelles, whereas in the red alga *Cyanidioschyzon merolae*, plastids and mitochondria contain different replication machinery components. The enzymes involved in organellar genome replication in green plants and red algae were derived from different origins, including proteobacterial, cyanobacterial, and eukaryotic lineages. In the present review, we summarize the available data for enzymes related to organellar genome replication in green plants and red algae. In addition, based on the type and distribution of replication enzymes in photosynthetic eukaryotes, we discuss the transitional history of replication enzymes in the organelles of plants.

**Keywords:** plastids, mitochondria, organellar genome, replication

## INTRODUCTION

Plastids and mitochondria are semi-autonomous organelles that contain their own genomes, encoding the genes necessary to perform their respective metabolic functions. These organellar genomes are replicated by specific enzymes, such as DNA polymerase, DNA primase, and DNA helicase, as occurs in bacteria and the nuclei of eukaryotes. In contrast, plant organellar genomes do not encode these replicative proteins, and are instead replicated by nucleus-encoded enzymes that are transported to the organelles. However, the mechanism of plant organellar genome replication is not clearly understood because more than one mode of replication is possible and include recombination-dependent, double D-loop, and rolling circle replication mechanisms (reviewed in Maréchal and Brisson, 2010; Nielsen et al., 2010; Gualberto et al., 2014).

Mitochondria and plastids are thought to have been acquired through endosymbiotic events with ancestors of  $\alpha$ -proteobacteria and cyanobacteria, respectively. Studies of bacterial replication enzymes (Langston et al., 2009; Sanyal and Doig, 2012) have shown that homologs of these enzymes also function in plant organelles. In bacterial DNA replication, DnaB helicase unwinds double-stranded DNA (dsDNA) at the replication fork, and the unwound DNA is then prevented from re-annealing with other single-stranded DNAs (ssDNAs) by single-stranded DNA-binding protein (SSB). A type II DNA topoisomerase (gyrase) consisting of A and B subunits alleviates the mechanical strain of unwound DNA. DnaG primase synthesizes an RNA primer, which is elongated by DNA polymerase III and is then removed by nick translation with 5'-3' exonuclease and the polymerase activity of

DNA polymerase I (PolI). The nicked DNA is combined by the NAD<sup>+</sup>-dependent DNA ligase LigA.

All of the enzymes involved in mitochondrial DNA replication in vertebrates have been identified (Arnold et al., 2012; Kasiviswanathan et al., 2012). DNA polymerase  $\gamma$  (Poly $\gamma$ ) functions in the replication and repair of animal mitochondrial DNA. Animal Poly consists of two subunits: a large subunit with DNA polymerase and 3'-5' exonuclease activities, and a small subunit that functions in primer recognition and enhances processivity (Kaguni, 2004; Wanrooij and Falkenberg, 2010). The animal mitochondrial primase, POLRMT, which has homology to the RNA polymerase of T3/T7 phage, was recently indicated to function in both primer synthesis and transcription, although it was previously thought to function only in transcription (Wanrooij et al., 2008). Plants also have homolog(s) of T3/T7 phage RNA polymerase, which are named RPOTs (RNA polymerase of the T3/T7 type) and are localized to plastids and/or mitochondria, where they function in transcription (Kühn et al., 2007). Mitochondrial helicase is called TWINKLE and has homology to the gp protein of T7 phage, which contains primase and helicase domains at the N- and C-termini, respectively. However, although the TWINKLE found in bikonts (plants and protists) has both primase and helicase activities, animal TWINKLE only shows helicase activity (Shutt and Gray, 2006). A number of other replication-related enzymes, including topoisomerases 1 and 3a, SSB, ligase 3, and RNase H1, have also been identified in human mitochondria. An *in vitro* reconstituted mitochondrial replisome composed of Poly, TWINKLE, and SSB displayed rolling-circle replication with high processivity (Korhonen et al., 2004).

Several of the replicative enzymes found in bacteria and animal mitochondria are also encoded by plant nuclear genomes. In addition to these common enzymes, a number of plant-specific enzymes for DNA replication and recombination have recently been identified, and their subcellular localization has been examined in both plants and algae. In this article, we summarize the current knowledge on enzymes related to organellar replication in photosynthetic eukaryotes and also discuss the evolution of these replication-related enzymes based on their distribution in photosynthetic eukaryotes.

## REPLICATION DNA POLYMERASE, POP

### HISTORY OF STUDIES ON ORGANELLAR DNA POLYMERASES IN PLANTS

DNA replication activity was first detected in isolated organelles from plants, yeasts, and animals in the late 1960s (Wintersberger, 1966; Parsons and Simpson, 1967; Spencer and Whitfeld, 1967; Tewari and Wildman, 1967). In the following decade, DNA polymerases were purified from isolated chloroplasts and mitochondria of various photosynthetic eukaryotes (summarized in Moriyama and Sato, 2013). Biochemical data suggested that plant organellar and  $\gamma$ -type DNA polymerases, which are responsible for replication of the mitochondrial genome in fungi and animals (Lecrenier and Foury, 2000; Kaguni, 2004), had similar optimal enzymatic conditions, particularly pH and the concentration of monovalent and divalent ions, sensitivity to DNA polymerase inhibitors, molecular size, and template preference. Despite such biochemical evidence, no gene encoding a homolog of Poly has been found in the sequenced genomes of bikonts, including those of plants and protists, and the organellar DNA polymerase in photosynthetic organisms remains unidentified. Sakai et al. (1999) and Sakai (2001) detected DNA synthetic activity in the nucleoid fraction isolated from chloroplasts and mitochondria of tobacco and determined that the apparent molecular mass of the enzyme exhibiting the activity was similar to the Klenow fragment of PolI in *Escherichia coli*. This finding led to the identification of a gene(s) encoding a DNA polymerase with distant homology to *E. coli* PolI in the genomes of bikonts. The identified DNA polymerase was first isolated from plastids of rice, and its localization was confirmed by immunoblot analysis of isolated plastids (Kimura et al., 2002). Subsequent studies using GFP-fusion proteins and/or immunoblotting demonstrated that the polymerases, which were named PolI-like, PolI or Poly, are localized to both plastids and mitochondria in *Arabidopsis thaliana* and tobacco (Christensen et al., 2005; Mori et al., 2005; Ono et al., 2007; Parent et al., 2011; Cupp and Nielsen, 2013). We also identified this type of DNA polymerase in algae and ciliates (Moriyama et al., 2008, 2011, 2014). In these reports, phylogenetic analysis of Family A DNA polymerases revealed that plant organellar DNA polymerases belong to a clade that is distinct from that of bacterial PolI and Poly (Figure 1). In addition, red algae were found to encode a DNA polymerase with high homology to *E. coli* PolI (Moriyama et al., 2008). Therefore, we proposed that this type of organellar DNA polymerase be named POP (plant and protist organellar DNA polymerase), because the genes encoding the polymerases are present in both photosynthetic eukaryotes and protists.

### ENZYMATIC CHARACTERISTICS OF POPs

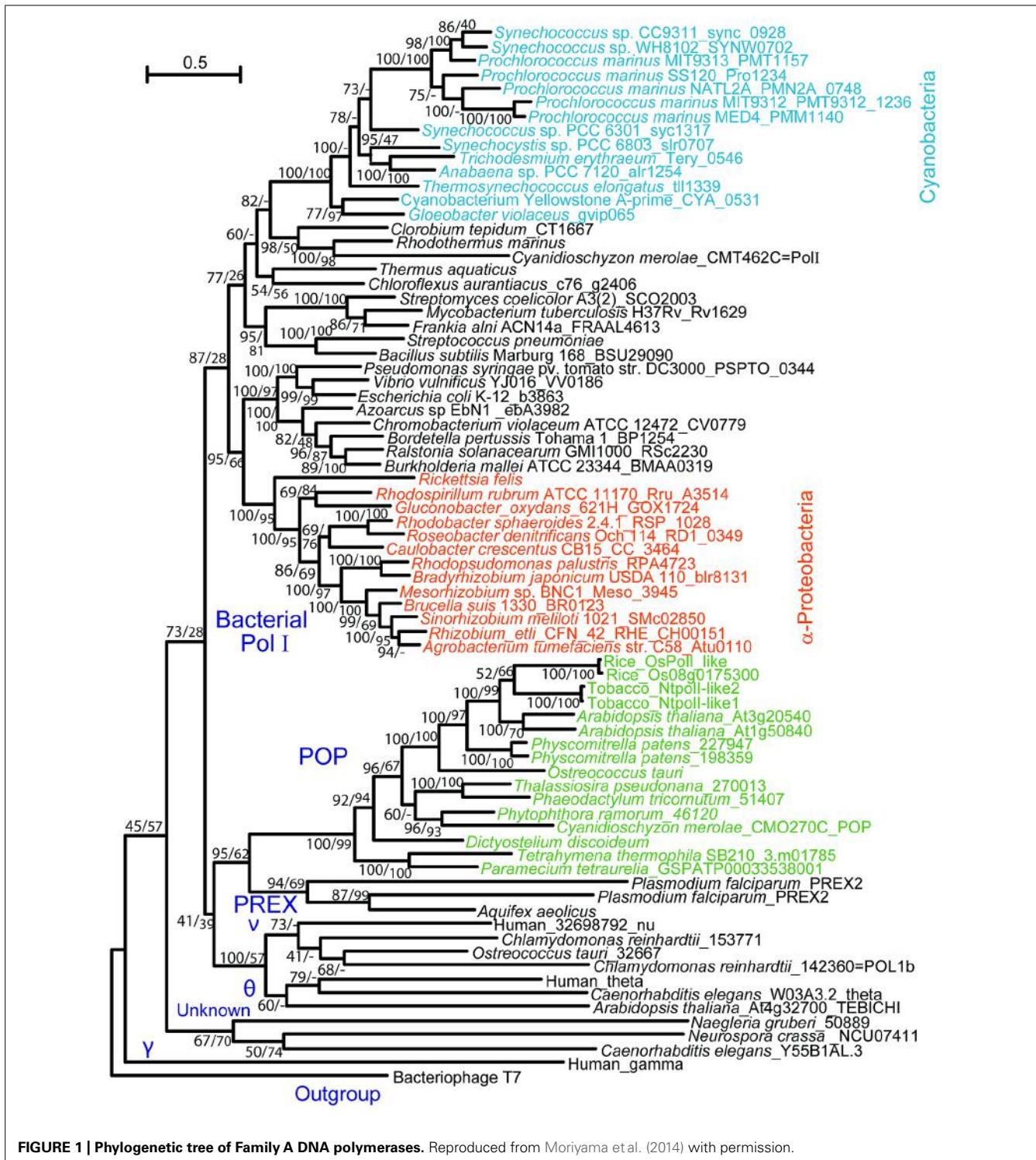
#### *Characterization of DNA polymerase activity of POPs*

The DNA polymerase activity of POPs has been characterized using recombinant (Kimura et al., 2002; Ono et al., 2007; Takeuchi et al., 2007) and native proteins purified from *Cyanidioschyzon merolae* (red alga) and isolated mitochondria of *Tetrahymena thermophila* (Moriyama et al., 2008, 2011). The optimal KCl concentration for POP polymerase activity is 50–150 mM. POPs also show divalent metal ion-dependent activity, and the optimal MgCl<sub>2</sub> concentration for their activity is 2.5–5 mM. POPs exhibit the highest activity with poly(dA)/oligo(dT) as a template, rather than with activated calf thymus DNA. However, poly(rA)/oligo(dT) can also be used as a template, indicating that POPs have reverse transcriptase activity, similar to Poly, although the *in vivo* role of this activity has not been elucidated.

DNA polymerase enzymes bind to and dissociate from template DNA repeatedly during the replication or repair process. The number of synthesized nucleotides added by the DNA polymerase per one binding event is defined as processivity. POPs show markedly high processivity values, ranging from 600 to 900 nt for recombinant rice POP and 1,300 nt for *Cyanidioschyzon merolae* POP (Figure 2, Moriyama et al., 2008). In comparison, *E. coli* PolI has mid-range processivity of <15 nt (Takeuchi et al., 2007). The *Cyanidioschyzon merolae* genome encodes a *Poll* gene (*CmPoll*) having high homology with *E. coli* PolI, and CmPoll also has mid-range processivity of <70 nt (Moriyama et al., 2008). Alignment analysis of POPs with other Family A DNA polymerases revealed that POP proteins have additional sequences that are involved in DNA binding and synthesis activities, suggesting that the high processivity of POPs might be attributable to these extra sequences (Takeuchi et al., 2007). In animals, an accessory subunit (PolyB) of Poly enhances the processivity of PolyA. For example, the processivities of the *Drosophila* PolyA subunit and a Poly holoenzyme consisting of PolyA and PolyB are <40 nt and >1,000 nt, respectively (Williams et al., 1993). In contrast, POPs show high processivity as a single subunit enzyme, and to our knowledge, no accessory proteins associated with POP have been identified.

#### *Sensitivity to DNA polymerase inhibitors*

The effect of inhibitors, such as aphidicolin, N-ethylmaleimide (NEM), dideoxyTTP (ddTTP), and phosphonoacetate (PAA), on the DNA synthesis activity of POPs was evaluated (Kimura et al., 2002; Ono et al., 2007; Moriyama et al., 2008, 2011). Aphidicolin is an inhibitor of eukaryotic nuclear DNA polymerases  $\alpha$ ,  $\delta$ , and  $\epsilon$  (Holmes, 1981; Weiser et al., 1991), and did not inhibit POP activity (Moriyama et al., 2008). NEM is an inhibitor of DNA polymerases  $\alpha$ ,  $\gamma$ ,  $\delta$ , and  $\epsilon$  (Chavalitshekoon-Petmitr et al., 2000), and also did not inhibit POP. ddTTP, which severely impairs the activity of DNA polymerases  $\beta$  and  $\gamma$  (Kornberg and Baker, 1992), had various inhibitory effects on POP depending on the organism, with the half maximal inhibitory concentrations (IC<sub>50</sub>) ranging from 4 to 615  $\mu$ M for the POPs of *A. thaliana*, *Cyanidioschyzon merolae*, and the ciliate *Tetrahymena* (Moriyama et al., 2011). PAA was originally identified as an inhibitor of viral DNA polymerases and reverse transcriptases, and functions by interacting with pyrophosphate binding sites, leading to an alternative reaction pathway (Leinbach et al., 1976; Shiraki et al., 1989). PAA

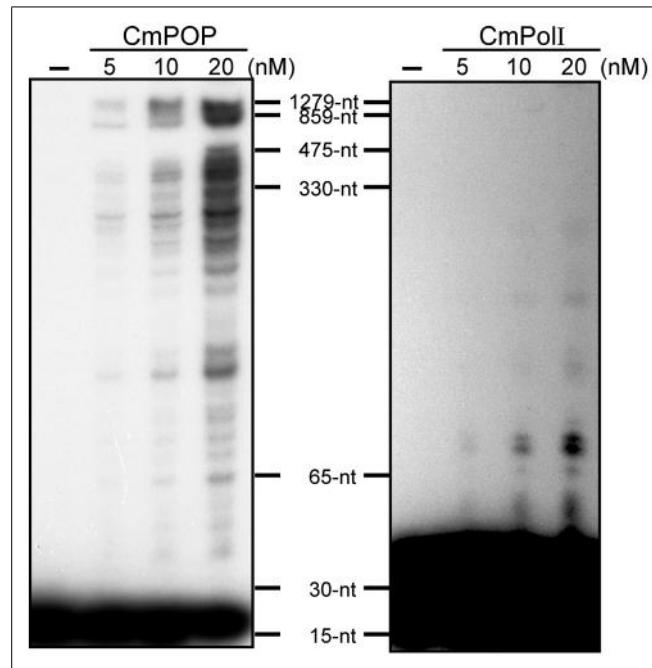


**FIGURE 1 |** Phylogenetic tree of Family A DNA polymerases. Reproduced from Moriyama et al. (2014) with permission.

severely inhibits the activity of POPs, with IC<sub>50</sub> values of 1–25  $\mu$ M (Moriyama et al., 2011). Because PAA at low concentrations does not inhibit other Family A DNA polymerases, such as PolI and Poly, sensitivity to PAA is a useful marker for the classification of organelle-localized DNA polymerases in unsequenced eukaryotes (Moriyama et al., 2011).

### 3'-5' exonuclease activity

Plant and protist organellar DNA polymerases have a 3'-5' exonuclease domain at the N-terminus consisting of three conserved regions, Exo I, Exo II, and Exo III. This exonuclease activity was shown in rice (Takeuchi et al., 2007) and *Cyanidioschyzon* (Moriyama et al., 2008). In rice POP, a mutant protein containing



**FIGURE 2 | Processivity of CmPOP and CmPolI.** Reactions were performed in reaction buffer (50 mM Tris-HCl [pH 8.0], 5 mM MgCl<sub>2</sub>, 1 mM DTT, 0.01% BSA, 50 μM dATP, 50 μM dTTP, 50 μM dGTP, 50 μM dCTP, and 10 nM <sup>32</sup>P-labeled M13mp18-M1) at 37°C for 5 min. The reaction products were analyzed in 5% polyacrylamide gels containing 7 M urea. Reproduced from Moriyama et al. (2008) with permission.

a replacement of Asp365 with Ala in the Exo II domain lost 3'-5' exonuclease activity. In terms of 3'-5' exonuclease proof-reading activity, rice POP exhibited a relatively high fidelity at a base substitution rate of 10<sup>-4</sup> to 10<sup>-5</sup> (Takeuchi et al., 2007).

#### EXPRESSION OF POP IN PLANTS

In tobacco BY-2 cells, the amount of POP transcripts and proteins increases at the initial phase of plastidial and mitochondrial DNA replication (Ono et al., 2007). The spatial expression patterns of POPs were analyzed in *A. thaliana* and rice by *in situ* hybridization, which revealed that POP genes are strongly expressed in the apical meristems of roots and shoots, leading to high POP protein levels in these tissues (Kimura et al., 2002; Mori et al., 2005). In *A. thaliana*, the expression of two POPs, AtPOP1 (At1g80840) and AtPOP2 (At3g20540), were compared by quantitative RT-PCR (Cupp and Nielsen, 2013). The analysis demonstrated that AtPOP1 is mainly expressed in rosette leaves, whereas AtPOP2 is predominantly found in the meristems of roots and shoots.

#### EXPRESSION OF POP IN ALGAE

The unicellular red alga *Cyanidioschyzon merolae* contains a single plastid and mitochondrion (Matsuzaki et al., 2004), whose division cycles are synchronized with the cell cycle. In synchronous cultures of *Cyanidioschyzon merolae* established using light-dark cycles (Suzuki et al., 1994), cells divide at ~12 h from the beginning of the light phase, and nuclear DNA increases at or just

before the M-phase. Replication of the mitochondrial genome appears to be at least partially synchronized with the cell cycle, as mitochondrial DNA begins to replicate from the light phase, and reaches a two-fold increase at or near the M-phase. In contrast, plastid DNA increases gradually throughout the entire cell cycle, even after cell division is complete (Moriyama et al., 2010). In contrast to land plants, which typically encode two or more copies of POP, *Cyanidioschyzon merolae* only has a single POP. The mRNA level of CmPOP changes during the cell cycle and reaches a peak that correlates with the rise in the mitotic index (Moriyama et al., 2008). However, the protein level of POP remains nearly unchanged throughout the cell cycle, with only small increases and decreases occurring during the light and dark phases, respectively. The observed expression of CmPOP is consistent with the results of organellar DNA replication during the cell cycle of *Cyanidioschyzon merolae*.

#### PHENOTYPES OF POP MUTANTS IN PLANTS

POP mutants of *A. thaliana* have been characterized by two research groups (Parent et al., 2011; Cupp and Nielsen, 2013). The *A. thaliana* genome encodes two POPs, AtPOP1 and AtPOP2, which are both localized to plastids and mitochondria. Double mutation of AtPOP1 and AtPOP2 was lethal, whereas each single mutant showed reduced DNA levels in both plastids and mitochondria. Additionally, the *Atpop2* mutant displayed high sensitivity to ciprofloxacin, an inducer of DNA double-strand breaks. These results indicate that two distinct POPs are involved in genome replication for plastids and mitochondria, and that AtPOP2 also functions in DNA repair in both organelles.

#### OTHER REPLICATION ENZYMES OF ORGANELLAR GENOMES

##### DNA PRIMASE AND HELICASE

DNA helicase unwinds dsDNA to allow DNA replication by DNA polymerase. In *E. coli*, primase synthesizes an RNA primer at the origin of replication in the leading strand and every ~1 kb in the lagging strand. TWINKLE (T7 gp4-like protein with intramitochondrial nucleoid localization), which is a homolog of the T7 phage gp4 protein with primase and helicase activities, was originally reported to function as a hexameric DNA helicase in human mitochondria (Spelbrink et al., 2001). In *A. thaliana*, TWINKLE functions as a DNA helicase and primase (Diray-Arce et al., 2013), and was shown to be localized to both chloroplasts and mitochondria by GFP-tagging experiments (Carrie et al., 2009). Dual-targeted enzymes to the mitochondria and chloroplasts of plants are summarized in the review by Carrie and Small (2013). In an assay using single-stranded M13 DNA as a template, a recombinant TWINKLE protein of *A. thaliana* (AtTWINKLE) showed ATP-dependent helicase and primase activities, synthesizing RNA primers of >15 nt that were then extended by *E. coli* PolI into high-molecular-weight DNA (Diray-Arce et al., 2013). The protein and mRNA of AtTWINKLE are mainly expressed in the meristem and young leaves, which is similar to the expression pattern of *A. thaliana* POPs, particularly AT3G20540. The *A. thaliana* genome encodes a second TWINKLE gene whose protein product only has the N-terminal primase domain of TWINKLE and is localized to chloroplasts,

according to unpublished data in the review by Cupp and Nielsen (2014).

Red algae also have a TWINKLE protein; however, it is localized to only mitochondria (Moriyama et al., 2014). Red algae and diatoms have a plastid-encoded DnaB helicase and a nucleus-encoded DnaG primase. In our analysis using GFP in *Cyanidioschyzon merolae*, DnaG was localized to the plastid. We also confirmed the plastid-localization of DnaG in the red alga *Porphyridium purpureum*. Based on these data, it appears that red algae and diatoms, the latter of which is thought to have originated from the secondary endosymbiosis with a red alga, utilize DnaB/DnaG in plastids and TWINKLE in mitochondria.

### DNA TOPOISOMERASE

*Arabidopsis thaliana* has a single gyrase A (AtGYRA) that is localized to both plastids and mitochondria, and has two gyrase B enzymes that are localized to either chloroplasts (AtGYRB1) or mitochondria (AtGYRB2; Wall et al., 2004). T-DNA insertion mutation of AtGYRA leads to an embryo-lethal phenotype, whereas T-DNA insertion mutations of both plastidial AtGYRB1 and mitochondrial AtGYRB2 result in seedling-lethal phenotypes (Wall et al., 2004). In *Nicotiana benthamiana*, virus-induced silencing of the genes encoding GYRA and GYRB resulted in abnormal nucleoid content and structure of chloroplasts and mitochondria (Cho et al., 2004). *A. thaliana* also encodes a gyrase B-like gene, GYRB3; however, it is recently reported that AtGYRB3 does not encode a gyrase subunit, as AtGYRB3 showed no supercoiling activity and did not interact with AtGYRA (Evans-Roberts et al., 2010). In addition to gyrases, plant organelles contain A-type topoisomerase I, which is a homolog of bacterial topoisomerase I (TopA). Based on localization analysis using a GFP-fusion protein, AtTOP1 was shown to be localized to both chloroplasts and mitochondria (Carrie et al., 2009).

The genome of *Cyanidioschyzon merolae* encodes genes for GYRA and GYRB, which are localized only to the plastid (Moriyama et al., 2014). Similarly, *Cyanidioschyzon merolae* TOP1 (type IA) is also localized only to the plastid. To search for mitochondrial topoisomerases, we examined the subcellular localization of topoisomerases encoded in the *Cyanidioschyzon merolae* genome, and showed that a homolog of eukaryotic TOP2 is targeted to mitochondria. To date, organellar localization of eukaryotic TOP2 has not been reported in plants. In *Cyanidioschyzon merolae*, the gyrase specific inhibitor nalidixic acid arrests not only replication of the plastid genome, but also that of the mitochondrial and nuclear genomes (Itoh et al., 1997; Kobayashi et al., 2009). The localization results of gyrases in *Cyanidioschyzon merolae* suggest that defective plastid replication leads to the arrest of mitochondrial and nuclear replication by a yet unknown mechanism.

### DNA LIGASE

DNA ligase is required for DNA replication, repair, and recombination, as it seals nicked-DNA ends of single-stranded breaks or joins DNA ends after double-stranded breaks. Four DNA ligases have been identified in the *A. thaliana* genome. *A. thaliana* DNA

ligase 1 (AtLIG1) is targeted to either the mitochondria or the nucleus when the gene transcript is translated from the first and second initiation codons, respectively (Sunderland et al., 2004, 2006). AtLIG1 is expressed in all tissues of *A. thaliana*, but higher transcript levels are found in young leaves and tissues containing meristem (Taylor et al., 1998). Plastid-targeting of AtLIG1 was not observed for any AtLIG1-GFP constructs translated from possible initiation codons, and the plastidial enzyme functioning as DNA ligase remains unclear. However, it has been noted that AtLIG6 has a putative plastid-targeting peptide at the N-terminus and might therefore be targeted to plastids (Sunderland et al., 2006).

*Cyanidioschyzon merolae* has a single gene encoding DNA ligase. *Cyanidioschyzon merolae* DNA ligase 1 (CmLIG1) has two methionine residues in its N-terminal region and is targeted to both mitochondria and plastids when the transcript is translated from the first and second initiation codons (Moriyama et al., 2014). In our analysis, no nuclear localization was observed when the N-terminal peptide of CmLIG1 was fused with GFP. However, the protein subcellular localization prediction software WolfPSORT (<http://wolfpsort.seq.cbrc.jp/>) detected a nuclear localization signal in CmLIG1. Therefore, CmLIG1 appears to have triple localization in plastids, mitochondria, and the nucleus.

### SINGLE-STRANDED DNA (ssDNA)-BINDING PROTEIN

An SSB, AtSSB1, was identified in *A. thaliana* (Edmondson et al., 2005). AtSSB1 is localized to mitochondria, but was also reported to be localized to chloroplasts in the review by Cupp and Nielsen (2014). AtSSB1 binds to ssDNA, but not to dsDNA, and stimulates RecA-mediated strand exchange activity.

Organellar ssDNA-binding proteins (OSBs) comprise the second class of SSBs in plants (Zaegel et al., 2006). OSBs have an SSB-like domain in the central region and one, two, or three C-terminal PDF motifs, which consist of 50-amino acids and are responsible for ssDNA binding. PDF motifs are conserved only in green plants, including *Chlamydomonas reinhardtii*. *A. thaliana* has four OSBs: AtOSB1 and 2 are localized to mitochondria and chloroplasts, respectively, whereas AtOSB3 is localized to both chloroplasts and mitochondria. AtOSB1 and AtOSB2 have been purified as recombinant proteins that showed preferential binding activity to ssDNA, as compared to dsDNA or RNA. Expression analysis of AtOSB1 using a  $\beta$ -glucuronidase (GUS) assay demonstrated that AtOSB1 is mainly expressed in gametophytic cells. T-DNA insertion mutation of the OSB1, OSB2, and OSB3 genes revealed that osb1 mutants accumulate homologous recombination products of mitochondrial DNA, whereas osb2 and osb3 mutants have no visible phenotype. These findings, together with the expression analysis for AtOSB1, indicate that AtOSB1 is involved in mitochondrial DNA recombination in gametophytic cells (Zaegel et al., 2006).

Replication protein A (RPA) is a nucleus-localized BBC in eukaryotes and is comprised of three subunits, RPA70, RPA32, and RPA14. The rice genome encodes three RPA70s, three RPA32s, and one RPA14. These RPA subunits combine in different variations to make three types of complexes: type A, B, and C. Among these

RPA complexes, the type A complex is localized to chloroplasts in rice (Ishibashi et al., 2006).

The *Cyanidioschyzon merolae* genome encodes a single gene for SSB, but does not contain a gene for OSB. In our analysis, the SSB of *Cyanidioschyzon merolae* is localized only in the mitochondrion, unlike that of *A. thaliana* (Moriyama et al., 2014). We performed the localization analysis using a construct starting from the second methionine codon or starting from the ATA codon located upstream of the first methionine codon; however, none of the constructs showed plastid localization. We also examined the organellar localization of RPAs in *Cyanidioschyzon merolae*, and even though they have no extension sequence at the N-terminus, they were localized to the nucleus. Based on these findings, the plastidial SSB in red algae remains unidentified.

### PRIMER REMOVAL ENZYME

In *E. coli*, RNA primers are removed by nick translation with the 5'-3' exonuclease and polymerase activities of DNA polymerase I (Langston et al., 2009; Sanyal and Doig, 2012). In contrast, RNaseH1 performs this role in human mitochondria (Kasiviswanathan et al., 2012). Although there are no reports of RNA primer removal enzymes that are specific to the organelles in green plants, two 5'-3' exonucleases (5'-3'EXO1 and 2) having sequence homology to the 5'-3' exonuclease domain of *E. coli* PolI are predicted to be localized to chloroplasts or mitochondria (Sato et al., 2003).

*Cyanidioschyzon merolae* has a gene with high sequence homology to bacterial PolII (Moriyama et al., 2008). The corresponding protein, CmPolII, contains 5'-3' exonuclease and polymerase domains. We demonstrated the plastid localization of CmPolII by immunoblotting and observation of a CmPolII-GFP fusion protein (Moriyama et al., 2008, 2014). Because CmPolII has low processive polymerase activity and no 3'-5' exonuclease activity, the enzyme appears to function in repair and primer removal by nick translation, similar to PolII in *E. coli*. We also showed that *Cyanidioschyzon merolae* RNase HII is localized to the mitochondrion, and that DNA2 nuclease/helicase and FEN1 are localized to the nucleus (Moriyama et al., 2014).

### PHYLOGENETIC DISTRIBUTION OF ORGANELLAR REPLICATIVE ENZYMES

#### ORIGIN OF ENZYMES RELATED TO ORGANELLAR GENOME REPLICATION

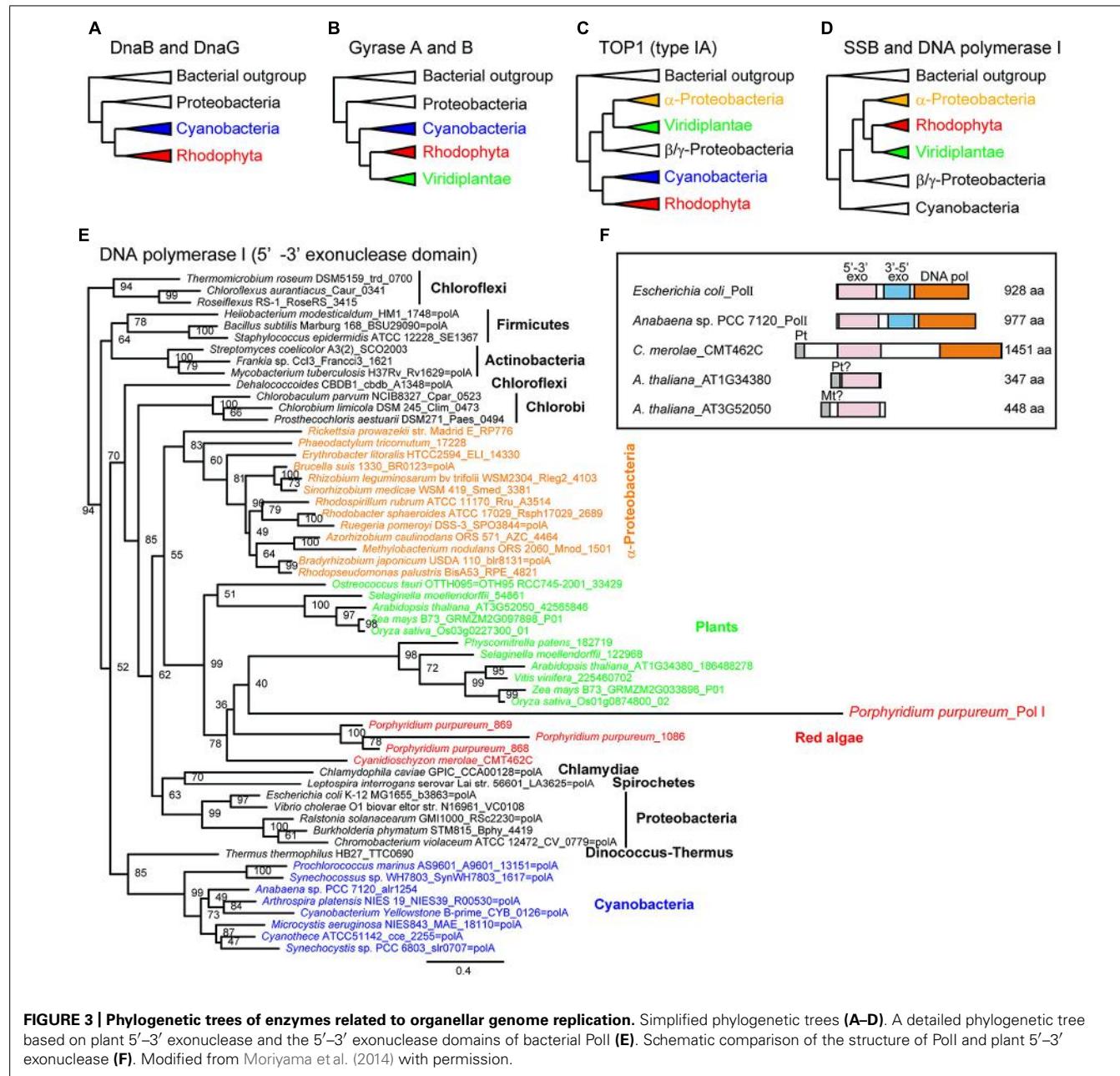
Phylogenetic analyses of bacterial-type replicative enzymes have been performed (Moriyama et al., 2014). We previously suggested that POP did not originate from the PolI of  $\alpha$ -proteobacteria or cyanobacteria (Figure 1, Moriyama et al., 2008). However, the analyses indicated that red algal DnaB helicase and DnaG primase originated from cyanobacteria (Figure 3A). Gyrase A and B also originated from cyanobacteria in both green plants and red algae (Figure 3B). Type IA topoisomerase was derived from  $\alpha$ -proteobacteria in green plants and from cyanobacteria in red algae (Figure 3C). SSB and 5'-3' exonuclease/PolII originated from  $\alpha$ -proteobacteria in green plants and red algae (Figure 3D). With regard to PolI, as green plants have enzymes with 5'-3' exonuclease domain, but lack 3'-5' exonuclease and DNA polymerase domains (Figure 3F), phylogenetic analysis of the 5'-3' exonuclease domain in bacteria and photosynthetic eukaryotes was

performed (Figures 3D,E). These results suggest that the organelle replication apparatus of both green plants and red algae is composed of enzymes of various origins, including  $\alpha$ -proteobacteria, cyanobacteria, and eukaryotes.

### REPERTOIRE OF ENZYMES RELATED TO ORGANELLAR GENOME REPLICATION IN PLANTS AND ALGAE

The enzymes related to organelar DNA replication and recombination in a species of angiosperm, fern, moss, filamentous terrestrial alga, two green algae, and two red algae are listed in Table 1. The proteins conserved in all examined species are POP, TWINKLE, Gyrases, type IA-TOP1, TOP2, and LIG1. DnaB and DnaG are conserved only in red algae. The retention of SSBs is highly variable in photosynthetic eukaryotes. Bacterial-type SSB proteins are conserved in land plants and *Cyanidioschyzon merolae*, whereas OSB proteins are conserved among land plants, including *A. thaliana*, *Physcomitrella patens*, and *Klebsormidium flaccidum*. It was reported that OSB proteins contain a few PDF motifs in addition to an SSB-like domain (Zaegel et al., 2006). According to this classification, *Physcomitrella patens* and *K. flaccidum* have a single OSB, which contains one and two PDF motifs, respectively, in addition to the SSB-like domain. RECA and Whirly (WHY) are recombination-related proteins and are also not uniformly conserved in photosynthetic eukaryotes. For example, *Selaginella moellendorffii* and *Porphyridium purpureum* do not have RECA, red algae do not encode WHY, and *Physcomitrella patens* has neither RECA nor WHY. Conservation of origin-binding protein (ODB) is more limited, as only land plants have this protein. PolI containing 5'-3' exonuclease and DNA polymerase domains is retained only in red algae. However, a protein with the 5'-3' exonuclease domain is found in photosynthetic eukaryotes, with the exception of *Cyanidioschyzon merolae*. Therefore, all photosynthetic eukaryotes contain proteobacteria-derived PolI. RNase H having high homology to RNase HII of *Cyanidioschyzon merolae* is conserved in most plants and algae, with the notable exception of *A. thaliana*. The observed distribution of enzymes that play key roles in replication indicates that they are essentially conserved in all plants and algae. In contrast, because recombination-related enzymes and SSBs are non-uniformly distributed among plants and algae, these enzymes are considered to exhibit high plasticity during evolution.

Based on the presence of enzymes related to organelar genome replication in plant genomes, we propose a model for the substitution of these enzymes in photosynthetic eukaryotes (Figure 4). In this model, the ancestor of photosynthetic eukaryotes, prior to its divergence into green and red lineages, contained POP, TWINKLE, DnaB, DnaG, gyrases, type-IA TOP1, TOP2, PolI, RNase HII, SSB, and LIG1. Several of these enzymes, including POP, TWINKLE, gyrases, type-IA TOP1, TOP2, LIG1, and PolI (or 5'-3' exonuclease), were retained by all plants. In the early green lineage, WHY was obtained, and mitochondrial TWINKLE became dually localized in mitochondria and plastids. However, the dual-localization of TWINKLE resulted in the loss of DnaB and DnaG in the green lineage. After the divergence of Chlorophyta, the ancestor of land plants obtained the SSBs, OSB, and ODB. The acquisition of these SSBs, in addition to WHY, in the



**FIGURE 3 | Phylogenetic trees of enzymes related to organellar genome replication.** Simplified phylogenetic trees (A–D). A detailed phylogenetic tree based on plant 5'-3' exonuclease and the 5'-3' exonuclease domains of bacterial PolI (E). Schematic comparison of the structure of PolI and plant 5'-3' exonuclease (F). Modified from Moriyama et al. (2014) with permission.

green lineage may have partially contributed to exploration of terrestrial habitats, as the high recombinant/repair activity of these enzymes would have potentially allowed the repair of DNA damaged by direct sunlight. The sets of RECA and WHY, and RECA and OSB were lost in Bryophyta and Pteridophyta, respectively, followed by the loss of RNase HII (homolog of *Cyanidioschyzon merolae* CMT626C) in Angiospermae. In red algae, most replication enzymes in the ancestor of photosynthetic eukaryotes are found in present-day species.

## CONCLUDING REMARKS

In the past decade, most enzymes related to plastid and mitochondrial DNA replication in plants and algae have been identified.

These studies have revealed that the core enzymes and components involved in replication are identical in the plastids and mitochondria of land plants. Because the nuclear genomes of green plants and algae encode these core replicative enzymes, such as POP, TWIN-KLE, gyrases, TOP1, TOP2, LIG1, and 5'-3' exonuclease, which frequently contain putative dual-targeting sequences at the N-terminus (Table 1), it is presumed that the green lineage contains a similar set of plastid and mitochondrial enzymes. In contrast, SSBs and recombination-related enzymes are not universally conserved in the green lineage, suggesting that these enzymes are possibly susceptible to exchange or loss during evolution, leading to the acquisition or creation of species-specific enzymes. Unlike the green lineage, red algae contain different replicative protein

**Table 1 |** List of replication-related enzymes possibly localized to plastids or mitochondria in photosynthetic eukaryotes.

| Function         | Angiospermae            |  | Pteridophyta                      |                              | Bryophyta                      |                                  | Charophyta                      |                                | Chlorophyta                   |                                | Rhodophyta |  |
|------------------|-------------------------|--|-----------------------------------|------------------------------|--------------------------------|----------------------------------|---------------------------------|--------------------------------|-------------------------------|--------------------------------|------------|--|
|                  | Protein name            | <i>Arabidopsis thaliana</i>              | <i>Selaginella moellendorffii</i> | <i>Physcomitrella patens</i> | <i>Klebsormidium flaccidum</i> | <i>Chlamydomonas reinhardtii</i> | <i>Ostreococcus lucimarinus</i> | <i>Cyanidioschyzon merolae</i> | <i>Porphyridiun purpureum</i> |                                |            |  |
| DNA polymerase   | POP                     | AT1G50840 ( <b>pt, mt</b> )              | 82310 <sup>a</sup>                | 008G080400 (pt, mt)          | kf00710_0020 (pt, mt)          | g8521 (pt, mt)                   | 34610 <sup>a</sup>              | CML0270C ( <b>pt, mt</b> )     | contig2347.14 (pt, mt)        |                                |            |  |
|                  |                         | AT3G20540 ( <b>pt, mt</b> )              |                                   | 014G054900 (pt, mt)          |                                |                                  |                                 |                                |                               |                                |            |  |
| Primase/helicase | TWNKLE                  | AT1G30680 ( <b>pt, mt</b> )              | 164801 <sup>a</sup>               | 024G060900                   | 012G092800 <sup>a</sup>        | kf00614_0010 (pt, mt)            | Cre12.g503150 (pt)              | 94562 <sup>a</sup>             | CMT452C ( <b>mt</b> )         | contig653.2 (partial)          |            |  |
|                  |                         | AT1G30660 (primase domain) ( <b>pt</b> ) |                                   |                              |                                | Cre12.g514950 (pt, mt)           |                                 |                                |                               | contig653.4 (partial)          |            |  |
| Primase          | DnaG                    | N.D.                                     | N.D.                              | N.D.                         | N.D.                           | N.D.                             | N.D.                            | N.D.                           | CML0286C ( <b>pt</b> )        | contig2059.10 (pt)             |            |  |
|                  |                         | DnaB                                     | N.D.                              | N.D.                         | N.D.                           | N.D.                             | N.D.                            | N.D.                           | CMV098C                       | contig622.1 (encoded by ptDNA) |            |  |
| Topoisomerase    | GYRA                    | AT3G10690 ( <b>pt, mt</b> )              | 111960 (pt)                       | 008G071000 (pt, mt)          | kf00194_0180 (pt, mt)          | Cre04.g22.1650 (pt)              | 28121 (pt, mt)                  | CMS243C ( <b>pt</b> )          | contig445.3 (pt)              |                                |            |  |
|                  |                         | GYRB                                     | AT3G10270 ( <b>pt</b> )           | 405784 (pt, mt)              | 019G017500                     | kf00564_0010 (mt)                | Cre10.g440750 (pt, mt)          | 30508 (pt, mt)                 | CMH166C ( <b>pt</b> )         | contig2132.10 (pt)             |            |  |
| TOP1             | AT5G04130 ( <b>mt</b> ) | 022G014300 (pt)                          | 123748 <sup>a</sup>               | 023G038200 <sup>a</sup>      | kf00696_0070 (pt, mt)          | Cre10.g442850 (pt, mt)           | 41359 (pt, mt)                  | CML252C ( <b>pt</b> )          | contig3627.1 (pt)             |                                |            |  |
| (type I)         | TOP2                    | AT3G23890 (pt, nuc)                      | 98330 (nuc)                       | 018G068300 (nuc)             | kf00053_0310 (nuc)             | Cre01.g009250 (nuc)              | 41836 (nuc)                     | CMB013C ( <b>nuc</b> )         | contig2019.13 (nuc)           |                                |            |  |
|                  |                         |  | 127267 (nuc)                      | 019G011100 (nuc)             | kf00219_0040 (pt, mt)          |                                  |                                 | CML330C ( <b>mt</b> )          | contig3477.2 (mt)             |                                |            |  |

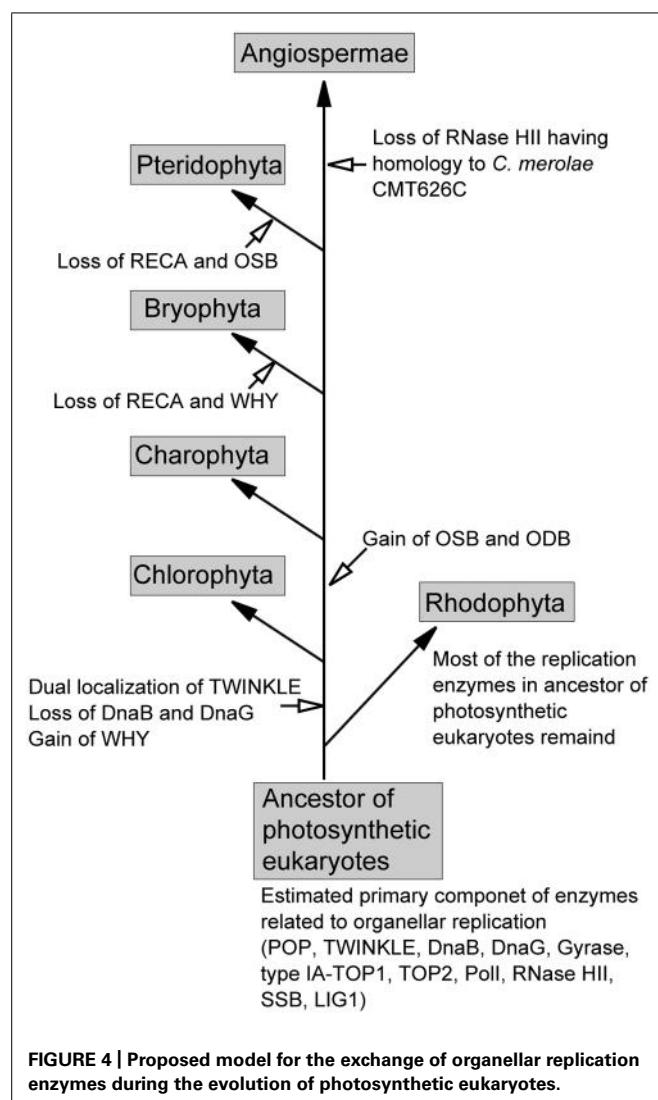
(Continued)

Table 1 | Continued

| Function               | Angiospermae     |                              | Pteridophyta                      |                              | Bryophyta                               |                                  | Charophyta                      |                                | Chlorophyta                    |      | Rhodophyta |  |
|------------------------|------------------|------------------------------|-----------------------------------|------------------------------|---|----------------------------------|---------------------------------|--------------------------------|--------------------------------|------|------------|--|
|                        | Protein name     | <i>Arabidopsis thaliana</i>  | <i>Selaginella moellendorffii</i> | <i>Physcomitrella patens</i> | <i>Klebsormidium flaccidum</i>          | <i>Chlamydomonas reinhardtii</i> | <i>Ostreococcus lucimarinus</i> | <i>Cyanidioschyzon merolae</i> | <i>Porphyrithium purpureum</i> |      |            |  |
| ssDNA binding protein  | SSB              | AT3G18580 (mt)               | 115961 <sup>a</sup>               | 005G033200 (mt, nuc)         | N.D.                                    | N.D.                             | 10587 <sup>a</sup>              | CML135C (mt)                   | N.D.                           |      |            |  |
|                        |                  | AT4G11060 ( <b>pt, mt</b> )  | 440287 (pt, mt)                   | 006G0532001 <sup>a</sup>     | 016G055400 (pt, mt)                     | kf00106_0250 (pt, mt)            | N.D.                            | N.D.                           | N.D.                           |      |            |  |
|                        | OSB              | AT1G31010 ( <b>mt</b> )      | N.D.                              | 021G006400 (pt, mt)          |   |                                  |                                 |                                |                                |      |            |  |
|                        |                  | AT1G47720 ( <b>mt</b> )      |                                   |                              |   |                                  |                                 |                                |                                |      |            |  |
|                        |                  | AT4G20010 ( <b>pt</b> )      |                                   |                              |   |                                  |                                 |                                |                                |      |            |  |
|                        |                  | AT5G44785 ( <b>pt, mt</b> )  |                                   |                              |   |                                  |                                 |                                |                                |      |            |  |
| Ligase                 | LG1              | AT1G08130 ( <b>mt, nuc</b> ) | 96803 <sup>a</sup>                | 013G068900 (pt)              | kf00242_0010 (none)                     | g7530 (nuc)                      | 16988 (none)                    | CML235C ( <b>pt, mt</b> )      | contig730.1 (pt, mt)           |      |            |  |
|                        |                  | AT1G49250 (pt, nuc)          | 97073 <sup>a</sup>                |                              | kf00698_0030 (none)                     | g6320 (pt, mt)                   |                                 |                                |                                |      |            |  |
|                        |                  | AT1G34380 (pt)               | 46231 <sup>a</sup>                | 001G092400 (pt)              | kf00249_0020 <sup>a</sup>               | g4419 <sup>a</sup>               | 5068 <sup>a</sup>               | N.D.                           | contig3646.3 (pt, mt)          |      |            |  |
| Primer removal         | 5'-3'EXO         | AT3G52050 (mt)               | 46232 <sup>a</sup>                | 016G072100 (pt, mt)          | kf00340_0150 <sup>a</sup>               |                                  |                                 |                                | contig3646.4 (pt, mt)          |      |            |  |
|                        |                  |                              | 54861 <sup>a</sup>                |                              |   |                                  |                                 |                                |                                |      |            |  |
|                        |                  | 105035 <sup>a</sup>          |                                   |                              |   |                                  |                                 |                                |                                |      |            |  |
|                        |                  | 105035 <sup>a</sup>          |                                   |                              |   |                                  |                                 |                                |                                |      |            |  |
|                        |                  | 416415 (pt)                  |                                   |                              |   |                                  |                                 |                                |                                |      |            |  |
|                        | PoII (5'-3')     | N.D.                         | N.D.                              | N.D.                         | N.D.                                    | N.D.                             | N.D.                            | CML1462C ( <b>pt</b> )         | contig2156.11 (nuc)            |      |            |  |
|                        | EXO & PoI)       |                              |                                   |                              |   |                                  |                                 |                                |                                |      |            |  |
|                        | RNase            | N.D.                         | 91138 <sup>a</sup>                | 002G083400 (pt, mt)          | kf00183_0070 (pt, mt)                   | Cre13_g561900 (pt, mt)           | 7887 <sup>a</sup>               | CML626C ( <b>mt</b> )          | contig4525.1 (pt, mt)          |      |            |  |
|                        | HII <sup>b</sup> |                              |                                   |                              |   |                                  |                                 |                                |                                |      |            |  |
|                        | RECA             | AT1G79050 ( <b>pt</b> )      | N.D.                              | N.D.                         | kf00358_0070 (pt, mt)                   | Cre07_g314650 (pt)               | 26240 <sup>a</sup>              | CML061C (pt, mt)               | N.D.                           |      |            |  |
|                        |                  | AT2G19490 ( <b>pt, mt</b> )  |                                   |                              | kf00493_0080 (pt, mt)                   |                                  |                                 |                                |                                |      |            |  |
|                        |                  | AT3G10140 ( <b>mt</b> )      |                                   |                              |   |                                  |                                 |                                |                                |      |            |  |
|                        |                  | AT3G32920 ( <b>mt</b> )      |                                   |                              |   |                                  |                                 |                                |                                |      |            |  |
| Recombination mediator | WHY              | AT1G14410 ( <b>pt</b> )      | 84703 (nuc)                       | N.D.                         | kf00637_0100 (pt, mt)                   | Cre02_g091550 (pt, mt)           | 8400 <sup>a</sup>               | N.D.                           | N.D.                           | N.D. | N.D.       |  |
|                        |                  | AT1G71260 ( <b>mt</b> )      |                                   |                              |   |                                  |                                 |                                |                                |      |            |  |
|                        | ODB              | AT1G71310 ( <b>mt, nuc</b> ) | 68768 <sup>a</sup>                | 001G130700 (pt)              | kf00130_0080 <sup>a</sup> (pt, mt, nuc) | N.D.                             | N.D.                            | N.D.                           | N.D.                           | N.D. | N.D.       |  |
|                        |                  | AT5G47870 ( <b>pt, nuc</b> ) |                                   | 015G028700 (pt, nuc)         |   |                                  |                                 |                                |                                |      |            |  |

Enzymes were estimated by BLAST searches at the websites of Phytozome v9.1 (<http://www.phytozome.net/>), Klebsormidium flaccidum Genome Project ([http://www.plantmolphogenesis.bio.titech.ac.jp/~algae\\_genome\\_project/klebsormidium/index.html](http://www.plantmolphogenesis.bio.titech.ac.jp/~algae_genome_project/klebsormidium/index.html)), and Porphyridium purpureum genome project (<http://cyanophora.rutgers.edu/porporphyrithium/>), and protein IDs are shown. The results of subcellular localization reported in *A. thaliana* and *Cyanidioschyzon* are indicated by **pt** (plastids), **mt** (mitochondria), or **nuc** (nucleus) in bold in parentheses, and references are indicated in the body text. Localization of enzymes in other species was estimated by the subcellular-targeting estimation programs TargetP (<http://www.cbs.dtu.dk/services/TargetP/>) and WolfPSORT (<http://wolfpsort.org/>).

<sup>a</sup>The obtained protein sequence did not appear to encode for a full-length protein containing an initiation methionine residue; <sup>b</sup>RNase HII with high homology to *C. merolae* protein (CMT626C); N.D., not detected.



profiles in plastids and mitochondria. Red algal plastids contain numerous replication proteins that originated from cyanobacteria (Moriyama et al., 2014), suggesting that the mechanism of genome replication in these plastids might be similar to that found in bacteria.

To date, a number of organelle-localized enzymes have been identified. However, biochemical data are lacking for the majority of organellar replication enzymes in plants. The role of an enzyme predicted by homology searches against known enzymes might differ from its actual function or properties. For example, soybean plastid DNA replication ODB shares homology with bacterial and *A. thaliana* SSB, but only binds to dsDNA of the *oriA* sequence in plastid DNA, and not to ssDNA (Lassen et al., 2011).

The regulatory mechanisms controlling the initiation of plant organellar genome replication and the number of organellar DNA copies remains to be explored. Recently, chloroplast DNA replication was shown to be regulated by the cellular redox state in the green alga *Chlamydomonas reinhardtii* (Kabeya and Miyagishima, 2013). Specifically, chloroplast DNA replication was activated

and inactivated by the addition of reducing and oxidative agents, respectively, in both *in vivo* and *in vitro* assays. Light-dependent genome replication was also reported in cyanobacteria, in which DCMU [3-(3,4-dichlorophenyl)-1,1-dimethylurea], an inhibitor of electron transport between the PSII complex and plastoquinone pool, inhibits DNA replication initiation, and DBMIB (2,5-dibromo-3-isopropyl-6-methyl-p-benzoquinone), an inhibitor of electron transport between plastoquinone and cytochrome *b6f* complex, inhibits the initiation and elongation of replication (Watanabe et al., 2012; Ohbayashi et al., 2013). Thus, the light-mediated replication of plastid DNA in algae may have originated from cyanobacteria. However, organellar replication in land plants and multicellular plants appears to be regulated by other mechanisms. In land plants, the replication of organellar genomes is restricted to meristematic tissues, and is not associated with the cycle or organellar division (Hashimoto and Possingham, 1989; Fujie et al., 1993). These findings suggest that land plants have more complex regulatory mechanisms controlling the replication of organellar genomes than those operating in algae.

## ACKNOWLEDGMENTS

This work was supported in part by Core Research for Evolutional Science and Technology (CREST) from the Japan Science and Technology Agency (JST), and Grants-in-Aid for Young Scientists (B) from the Japan Society for the Promotion of Science (JSPS; no. 25870155).

## REFERENCES

- Arnold, J., Smidansky, E., Moustafa, I., and Cameron, C. (2012). Human mitochondrial RNA polymerase: structure-function, mechanism and inhibition. *Biochim. Biophys. Acta* 1819, 948–960. doi: 10.1016/j.bbaparam.2012.04.002
- Carrie, C., Kühn, K., Murcha, M., Duncan, O., Small, I., O’Toole, N., et al. (2009). Approaches to defining dual-targeted proteins in *Arabidopsis*. *Plant J.* 57, 1128–1139. doi: 10.1111/j.1365-313X.2008.03745.x
- Carrie, C., and Small, I. (2013). A reevaluation of dual-targeting of proteins to mitochondria and chloroplasts. *Biochim. Biophys. Acta* 1833, 253–259. doi: 10.1016/j.bbamcr.2012.05.029
- Chavalitshewinkoon-Petrmitr, P., Chawprom, S., Naesens, L., Balzarini, J., and Wilairat, P. (2000). Partial purification and characterization of mitochondrial DNA polymerase from *Plasmodium falciparum*. *Parasitol. Int.* 49, 279–288. doi: 10.1016/S1383-5769(00)00057-X
- Cho, H. S., Lee, S. S., Kim, K. D., Hwang, I., Lim, J., Park, Y., et al. (2004). DNA gyrase is involved in chloroplast nucleoid partitioning. *Plant Cell* 16, 2665–2682. doi: 10.1105/tpc.104.024281
- Christensen, A., Lyznik, A., Mohammed, S., Elowsky, C., Elo, A., Yule, R., et al. (2005). Dual-domain, dual-targeting organellar protein presequences in *Arabidopsis* can use non-AUG start codons. *Plant Cell* 17, 2805–2816. doi: 10.1105/tpc.105.035287
- Cupp, J. D., and Nielsen, B. L. (2013). *Arabidopsis thaliana* organellar DNA polymerase IB mutants exhibit reduced mtDNA levels with a decrease in mitochondrial area density. *Physiol. Plant.* 149, 91–103. doi: 10.1111/pl.12009
- Cupp, J. D., and Nielsen, B. L. (2014). Minireview: DNA replication in plant mitochondria. *Mitochondrion* doi: 10.1016/j.mito.2014.03.008 [Epub ahead of print].
- Diray-Arce, J., Liu, B., Cupp, J. D., Hunt, T., and Nielsen, B. L. (2013). The *Arabidopsis* At1g30680 gene encodes a homologue to the phage T7 gp4 protein that has both DNA primase and DNA helicase activities. *BMC Plant Biol.* 13:36. doi: 10.1186/1471-2229-13-36
- Edmondson, A. C., Song, D., Alvarez, L. A., Wall, M. K., Almond, D., McClellan, D. A., et al. (2005). Characterization of a mitochondrially targeted single-stranded DNA-binding protein in *Arabidopsis thaliana*. *Mol. Genet. Genomics* 273, 115–122. doi: 10.1007/s00438-004-1106-5

- Evans-Roberts, K. M., Breuer, C., Wall, M. K., Sugimoto-Shirasu, K., and Maxwell, A. (2010). *Arabidopsis thaliana* GYRB3 does not encode a DNA gyrase subunit. *PLoS ONE* 5:e9899. doi: 10.1371/journal.pone.0009899
- Fujie, M., Kuroiwa, H., Kawano, S., and Kuroiwa, T. (1993). Studies on the behavior of organelles and their nucleoids in the root apical meristem of *Arabidopsis thaliana* (L.) Col. *Planta* 189, 443–452. doi: 10.1007/BF00194444
- Gualberto, J. M., Milesina, D., Wallet, C., Niazi, A. K., Weber-Lotfi, F., and Dietrich, A. (2014). The plant mitochondrial genome: dynamics and maintenance. *Biochimie* 100, 107–120. doi: 10.1016/j.biochi.2013.09.016
- Hashimoto, H., and Possingham, J. (1989). DNA levels in dividing and developing plastids in expanding primary leaves of *Avena sativa*. *J. Exp. Bot.* 40, 257–262. doi: 10.1093/jxb/40.2.257
- Holmes, A. M. (1981). Studies on the inhibition of highly purified calf thymus 8S and 7.3S DNA polymerase alpha by aphidicolin. *Nucleic Acids Res.* 9, 161–168. doi: 10.1093/nar/9.1.161
- Ishibashi, T., Kimura, S., and Sakaguchi, K. (2006). A higher plant has three different types of RPA heterotrimeric complex. *J. Biochem.* 139, 99–104. doi: 10.1093/jb/mvj014
- Itoh, R., Takahashi, H., Toda, K., Kuroiwa, H., and Kuroiwa, T. (1997). DNA gyrase involvement in chloroplast-nucleoid division in *Cyanidioschyzon merolae*. *Eur. J. Cell Biol.* 73, 252–258.
- Kabeya, Y., and Miyagishima, S. (2013). Chloroplast DNA replication is regulated by the redox state independently of chloroplast division in *Chlamydomonas reinhardtii*. *Plant Physiol.* 161, 2102–2112. doi: 10.1104/pp.113.216291
- Kaguni, L. (2004). DNA polymerase gamma, the mitochondrial replicase. *Annu. Rev. Biochem.* 73, 293–320. doi: 10.1146/annurev.biochem.72.121801.161455
- Kasiviswanathan, R., Collins, T., and Copeland, W. (2012). The interface of transcription and DNA replication in the mitochondria. *Biochim. Biophys. Acta* 1819, 970–978. doi: 10.1016/j.bbagr.2011.12.005
- Kimura, S., Uchiyama, Y., Kasai, N., Namekawa, S., Saotome, A., Ueda, T., et al. (2002). A novel DNA polymerase homologous to *Escherichia coli* DNA polymerase I from a higher plant, rice (*Oryza sativa* L.). *Nucleic Acids Res.* 30, 1585–1592. doi: 10.1093/nar/30.7.1585
- Kobayashi, Y., Kanesaki, Y., Tanaka, A., Kuroiwa, H., Kuroiwa, T., and Tanaka, K. (2009). Tetrapyrrole signal as a cell-cycle coordinator from organelle to nuclear DNA replication in plant cells. *Proc. Natl. Acad. Sci. U.S.A.* 106, 803–807. doi: 10.1073/pnas.0804270105
- Korhonen, J., Pham, X., Pellegrini, M., and Falkenberg, M. (2004). Reconstitution of a minimal mtDNA replisome in vitro. *EMBO J.* 23, 2423–2429. doi: 10.1038/sj.emboj.7600257
- Kornberg, A., and Baker, T. A. (1992). *DNA Replication*, 2nd Edn. New York: W. H. Freeman and Company.
- Kühn, K., Bohne, A., 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. doi: 10.1105/tpc.106.046839
- Langston, L. D., Indiani, C., and O'Donnell, M. (2009). Whither the replisome: emerging perspectives on the dynamic nature of the DNA replication machinery. *Cell Cycle* 8, 2686–2691. doi: 10.4161/cc.8.17.9390
- Lassen, M. G., Kochhar, S., Kocchar, S., and Nielsen, B. L. (2011). Identification of a soybean chloroplast DNA replication origin-binding protein. *Plant Mol. Biol.* 76, 463–471. doi: 10.1007/s11103-011-9736-6
- Lecrenier, N., and Foury, F. (2000). New features of mitochondrial DNA replication system in yeast and man. *Gene* 246, 37–48. doi: 10.1016/S0378-1119(00)00107-4
- Leinbach, S. S., Reno, J. M., Lee, L. F., Isbell, A. F., and Boezi, J. A. (1976). Mechanism of phosphonoacetate inhibition of herpesvirus-induced DNA polymerase. *Biochemistry* 15, 426–430. doi: 10.1021/bi00647a029
- Maréchal, A., and Brisson, N. (2010). Recombination and the maintenance of plant organelle genome stability. *New Phytol.* 186, 299–317. doi: 10.1111/j.1469-8137.2010.03195.x
- Matsuzaki, M., Misumi, O., Shin-I, T., Maruyama, S., Takahara, M., Miyagishima, S. Y., et al. (2004). Genome sequence of the ultrasmall unicellular red alga *Cyanidioschyzon merolae* 10D. *Nature* 428, 653–657. doi: 10.1038/nature02398
- Mori, Y., Kimura, S., Saotome, A., Kasai, N., Sakaguchi, N., Uchiyama, Y., et al. (2005). Plastid DNA polymerases from higher plants, *Arabidopsis thaliana*. *Biochem. Biophys. Res. Commun.* 334, 43–50. doi: 10.1016/j.bbrc.2005.06.052
- Moriyama, T., and Sato, N. (2013). “The plant and protist organellar DNA replication enzyme POP showing up in place of DNA polymerase gamma may be a suitable antiprotozoal drug target,” in *The Mechanisms of DNA Replication*, ed. D. D. Stuart (Rijeka: InTech), 287–311.
- Moriyama, T., Tajima, N., Sekine, K., and Sato, N. (2014). Localization and phylogenetic analysis of enzymes related to organellar genome replication in the unicellular rhodophyte *Cyanidioschyzon merolae*. *Genome Biol. Evol.* 6, 228–237. doi: 10.1093/gbe/evu009
- Moriyama, T., Terasawa, K., Fujiwara, M., and Sato, N. (2008). Purification and characterization of organellar DNA polymerases in the red alga *Cyanidioschyzon merolae*. *FEBS J.* 275, 2899–2918. doi: 10.1111/j.1742-4658.2008.06426.x
- Moriyama, T., Terasawa, K., and Sato, N. (2011). Conservation of POPs, the plant organellar DNA polymerases, in eukaryotes. *Protist* 162, 177–187. doi: 10.1016/j.protis.2010.06.001
- Moriyama, T., Terasawa, K., Sekine, K., Toyoshima, M., Koike, M., Fujiwara, M., et al. (2010). Characterization of cell-cycle-driven and light-driven gene expression in a synchronous culture system in the unicellular rhodophyte *Cyanidioschyzon merolae*. *Microbiology* 156, 1730–1737. doi: 10.1099/mic.0.037754-0
- Nielsen, B. L., Cupp, J. D., and Brammer, J. (2010). Mechanisms for maintenance, replication, and repair of the chloroplast genome in plants. *J. Exp. Bot.* 61, 2535–2537. doi: 10.1093/jxb/erq163
- Ohbayashi, R., Watanabe, S., Kaneko, Y., Narikawa, R., Chibazakura, T., Ikeuchi, M., et al. (2013). DNA replication depends on photosynthetic electron transport in cyanobacteria. *FEMS Microbiol. Lett.* 344, 138–144. doi: 10.1111/1574-6968.12166
- Ono, Y., Sakai, A., Takechi, K., Takio, S., Takusagawa, M., and Takano, H. (2007). NtPolI-like1 and NtPolI-like2, bacterial DNA polymerase I homologs isolated from BY-2 cultured tobacco cells, encode DNA polymerases engaged in DNA replication in both plastids and mitochondria. *Plant Cell Physiol.* 48, 1679–1692. doi: 10.1093/pcp/pcm140
- Parent, J., Lepage, E., and Brisson, N. (2011). Divergent roles for the two PolI-like organelle DNA polymerases of *Arabidopsis*. *Plant Physiol.* 156, 254–262. doi: 10.1104/pp.111.173849
- Parsons, P., and Simpson, M. V. (1967). Biosynthesis of DNA by isolated mitochondria: incorporation of thymidine triphosphate-2-C-14. *Science* 155, 91–93. doi: 10.1126/science.155.3758.91
- Sakai, A. (2001). In vitro transcription/DNA synthesis using isolated organelle-nuclei: application to the analysis of the mechanisms that regulate organelle genome function. *J. Plant Res.* 114, 199–211. doi: 10.1007/PL00013984
- Sakai, A., Suzuki, T., Nagata, N., Sasaki, N., and Miyazawa, Y. (1999). Comparative analysis of DNA synthesis activity in plastid-nuclei and mitochondrial-nuclei simultaneously isolated from cultured tobacco cells. *Plant Sci.* 140, 9–19. doi: 10.1016/S0168-9452(98)00207-6
- Sanyal, G., and Doig, P. (2012). Bacterial DNA replication enzymes as targets for antibacterial drug discovery. *Expert Opin. Drug Discov.* 7, 327–339. doi: 10.1517/17460441.2012.660478
- Sato, N., Terasawa, K., Miyajima, K., and Kabeya, Y. (2003). Organization, developmental dynamics, and evolution of plastid nucleoids. *Int. Rev. Cytol.* 232, 217–262. doi: 10.1016/S0074-7696(03)32006-6
- Shiraki, K., Okuno, T., Yamanishi, K., and Takahashi, M. (1989). Phosphonoacetic acid inhibits replication of human herpesvirus-6. *Antiviral Res.* 12, 311–318. doi: 10.1016/0166-3542(89)90058-2
- Shutt, T., and Gray, M. (2006). Twinkle, the mitochondrial replicative DNA helicase, is widespread in the eukaryotic radiation and may also be the mitochondrial DNA primase in most eukaryotes. *J. Mol. Evol.* 62, 588–599. doi: 10.1007/s00239-005-0162-8
- Spelbrink, J. N., Li, F. Y., Tiranti, V., Nikali, K., Yuan, Q. P., Tariq, M., et al. (2001). Human mitochondrial DNA deletions associated with mutations in the gene encoding Twinkle, a phage T7 gene 4-like protein localized in mitochondria. *Nat. Genet.* 28, 223–231. doi: 10.1038/90058
- Spencer, D., and Whitfield, P. R. (1967). DNA synthesis in isolated chloroplasts. *Biochem. Biophys. Res. Commun.* 28, 538–542. doi: 10.1016/0006-291X(67)90347-6
- Sunderland, P. A., West, C. E., Waterworth, W. M., and Bray, C. M. (2004). Choice of a start codon in a single transcript determines DNA ligase 1 isoform production and intracellular targeting in *Arabidopsis thaliana*. *Biochem. Soc. Trans.* 32, 614–616. doi: 10.1042/BST0320614

- Sunderland, P. A., West, C. E., Waterworth, W. M., and Bray, C. M. (2006). An evolutionarily conserved translation initiation mechanism regulates nuclear or mitochondrial targeting of DNA ligase I in *Arabidopsis thaliana*. *Plant J.* 47, 356–367. doi: 10.1111/j.1365-313X.2006.02791.x
- Suzuki, K., Ehara, T., Osafune, T., Kuroiwa, H., Kawano, S., and Kuroiwa, T. (1994). Behavior of mitochondria, chloroplasts and their nuclei during the mitotic cycle in the ultramicroalga *Cyanidioschyzon merolae*. *Eur. J. Cell Biol.* 63, 280–288.
- Takeuchi, R., Kimura, S., Saotome, A., and Sakaguchi, K. (2007). Biochemical properties of a plastidial DNA polymerase of rice. *Plant Mol. Biol.* 64, 601–611. doi: 10.1007/s11003-007-9179-2
- Taylor, R. M., Hamer, M. J., Rosamond, J., and Bray, C. M. (1998). Molecular cloning and functional analysis of the *Arabidopsis thaliana* DNA ligase I homologue. *Plant J.* 14, 75–81. doi: 10.1046/j.1365-313X.1998.00094.x
- Tewari, K. K., and Wildman, S. G. (1967). DNA polymerase in isolated tobacco chloroplasts and nature of the polymerized product. *Proc. Natl. Acad. Sci. U.S.A.* 58, 689–696. doi: 10.1073/pnas.58.2.689
- Wall, M., Mitchenall, L., and Maxwell, A. (2004). *Arabidopsis thaliana* DNA gyrase is targeted to chloroplasts and mitochondria. *Proc. Natl. Acad. Sci. U.S.A.* 101, 7821–7826. doi: 10.1073/pnas.0400836101
- Wanrooij, S., and Falkenberg, M. (2010). The human mitochondrial replication fork in health and disease. *Biochim. Biophys. Acta* 1797, 1378–1388. doi: 10.1016/j.bbabi.2010.04.015
- Wanrooij, S., Fusté, J., Farge, G., Shi, Y., Gustafsson, C., and Falkenberg, M. (2008). Human mitochondrial RNA polymerase primes lagging-strand DNA synthesis in vitro. *Proc. Natl. Acad. Sci. U.S.A.* 105, 11122–11127. doi: 10.1073/pnas.0805399105
- Watanabe, S., Ohbayashi, R., Shiwa, Y., Noda, A., Kanesaki, Y., Chibazakura, T., et al. (2012). Light-dependent and asynchronous replication of cyanobacterial multi-copy chromosomes. *Mol. Microbiol.* 83, 856–865. doi: 10.1111/j.1365-2958.2012.07971.x
- Weiser, T., Gassmann, M., Thömmes, P., Ferrari, E., Hafkemeyer, P., and Hübscher, U. (1991). Biochemical and functional comparison of DNA polymerases alpha, delta, and epsilon from calf thymus. *J. Biol. Chem.* 266, 10420–10428.
- Williams, A. J., Wernette, C. M., and Kaguni, L. S. (1993). Processivity of mitochondrial DNA polymerase from *Drosophila* embryos. Effects of reaction conditions and enzyme purity. *J. Biol. Chem.* 268, 24855–24862.
- Wintersberger, E. (1966). Occurrence of a DNA-polymerase in isolated yeast mitochondria. *Biochem. Biophys. Res. Commun.* 25, 1–7. doi: 10.1016/0006-291X(66)90630-9
- Zaegel, V., Guermann, B., Le Ret, M., Andrés, C., Meyer, D., Erhardt, M., et al. (2006). The plant-specific ssDNA binding protein OSB1 is involved in the stoichiometric transmission of mitochondrial DNA in *Arabidopsis*. *Plant Cell* 18, 3548–3563. doi: 10.1105/tpc.106.042028

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 22 July 2014; paper pending published: 16 August 2014; accepted: 30 August 2014; published online: 17 September 2014.

Citation: Moriyama T and Sato N (2014) Enzymes involved in organellar DNA replication in photosynthetic eukaryotes. *Front. Plant Sci.* 5:480. doi: 10.3389/fpls.2014.00480

This article was submitted to Plant Physiology, a section of the journal *Frontiers in Plant Science*.

Copyright © 2014 Moriyama and Sato. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Nuclear-encoded factors associated with the chloroplast transcription machinery of higher plants

**Qing-Bo Yu<sup>1,2</sup>, Chao Huang<sup>1,2</sup> and Zhong-Nan Yang<sup>1,2\*</sup>**

<sup>1</sup> Department of Biology, College of Life and Environmental Sciences, Shanghai Normal University, Shanghai, China

<sup>2</sup> Institute for Plant Gene Function, Department of Biology, Shanghai Normal University, Shanghai, China

**Edited by:**

Thomas Pfannschmidt, University Joseph Fourier Grenoble, France

**Reviewed by:**

Jin Chen, Michigan State University, USA

Takashi Shiina, Kyoto Prefectural University, Japan

**\*Correspondence:**

Zhong-Nan Yang, Department of Biology, College of Life and Environmental Sciences, Shanghai Normal University, No.100,

Rd. GuiLin, Shanghai 200234, China  
e-mail: znyang@shnu.edu.cn

Plastid transcription is crucial for plant growth and development. There exist two types of RNA polymerases in plastids: a nuclear-encoded RNA polymerase (NEP) and plastid-encoded RNA polymerase (PEP). PEP is the major RNA polymerase activity in chloroplast. Its core subunits are encoded by the plastid genome, and these are embedded into a larger complex of nuclear-encoded subunits. Biochemical and genetics analysis identified at least 12 proteins are tightly associated with the core subunit, while about 34 further proteins are associated more loosely generating larger complexes such as the transcriptionally active chromosome (TAC) or a part of the nucleoid. Domain analyses and functional investigations suggested that these nuclear-encoded factors may form several functional modules that mediate regulation of plastid gene expression by light, redox, phosphorylation, and heat stress. Genetic analyses also identified that some nuclear-encoded proteins in the chloroplast that are important for plastid gene expression, although a physical association with the transcriptional machinery is not observed. This covers several PPR proteins including CLB19, PDM1/SEL1, OTP70, and YS1 which are involved in the processing of transcripts for PEP core subunit as well as AtECB2, Prin2, SVR4-Like, and NARA5 that are also important for plastid gene expression, although their functions are unclear.

**Keywords:** plastid transcription, RNA polymerase, PEP, NEP, functional modules

## INTRODUCTION

Plastids are specific organelles in plant and algal cells that are responsible for photosynthesis and some important metabolic pathways. They possess their own genetic material and are generally considered to be of endosymbiotic origin (McFadden and van Dooren, 2004). Similar to bacteria, the DNA is organized into dense particles, the nucleoids (Pfalz and Pfannschmidt, 2013). The genome size from plastids of vascular plants ranges from 120 to 180 kbp and the encoded gene set is highly conserved (Sugiura, 1992). They can be categorized into three groups according to their molecular function of the encoded components: (1) Components of the plastid gene expression machinery (RNA polymerase, ribosomal proteins, tRNAs, and rRNAs); (2) Subunits of photosynthesis-related complexes (Rubisco, PSII, the cytochrome b6f complex, PSI, NAPH dehydrogenase, and ATP synthase), and (3) a few proteins involved in other processes (e.g., ClpP1 and YCF3) (Sugiura, 1992). The chloroplast proteome is estimated to be between 2100 and 3600 proteins (Leister, 2003). Most of the chloroplast proteins are encoded by the nuclear genome and are imported from the cytosol (Li and Chiu, 2010), due to the limited coding capacity of the chloroplast genome. However, chloroplast gene expression is still essential for the development of chloroplasts and the maintenance of chloroplast functions. It involves the action of numerous nuclear-encoded factors, besides proteins encoded by the plastome. Recently, proteomics data (Pfannschmidt et al., 2000; Ogrzewalla et al., 2002;

Suzuki et al., 2004; Pfalz et al., 2006; Steiner et al., 2011; Melonek et al., 2012) and genetic analysis (Chi et al., 2008; Ogawa et al., 2009; Wu and Zhang, 2010; Qiao et al., 2011, 2013; Kindgren et al., 2012; Pyo et al., 2013; Yu et al., 2013) identified that numerous nuclear-encoded proteins with various functions are associated with the transcriptional machinery and are involved in chloroplast gene expression. In this paper, we focused on these nuclear-encoded factors for chloroplast transcription.

## TWO TYPES OF PLASTID RNA POLYMERASES IN HIGHER PLANTS

Plastid genes are transcribed by two RNA polymerases, the nuclear-encoded RNA polymerase (NEP) and the plastid-encoded RNA polymerase (PEP). NEP is a phage-type RNA polymerase with a single subunit (Chang et al., 1999; Lerbs-Mache, 2011). In *Arabidopsis*, the nuclear genome encodes three NEPs. RpoTp is targeted to chloroplast, RpoTm is targeted to mitochondria, and RpoTmp is dually targeted to both organelles (Hess and Borner, 1999). NEP is important for plant development. Inactivation of RpoTp results in defects in plastid gene expression and leaf development (Hricová et al., 2006; Swiatecka-Hagenbruch et al., 2008) while plants with inactivated RpoTmp exhibit several defects, including a plastid gene expression defect, delayed greening and growth retardation of leaves and roots (Courtois et al., 2007). The dysfunction of both NEPs resulted in seedling lethality at a very early developmental stage (Hricová et al., 2006). Although NEP is generally considered to be a single

subunit RNA polymerase, recent biochemical analysis revealed that RPOTmp interacts with a thylakoid RING-H2 protein. This protein might mediate the fixation of RPOTmp to thylakoid membranes in order to regulate the transcription of the plastid *rrn* genes (Azevedo et al., 2008).

PEP is composed of four core subunits encoded by the genes *rpoA*, *rpoB*, *rpoC1*, and *rpoC2* that are located on the plastid genome. PEP exhibits a certain sensitivity to inhibitors of bacterial transcription, such as tagetitoxin, and the group of rifampicin-related drugs, indicating a distinct degree of conservation of these eubacterial-type RNA polymerase during evolution (Liere et al., 2011). Like for bacterial RNA polymerases, the activity/specificity of the PEP core enzyme is regulated by sigma-like transcription factors that are encoded by the nuclear genome of higher plants. In *Arabidopsis*, there exist six chloroplast sigma factors (SIG1–SIG6). These sigma factors might have overlapping as well as specific functions for recognizing a specific set of promoters during chloroplast development (Schweer, 2010; Liere et al., 2011). Besides the sigma factors, however, the core subunits of PEP are associated also with additional proteins (see below) that mediate a number of additional functions to the PEP complex.

NEP and PEP play different roles in plastid gene transcription during plastid development and plant growth (Liere et al., 2011). Based on their transcription by the different RNA polymerases, plastid genes can be grouped into three classes (Hajdukiewicz et al., 1997; Ishizaki et al., 2005). Transcription of photosynthesis-related genes (such as *psbA*, *psbD*, and *rbcL*) depend largely on PEP (class I), whereas a few house-keeping genes (mostly encoding components of the transcription/translation apparatus, such as *rpoB*) are exclusively transcribed by NEPs (class III). Most of plastid genes, however, are transcribed by both PEP and NEPs (class II). Generally, NEP is more active in the young, non-green tissues early in leaf development. It transcribes housekeeping genes including the four core subunits of PEP polymerase which primarily constitute the plastid gene expression machinery. Once PEP is formed in later developmental stages, it thereafter transcribes the photosynthesis-related genes (Hajdukiewicz et al., 1997; Lopez-Juez and Pyke, 2005; Schweer et al., 2010b) and plastid tRNAs (Williams-Carrier et al., 2014). In the mature chloroplast, the activity of NEP is barely detected, while PEP activity maintains high for chloroplast development and plant growth. Nevertheless, recent investigations demonstrated that both NEP and PEP are present in seeds, and PEP is also important for seed germination. This indicates that PEP exists also in non-photosynthetically active seed plastids (Demarsy et al., 2006).

#### PEP IS ASSOCIATED WITH NUMEROUS NUCLEAR-ENCODED PROTEINS

Early biochemical analysis demonstrated that two different forms of the PEP complex exist in higher plant, that is, PEP-A and PEP-B (Pfannschmidt and Link, 1994). PEP-B is composed only of the *rpo* core subunits and is present in both etioplasts and greening chloroplasts. During light-dependent chloroplast development, this PEP-B enzyme is reconfigured into an eukaryote-like enzyme complex, the PEP-A, by association of numerous proteins (Pfannschmidt and Link, 1997; Steiner et al., 2011; Pfalz and

Pfannschmidt, 2013). PEP-A is the major RNA polymerase in matured chloroplast of higher plant. Attempts have been focused on the isolation of the plastid RNA polymerase complex and its associated proteins for many years (Pfalz and Pfannschmidt, 2013). Biochemical analyses uncovered that the core *rpo* subunits of PEP are present in both the insoluble RNA polymerase preparation called transcriptionally active chromosome (TAC), and the soluble RNA polymerase preparation (sRNAP) (Krause and Krupinska, 2000; Pfalz et al., 2006; Melonek et al., 2012). The TAC fraction was isolated from lysed plastids through one or two gel filtration chromatography steps and subsequent ultracentrifugation, while the soluble RNA polymerase (sRNAP) is prepared from isolated and lysed plastids via several chromatographic purification steps without precipitation by centrifugation (Pfalz and Pfannschmidt, 2013). Based on gel filtration and mass spectrometry analysis from different organisms, including *Nicotiana tabacum* (Suzuki et al., 2004), *spinach* (Melonek et al., 2012), *mustard* (*Sinapis alba*) (Pfannschmidt et al., 2000; Pfalz et al., 2006; Steiner et al., 2011), and *Arabidopsis* (Pfalz et al., 2006) it is estimated that the TAC complex contains 43 nuclear-encoded proteins (Table 1). Ten proteins were reproducibly found to be tightly associated with PEP core subunits in *mustard* seedlings and, therefore, were named polymerase-associated proteins (PAPs) (Steiner et al., 2011). The other proteins were found in the previous reported TAC complex and might represent more loosely attached components of the transcription machinery (Pfalz et al., 2006). Two TAC components, pTAC7 (Yu et al., 2013) and MurE-like (Garcia et al., 2008), were not identified as PAPs in *mustard* (Steiner et al., 2011), however, based on their mutant phenotype in T-DNA inactivation mutants of *Arabidopsis* these two proteins were proposed to be PAPs (Pfalz and Pfannschmidt, 2013). One essential common feature of all PAPs is that they are essential for PEP activity. The *Arabidopsis* knock-out lines for the corresponding genes show all an albino/ivory or pale-green phenotype with severe defects in chloroplast development and PEP-dependent transcription (Table 1) (Pfalz et al., 2006; Garcia et al., 2008; Myouga et al., 2008; Arsova et al., 2010; Schröter et al., 2010; Gao et al., 2011; Steiner et al., 2011; Gilkerson et al., 2012; Yagi et al., 2012; Yu et al., 2013). The phenotype of these PAP mutants is identical to that of *rpo*-gene knock-out mutants in tobacco (Allison et al., 1996; Hajdukiewicz et al., 1997; De Santis-Maclossek et al., 1999). In the knockout mutants of AtECB1/SVR4/MRL7 (Qiao et al., 2011; Yu et al., 2014), PEP-Related Development Arrested 1 (PRDA1) (Qiao et al., 2013), and Delayed Greening 1 (DG1) (Chi et al., 2008), the expression of PEP-dependent chloroplast genes is also severely reduced. These proteins have not been identified in PEP complex by previous proteomic analyses (Krause and Krupinska, 2000; Suzuki et al., 2004; Pfalz et al., 2006; Steiner et al., 2011). Nevertheless, they interact with some members of the PEP/TAC complex (Chi et al., 2010; Qiao et al., 2011, 2013; Kindgren et al., 2012; Yu et al., 2014) and are either loosely or temporarily attached.

Based on proteomic analysis and protein interaction investigation, the TAC complex contains at least 50 proteins of which 46 are nuclear-encoded (Tables 1, 2). These nuclear-encoded proteins can be classified into several groups including

**Table 1 | Proteomics Identification of Chloroplast PEP complex Components in Higher Plant.**

| Gene name                 | AGI number | Species*                    | Localization**                       | Molecular phenotype****               | Protein domain***   | Molecular function****   | References  |
|---------------------------|------------|-----------------------------|--------------------------------------|---------------------------------------|---|--|---|
| PAP1/pTAC3 <sup>a</sup>   | AT3G04260  | <i>Arabidopsis/ Mustard</i> | Chloroplast nucleoid <sup>b</sup>    | Low PEP activity                      | SAP domain  | DNA binding ability  | Pfälz et al., 2006; Steiner et al., 2011; Yagi et al., 2012                           |
| PAP2/pTAC2 <sup>a</sup>   | AT1G74850  | <i>Arabidopsis/ Mustard</i> | Chloroplast nucleoid <sup>b</sup>    | Low PEP activity                      | pentatricopeptide repeat protein S1 domain                      | Unknown  | Pfälz et al., 2006; Steiner et al., 2011  |
| PAP3/pTAC10 <sup>a</sup>  | AT3G48500  | <i>Arabidopsis/ Mustard</i> | Chloroplast nucleoid <sup>b</sup>    | Low PEP activity                      |   | RNA binding ability  | Pfälz et al., 2006; Steiner et al., 2011; Jeon et al., 2012                           |
| PAP4/FSD3 <sup>a</sup>    | AT5G23310  | <i>Arabidopsis/ Mustard</i> | Chloroplast nucleoid <sup>b</sup>    | Low PEP activity                      | Iron superoxide dismutase                                       | SOD enzymes activity   | Pfälz et al., 2006; Miyuga et al., 2008; Steiner et al., 2011                         |
| PAP5/pTAC12 <sup>a</sup>  | AT2G34640  | <i>Arabidopsis/ Mustard</i> | Chloroplast and Nucleus <sup>b</sup> | Low PEP activity and Low PEP activity | Structurally similar to the RAD23 pfkB-type carbohydrate kinase | Protein degradation in nuclear and unknown function in chloroplast Unknown | Pfälz et al., 2006; Chen et al., 2010; Steiner et al., 2011                           |
| PAP6/FLN1 <sup>a</sup>    | At3g54090  | <i>Arabidopsis/ Mustard</i> | Chloroplast nucleoid <sup>b</sup>    | Low PEP activity                      | pfkB-type carbohydrate kinase                                   |  | Pfälz et al., 2006; Steiner et al., 2011; Gilkerson et al., 2012                      |
| PAP7/pTAC14 <sup>a</sup>  | AT4G20130  | <i>Arabidopsis/ Mustard</i> | Chloroplast <sup>b</sup>             | Low PEP activity                      | SET domain  | Unknown  | Pfälz et al., 2006; Gao et al., 2011; Steiner et al., 2011                            |
| PAP8/pTAC6 <sup>a</sup>   | AT1G21600  | <i>Arabidopsis/ Mustard</i> | Chloroplast <sup>b</sup>             | Low PEP activity                      | PHB_acc_N   | Unknown  | Pfälz et al., 2006; Steiner et al., 2011  |
| PAP9/FSD1 <sup>a</sup>    | AT5G51100  | <i>Arabidopsis/ Mustard</i> | Chloroplast nucleoid <sup>b</sup>    | Low PEP activity                      | Iron superoxide dismutase                                       | SOD enzymes activity   | Pfälz et al., 2006; Miyuga et al., 2008; Steiner et al., 2011                         |
| PAP10/Trx Z <sup>a</sup>  | AT3G06730  | <i>Arabidopsis/ Mustard</i> | Chloroplast stroma <sup>b</sup>      | Low PEP activity                      | Thioredoxin   | disulfide reductase activity <i>in vitro</i>                               | Pfälz et al., 2006; Arsova et al., 2010   |
| PAP11/AtMurE <sup>a</sup> | AT1G63680  | <i>Arabidopsis/ Mustard</i> | Chloroplast <sup>b</sup>             | Low PEP activity                      | Murilagase family protein                                       | Unknown  | Pfälz et al., 2006; Garcia et al., 2008   |
| pTAC7/pTAC12 <sup>a</sup> | AT5G24314  | <i>Arabidopsis/ Mustard</i> | Chloroplast <sup>b</sup>             | Low PEP activity                      | Unknown   | Unknown  | Pfälz et al., 2006; Yu et al., 2013   |
| FLN2 <sup>a</sup>         | AT1G69200  | <i>Arabidopsis/ Mustard</i> | Chloroplast nucleoid <sup>b</sup>    | Low PEP activity                      | pfkB-type carbohydrate kinase                                   | Unknown  | Pfälz et al., 2006; Gilkerson et al., 2012; Huang et al., 2013                        |
| WHY1/pTAC1                | AT1G14410  | <i>Arabidopsis/ Mustard</i> | Chloroplast and Nucleus <sup>b</sup> |                                       | Chloroplast genome rearrangement                                | Chloroplast genome rearrangement   | Pfälz et al., 2006; Maréchal et al., 2009   |
| WHY3/pTAC11               | AT2G02240  | <i>Arabidopsis/ Mustard</i> | Chloroplast and Nucleus <sup>b</sup> |                                       | Chloroplast genome rearrangement                                | Maintaining chloroplast genome stability                                   | Pfälz et al., 2006; Maréchal et al., 2009   |
| pTAC4/Vipp1               | AT1G65260  | <i>Arabidopsis/ Mustard</i> | Chloroplast <sup>b</sup>             |                                       | Albino phenotype and Thylakoid biogenesis defects               | Unknown  | Pfälz et al., 2006  |
| pTAC8                     | AT2G46820  | <i>Arabidopsis/ Mustard</i> | Chloroplast <sup>b</sup>             |                                       | Thylakoid biogenesis  | Curt1 domain   | Thylakoid biogenesis  |
| pTAC5                     | AT4G13670  | <i>Arabidopsis/ Mustard</i> | Chloroplast nucleoid <sup>b</sup>    | Low PEP activity under heat stress    | Peptidoglycan binding domain                                    | Zinc-dependent disulfide isomerase activity                                | Pfälz et al., 2006; Armbruster et al., 2013<br>Pfälz et al., 2006; Zhong et al., 2013 |

(Continued)

**Table 1 | Continued**

| Gene name | AGI number | Species*                              | Localization**           | Molecular phenotype**** | Protein domain***  | Molecular function****                             | References   |
|-----------|------------|---------------------------------------|--------------------------|-------------------------|--|--|--|
| pTAC17    | AT1G80480  | <i>Arabidopsis/</i><br><i>Mustard</i> | Chloroplast              | N.A.                    | CobW domain-containing   | N.A.   | Pfälz et al., 2006                                     |
| pTAC18    | AT2G32180  | <i>Arabidopsis/</i><br><i>Mustard</i> | Chloroplast              | N.A.                    | Unknown domain   | N.A.   | Pfälz et al., 2006                                     |
| pTAC16    | AT3G46780  | <i>Arabidopsis/</i><br><i>Mustard</i> | Chloroplast              | No Phenotype            | adh_short, Epimerase(1)<br>NmxA(1) SPT2(1)                           | DNA binding protein                                | Pfälz et al., 2006                                     |
| pTAC13    | AT3G09210  | <i>Arabidopsis/</i><br><i>Mustard</i> | Chloroplast              | N.A.                    | KOW and NusG domain  | N.A.   | Pfälz et al., 2006                                     |
| PTAC9     | AT4G20010  | <i>Arabidopsis/</i><br><i>Mustard</i> | Chloroplast              | N.A.                    | Single-stranded DNA<br>binding                                       | N.A.   | Pfälz et al., 2006                                     |
| PTAC15    | AT5G54180  | <i>Arabidopsis/</i><br><i>Mustard</i> | Chloroplast <sup>b</sup> | No Phenotype            | Mitochondrial  | N.A.   | Pfälz et al., 2006                                     |
| GyrB      | AT3G10270  | <i>Arabidopsis/</i><br><i>Mustard</i> | Chloroplast              | Etiolated cytoledons    | transcription termination<br>factor                                  | DNA_gyraseB,<br>DNA_gyraseB_C,<br>HATPase_c_Toprim | Pfälz et al., 2006; Pfälz and<br>Pfannschmidt, 2013    |
| GyrA      | AT3G10690  | <i>Arabidopsis/</i><br><i>Mustard</i> | Chloroplast              | Embryo lethal           | DNA_gyraseA_C;<br>DNA_topoisoiV                                      | N.A.   | Pfälz et al., 2006; Pfälz and<br>Pfannschmidt, 2013    |
| PolA      | AT3G20540  | <i>Arabidopsis/</i><br><i>Mustard</i> | Chloroplast              | N.A.                    | 3_5_exonuc, DNA_pol_A  | N.A.   | Pfälz et al., 2006                                     |
| rPL12     | AT3G27830  | <i>Arabidopsis/</i><br><i>Mustard</i> | Chloroplast              | N.A.                    | Ribosomal_L12  |  | Pfälz et al., 2006                                     |
| RABE1B    | AT4G20360  | <i>Arabidopsis/</i><br><i>Mustard</i> | Chloroplast              | N.A.                    | ATP_bind; CbiA; cobW;<br>GTP_EFTU;<br>GTP_EFTU_D2;                   | N.A.   | Pfälz et al., 2006                                     |
| rPL29     | AT5G65220  | <i>Arabidopsis/</i><br><i>Mustard</i> | Chloroplast              | N.A.                    | GTP_EFTU_D3; Miro;<br>Ribosomal protein L29                          | N.A.   | Pfälz et al., 2006                                     |
| PTK/cpCK2 | AT2G23070  | <i>Arabidopsis/</i><br><i>Mustard</i> | Chloroplast              | N.A.                    | CK2 kinase domain  | Protein kinase activity                            | Loschelder et al., 2004                                |
| CSP41b    | AT1G09340  | <i>Arabidopsis/</i><br><i>Mustard</i> | Chloroplast              | Pale green seedlings    | RNA Binding  | Chloroplast ribosomal RNA<br>metabolism            | Pfannschmidt et al., 2000;<br>Ogrzewalla et al., 2002; |
| CSP41A    | AT3G63140  | <i>Arabidopsis/</i><br><i>Mustard</i> | Chloroplast              | No Phenotype            | N.A.   | N.A.   | Loschelder et al., 2004                                |
| Atann4    | AT2G38750  | <i>Arabidopsis/</i><br><i>Mustard</i> | Chloroplast              | N.A.                    | Calcium ion<br>binding,calcium-<br>dependent phospholipid<br>binding | Zim17-type zinc finger<br>protein                  | Pfannschmidt et al., 2000;<br>Loschelder et al., 2004  |
| ETCHED1   | AT1G68730  | <i>Maize</i>                          | Chloroplast <sup>b</sup> | N.A.                    |  | SWIB/MDM2 domain                                   | da Costa e Silva et al., 2004                          |
|           | AT2G35605  | <i>Spinach</i>                        | Chloroplast <sup>b</sup> | N.A.                    |  | N-CoR and TFIIB DNA<br>binding domains             | Melonek et al., 2012                                   |
|           | AT2g02060  | <i>Spinach</i>                        | Chloroplast              | N.A.                    |  |  | Melonek et al., 2012                                   |

(Continued)

**Table 1 | Continued**

| Gene name | AGI number | Species* | Localization**           | Molecular phenotype*** | Protein domain****                          | Molecular function***** | References           |
|-----------|------------|----------|--------------------------|------------------------|---|-------------------------|----------------------|
| AT3G14320 | AT3G14320  | Spinach  | Chloroplast              | N.A.                   | AP2 DNA binding                             | N.A.                    | Melonek et al., 2012 |
| AT2G27430 | AT2G27430  | Spinach  | Chloroplast              | N.A.                   | Armadillo-type fold, armadillo-like helical | N.A.                    | Melonek et al., 2012 |
| AT5G02320 | AT5G02320  | Spinach  | Chloroplast              | N.A.                   | N-Cor and TFIIB DNA binding domains         | N.A.                    | Melonek et al., 2012 |
| AT5G36780 | AT5G36780  | Spinach  | Chloroplast              | N.A.                   | RecF/RecN/SMC                               | N.A.                    | Melonek et al., 2012 |
| MFP1      | AT3G16000  | Spinach  | Chloroplast <sup>b</sup> | N.A.                   | N-terminal domain<br>DNA binding            | N.A.                    | Melonek et al., 2012 |

\* Species that chloroplast proteomics experiments were performed are indicated.

\*\* Localization information is from GFP-fusion data and/or chloroplast proteomics data/immune analysis.

\*\*\* Protein domain information is from PDB database.

\*\*\*\* Molecular function data is given based on the reference. N.A. means that its detailed molecular function remains unclear.

\*\*\*\*\* The phenotypes of the knockout lines in *Arabidopsis* are indicated. N.A. means that the phenotype remains unclear.

<sup>a</sup> These factors are essential for PEP activity.

<sup>b</sup> Localization information is from individual GFP-fusion experiment or immune analysis.

DNA/RNA binding proteins, thioredoxin proteins, kinases, ribosome proteins and proteins with unknown function (**Table 1**). Yeast two-hybrid and other biochemical assays revealed the relationship of some proteins in the PEP complex (**Figure 1**). The interactions between these PAPs are consistent with the biochemical experiments that identified these proteins in the PEP complex under the stringent condition (Steiner et al., 2011). Currently, proteins directly interacting with the PEP core subunits have not been identified in the PEP complex. Immunoprecipitation analysis demonstrated that pTAC3 is associated with the *rpo* subunits (Yagi et al., 2012). However, the direct interaction between pTAC3 and PEP core subunits has not been verified.

## PROTEINS IN THE PEP COMPLEX WITH DNA/RNA BINDING DOMAIN

The eukaryotic transcriptional machinery consists of RNA polymerases and various DNA binding proteins, such as transcription factors. These DNA-binding proteins recognize the promoter to regulate downstream gene transcription. In the TAC complex, there are at least 14 proteins with DNA-binding domains (**Table 1**) (Pfalz et al., 2006; Steiner et al., 2011; Pfalz and Pfannschmidt, 2013). pTAC3 belongs to the SAP protein family. The *ptac3* mutant exhibits an albino phenotype with reduced PEP-dependent plastid transcription. It is unclear yet if pTAC3 can bind to a specific DNA region in order to regulate plastid gene transcription (Yagi et al., 2012). pTAC6 is essential for chloroplast transcription (Pfalz et al., 2006) since the expression of the *psbA* gene was barely detectable in the *ptac6* mutant, compared with that in *ptac2* and *ptac12* (Pfalz et al., 2006). It is likely that pTAC6 is a specific regulator for *psbA* (Pfalz et al., 2006), however, to date its function remains enigmatic. In bacteria, there exist two transcription termination mechanisms; Rho-independent transcription termination and Rho-dependent termination. The mitochondrial transcription termination factor (mTERF) family was identified to regulate mitochondrial gene expression including transcription termination (Kleine, 2012). pTAC15 is a member of the mTERF protein family (Pfalz et al., 2006). Whether it can terminate the transcription of PEP-dependent plastid genes needs to be verified.

The TAC complex contains at least six RNA-binding proteins including ZmWhy1, pTAC10, the elongation factor EF-Tu, and three ribosomal proteins, S3, L12-A, and L26 (**Table 1**). Whirly proteins belong to a small nuclear transcription factor family commonly found in plants. In *Arabidopsis*, pTAC1/AtWhy1 and pTAC11/AtWhy3 can bind DNA (Xiong et al., 2009). They are required to maintain the stability of the plastid genome (Maréchal et al., 2009). The whirly 1 ortholog in maize (ZmWHY1/pTAC1) can bind both RNA and DNA, and co-immuno-precipitated with chloroplast RNA splicing 1 (CRS1) (Prikryl et al., 2008). pTAC10 contains a S1 domain and has RNA binding activity in tobacco (Jeon et al., 2012), and it may be one substrate of chloroplast-target casein kinase 2 (cpCK2) (Reiland et al., 2009). The phosphorylation of pTAC10 may affect its RNA binding. The detailed function of the elongation factor EF-Tu and the ribosomal proteins S3, L12-A, and L26 in chloroplast is not reported. The existence of these RNA-binding proteins, however, suggests that there exists a translation subdomain in the TAC/nucleoid.

**Table 2 | Individual mutant analysis identifies several factors which affect PEP-dependent chloroplast transcription in higher plant.**

| Gene name                                    | AGI number | Phenotype*    | Localization**       | Molecular phenotype*                                       | Protein domain***             | Molecular function**** | References                                |
|--|------------|---------------|----------------------|--|-------------------------------|------------------------|---|
| DG1 <sup>a</sup>                             | AT5G67570  | Delayed green | Chloroplast          | Low PEP activity   | PPR domain                    | Interacts with sgg6    | Chi et al., 2008; 2010                    |
| AtECB1/MRL7/SVR4 <sup>a</sup>                | AT4G28590  | Albino        | Chloroplast nucleoid | Low PEP activity   | Thioredoxin-Like fold         | Thioredoxin activity   | Qiao et al., 2011; Yu et al., 2014;       |
| PDRA1 <sup>a</sup>                           | AT5G48470  | Yellowish     | Chloroplast nucleoid | Low PEP activity   | U.K.                          | Function unknown       | Qiao et al., 2013                         |
| NARA5 <sup>b</sup>                           | AT4G27600  | Yellowish     | Chloroplast          | Low PEP activity   | pfkB-type carbohydrate kinase | U.K.                   | Ogawa et al., 2009                        |
| MRL7-Like/SVR4-Like/AtECB1-Like <sup>b</sup> | AT2G31840  | Albino        | Chloroplast stroma   | Low PEP activity   | Thioredoxin-like fold         | U.K.                   | Qiao et al., 2011; Powkowska et al., 2014 |
| Prn2 <sup>b</sup>                            | AT1G10522  | Yellowish     | Chloroplast nucleoid | Low PEP activity   | U.K.                          | RNA editing            | Kindgren et al., 2012                     |
| AtECB2/NAC1 <sup>b</sup>                     | AT1G15510  | Albino        | Chloroplast          | Low PEP activity, accD and Ndhf editing defect             | PPR domain                    | RNA editing            | Yu et al., 2009; Tseng et al., 2010       |
| OTP70 <sup>c</sup>                           | AT4G25270  | Pale green    | Chloroplast          | Low PEP activity, splicing of the plastid transcript rpoC1 | PPR domain                    | RNA splicing           | Chateigner-Boutin et al., 2011            |
| YS1 <sup>c</sup>                             | AT3G22690  | Delayed green | Chloroplast          | Low PEP activity, rpoB editing defect                      | PPR domain                    | RNA editing            | Zhou et al., 2009                         |
| SEL1/PDM1 <sup>c</sup>                       | AT4G18520  | Albino        | Chloroplast          | Low PEP activity, rpoA processing defect                   | PPR domain                    | RNA processing         | Pyo et al., 2013                          |
| CLB19 <sup>c</sup>                           | AT1G05750  | Albino        | Chloroplast          | Low PEP activity, RNA editing defect                       | PPR domain                    | RNA editing            | Chateigner-Boutin et al., 2008            |

\* Observed phenotype and Molecular phenotype of inactivation mutants is given.

\*\* Localization information is from GFP-fusion data.

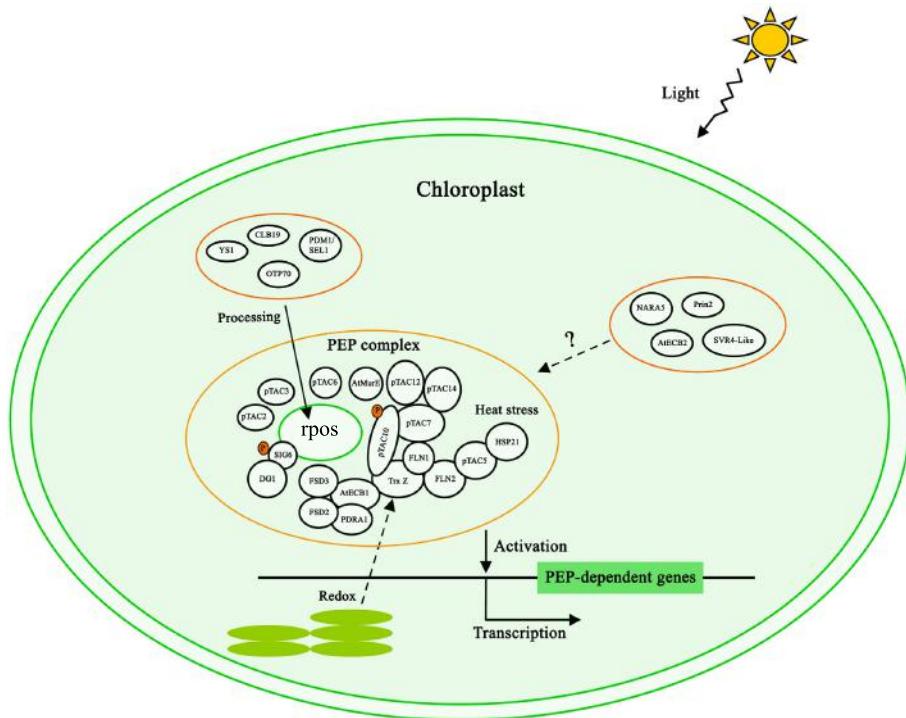
\*\*\* Protein domain information is from PfDB database.

\*\*\*\* Molecular function data is given based on the reference. U.K. means that its detailed molecular function remains unclear.

<sup>a</sup>These factors are associated with the PAPs or the core subunits.

<sup>b</sup>These factors regulate plastid transcription with unknown mechanism.

<sup>c</sup>These factors indirectly affect PEP activity through regulating the processing of chloroplast transcripts encoding the core subunits.



**FIGURE 1 | Model of the PEP complex regulating chloroplast transcription.** The reported interactions of PEP components are included. Light, redox status, phosphorylation and heat stress are involved in chloroplast transcription. Several PPR proteins including CLB19,

PDM1/SEL1, OTP70, and YS1 are involved in the processing of PEP core subunit transcripts to regulate PEP activity. AtECB2, Prin2, SVR4-Like and NARA5 are also important for plastid gene expression, but their mechanisms are unclear.

## CONNECTIONS OF REGULATORY MODULES WITH THE RNA POLYMERASE

Light plays highly important roles in the regulation of plastid gene transcription. The majority of PAPs (Pfalz and Pfannschmidt, 2013) and most sigma factor genes of higher plants are light-induced (Lerbs-Mache, 2011). Plastome-wide PEP-DNA association is also a light-dependent process (Finster et al., 2013). In plants, light plays an important role in almost every facet of plant growth and development through the action of photoreceptors. Interestingly, pTAC12 is an intrinsic subunit of the PEP complex (Pfalz et al., 2006; Steiner et al., 2011), but it was also identified as HEMERA and localized in both the nucleus and the chloroplast (Chen et al., 2010). pTAC12/HEMERA was considered as a proteolysis-related protein involved in phytochrome signaling in the nucleus (Chen et al., 2010). Its function in the PEP complex is unknown so far, but it was uncovered that pTAC12 interacts with pTAC14 in the yeast-two-hybrid system (Gao et al., 2011) suggesting that these two proteins might be also interaction partners in the native complex.

Chloroplasts are the site of photosynthesis that also produces reactive oxygen species (ROS). During photosynthesis, unbalanced excitation of the two photosystems affects the redox state of the electron transport chain which in turn serve as signals for plant acclimation responses. The PEP complex is a major target of such photosynthetic redox signals (Dietz and Pfannschmidt, 2011). Thioredoxin z (Trx Z) is a novel thioredoxin protein

with disulfide reductase activity *in vitro*. It interacts with two fructokinase-like proteins FLN1 and FLN2 in the yeast two hybrid system and is also a component of the PEP complex (Pfalz et al., 2006; Steiner et al., 2011) (Figure 1). Trx-Z mediated redox change of FLN2 during light-dark transitions (Arsova et al., 2010). Recent studies identified AtECB1/MRL7 as a thioredoxin-fold like protein with thioredoxin activity (Yu et al., 2014) that interacts with Trx Z in the PEP complex (Powikrowska et al., 2014; Yu et al., 2014). These two proteins thus may form a functional module to mediate redox signaling from thylakoids toward the RNA polymerase but the functional details of these interactions are completely unknown. Further redox mediators might be Fe Superoxide Dismutase 2 (FSD2) and FSD3, two iron superoxide dismutases, and PRDA1 is a chloroplast protein without any known domain. *prda1* and *fsd2 fsd3* knock out mutants are highly sensitive to oxidative stress (Myouga et al., 2008; Qiao et al., 2013). These proteins, therefore, may act as ROS scavengers in order to protect the PEP complex. The interactions between AtECB1 and PRDA1, FSD2, FSD3 suggest that the redox signaling pathway and ROS scavengers are eventually associated.

Protein phosphorylation is a very important post-translational modification in eukaryotic cells that regulates many cellular processes. In chloroplast, the phosphorylation of chloroplast proteins affects photosynthesis, metabolic functions and chloroplast transcription (Baginsky and Gruissem, 2009). The PEP complex appears to interact with a so-called plastid transcription kinase

(PTK), named cpCK2 (Ogrzewalla et al., 2002). The *Arabidopsis* sigma factor 6 was reported to be phosphorylated by cpCK2 (Schweer et al., 2010a). Furthermore, pTAC5, pTAC10, and pTAC16, were also predicated to be phosphorylated by cpCK2 (Reiland et al., 2009). The enzyme activity of cpCK2 was inhibited by GSH, which suggests that cpCK2 is generally under SH-group redox regulation (Baginsky et al., 1999; Turkeri et al., 2012). Biochemical analyses of mustard seedlings during photosynthetic acclimation suggested that redox signals in chloroplasts are linked to chloroplast transcription via the combined action of phosphorylation and thiol-mediated regulation events (Steiner et al., 2009). Proteins related with phosphorylation and redox signaling are closely located in the PEP complex which is in agreement with the results of the physiological studies for plastid gene expression.

Heat stress is a major abiotic factor for plants, that leads to severe retardation in plant growth and development. To maintain the process of chloroplast transcription under heat stress and to support the survival of the plant, the chloroplast transcriptional machinery needs to deal with heat stress to a certain extent. The protein pTAC5 is a C4-type zinc finger DnaJ protein with disulfide isomerase activity. Its expression is induced by heat stress (Zhong et al., 2013) and, subsequently, pTAC5 and Heat Shock Protein 21 (HSP21) form a heterocomplex, although they are not PAP members of the PEP complex (Zhong et al., 2013). pTAC5 as well as HSP21 may protect chloroplast transcription under heat stress.

#### OTHER NUCLEAR ENCODED FACTORS THAT REGULATE PEP ACTIVITY

In addition to the intrinsic components of PEP complex, multiple additional factors were identified to regulate the processing of PEP core subunit transcripts and PEP activity by individual mutant analysis. Both *Chloroplast Biogenesis19* (CLB19) (Chateigner-Boutin et al., 2008) and *Pigment-Deficient Mutant 1* (PDM1) (Wu and Zhang, 2010; Yin et al., 2012) genes encode pentatricopeptide repeat proteins. CLB19 is involved in the editing of the *rpoA* transcript (Chateigner-Boutin et al., 2008), while PDM1 is associated with *rpoA* polycistronic for *rpoA* cleavage (Wu and Zhang, 2010; Yin et al., 2012). Recent investigations demonstrated that PDM1/Seedling Lethal1 (SEL1) was also involved in *accD* RNA editing (Pyo et al., 2013). The PPR protein OTP70 was reported to affect the splicing of the *rpoC1* transcript (Chateigner-Boutin et al., 2011). The gene *Yellow Seedling 1* (YS1) encoding a PPR-DYW protein is required for editing of *rpoB* transcripts (Zhou et al., 2009). The common feature of the *Arabidopsis* knockout lines for all these proteins is that the plastid expression pattern in these mutants is similar to that of *rpo*-gene knock-out mutants in tobacco (Chateigner-Boutin et al., 2008, 2011; Zhou et al., 2009; Wu and Zhang, 2010; Pyo et al., 2013).

Functional analyses revealed that several proteins including *Arabidopsis* Early Chloroplast Biogenesis 2 (AtECB2) (Yu et al., 2009), Plastid redox insensitive 2 (Prin2) (Kindgren et al., 2012), SVR4-Like (Powikrowska et al., 2014), and NARA5 (Ogawa et al., 2009), are also essential for PEP-dependent chloroplast transcription. However, it is unclear if they are directly associated with the PEP complex. AtECB2 encodes a pentatricopeptide repeat protein, and is involved in editing of *accD* and *ndhF* chloroplast transcripts (Yu et al., 2009; Tseng et al., 2010). The defective editing in *ecb2* is unlikely to affect PEP-dependent plastid gene expression. How AtECB2 affects plastid gene expression is still unclear.

NARA5 encodes a chloroplast-localized phosphofructokinase B-type carbohydrate kinase family protein, which might be involved in massive expressions of plastid-encoded photosynthetic genes in *Arabidopsis* (Ogawa et al., 2009). The Prin2 is a small protein possibly involved in redox-mediated retrograde signaling in chloroplast (Kindgren et al., 2012) and the SVR4-like is a homolog of AtECB1/SVR4/MRL7, encoding a chloroplast protein essential for proper function of the chloroplast in *Arabidopsis* (Powikrowska et al., 2014). All these proteins may reversibly associate with the PEP complex but detailed studies are necessary to understand their functional roles and connections with the RNA polymerase. Alternatively, these proteins may act as signaling factors in order to mediate environmental stimuli and plastid gene expression.

#### CONCLUDING REMARKS

Plants grow under very different environment conditions and photosynthesis is the major function of chloroplast which is important for plant growth and development. Plastid gene expression is essential for chloroplast development and normal functions including photosynthesis. The PEP complex is the major RNA polymerase activity in mature chloroplasts. Proteomic and genetic analyses identified that at least 50 nuclear-encoded proteins in higher plant are important for PEP dependent plastid gene expression. These proteins may form several functional modules within the nucleoid or TAC in order to mediate plastid gene expression in response to light, redox changes, phosphorylation and heat stress or to protect the PEP complex from ROS damage. The large number of nuclear-encoded proteins reveals the complexity of plastid gene expression and regulation that is greatly different from the gene expression in the nucleus or in prokaryotes. However, the current knowledge about plastid transcription is quite limited and the investigation of the relationship between transcription, post-transcriptional processing as well as translation in the nucleoid could provide novel insights into chloroplast gene expression.

#### ACKNOWLEDGMENTS

We are grateful to Prof. Thomas Pfannschmidt from Grenoble University for his help in editing the manuscript. This work was supported by grants from the National Science Foundation of China (Grant no. 31100965 and Grant no. 31370271).

#### REFERENCES

- Allison, L. A., Simon, L. D., and Maliga, P. (1996). Deletion of *rpoB* reveals a second distinct transcription system in plastids of higher plants. *EMBO J.* 15, 2802–2809.
- Armbruster, U., Labs, M., Pribil, M., Viola, S., Xu, W., Scharfenberg, M., et al. (2013). *Arabidopsis* CURVATURE THYLAKOID1 proteins modify thylakoid architecture by inducing membrane curvature. *Plant Cell* 25, 2661–2678. doi: 10.1105/tpc.113.113118
- Arsova, B., Hoja, U., Wimmelbacher, M., Greiner, E., Ustün, S., Melzer, M., et al. (2010). Plastidial thioredoxin z interacts with two fructokinase-like proteins in a thiol-dependent manner: evidence for an essential role in chloroplast development in *arabidopsis* and *nicotiana benthamiana*. *Plant Cell* 22, 1498–1515. doi: 10.1105/tpc.109.071001
- Azevedo, J., Courtois, F., Hakimi, M. A., Demarsy, E., Lagrange, T., Alcaraz, J. P., et al. (2008). Intraplastidial trafficking of a phage-type RNA polymerase is mediated by a thylakoid RING-H2 protein. *Proc. Natl. Acad. Sci. U.S.A.* 105, 9123–9128. doi: 10.1073/pnas.0800909105

- Baginsky, S., and Grussem, W. (2009). The chloroplast kinase network: new insights from large-scale phosphoproteome profiling. *Mol. Plant* 2, 1141–1153. doi: 10.1093/mp/ssp058
- Baginsky, S., Tiller, K., Pfannschmidt, T., and Link, G. (1999). PTK, the chloroplast RNA polymerase-associated protein kinase from mustard (*Sinapis alba*), mediates redox control of plastid *in vitro* transcription. *Plant Mol. Biol.* 39, 1013–1023.
- Chang, C. C., Sheen, J., Bligny, M., Niwa, Y., Lerbs-Mache, S., and Stern, D. B. (1999). Functional analysis of two *maize* cDNAs encoding T7-like RNA polymerases. *Plant Cell* 11, 911–926. doi: 10.1105/tpc.11.5.911
- Chateigner-Boutin, A. L., des Francs-Small, C. C., Delannoy, E., Kahlau, S., Tanz, S. K., de Longevialle, A. F., et al. (2011). OTP70 is a pentatricopeptide repeat protein of the E subgroup involved in splicing of the plastid transcript rpoC1. *Plant J.* 65, 532–542. doi: 10.1111/j.1365-313X.2010.04441.x
- Chateigner-Boutin, A. L., Ramos-Vega, M., Guevara-García, A., Andrés, C., de la Luz Gutiérrez-Navala, Luz Gutiérrez-Navala, M., Cantero, A., et al. (2008). CLB19, a pentatricopeptide repeat protein required for editing of rpoA and clpP chloroplast transcripts. *Plant J.* 56, 590–602. doi: 10.1111/j.1365-313X.2008.03634.x
- Chen, M., Galvão, R. M., Li, M., Burger, B., Bugea, J., Bolado, J., et al. (2010). *Arabidopsis* HEMERA/pTAC12 initiates photomorphogenesis by phytochromes. *Cell* 141, 1230–1240. doi: 10.1016/j.cell.2010.05.007
- Chi, W., Ma, J., Zhang, D., Guo, J., Chen, F., Lu, C., et al. (2008). The pentatricopeptide repeat protein DELAYED GREENING1 is involved in the regulation of early chloroplast development and chloroplast gene expression in *Arabidopsis*. *Plant Physiol.* 147, 573–584. doi: 10.1104/pp.108.116194
- Chi, W., Mao, J., Li, Q., Ji, D., Zou, M., Lu, C., et al. (2010). Interaction of the pentatricopeptide-repeat protein DELAYED GREENING 1 with sigma factor SIG6 in the regulation of chloroplast gene expression in *Arabidopsis* cotyledons. *Plant J.* 64, 14–25. doi: 10.1111/j.1365-313X.2010.04304.x
- Courtois, F., Merendino, L., Demarsy, E., Mache, R., and Lerbs-Mache, S. (2007). Phage-type RNA polymerase RPOTmp transcribes the rrn operon from the PC promoter at early developmental stages in *Arabidopsis*. *Plant Physiol.* 145, 712–721. doi: 10.1104/pp.107.103846
- da Costa e Silva, O., Lorbicke, R., Garg, P., Müller, L., Wassmann, M., Lautert, P., et al. (2004). The Etched1 gene of *Zea mays* (L.) encodes a zinc ribbon protein that belongs to the transcriptionally active chromosome (TAC) of plastids and is similar to the transcription factor TFIIS. *Plant J.* 38, 923–939. doi: 10.1111/j.1365-313X.2004.02094.x
- Demarsy, E., Courtois, F., Azevedo, J., Buhot, L., and Lerbs-Mache, S. (2006). Building up of the plastid transcriptional machinery during germination and early plant development. *Plant Physiol.* 142, 993–1003. doi: 10.1104/pp.106.085043
- De Santis-Maclossek, G., Kofer, W., Bock, A., Schoch, S., Maier, R. M., Wanner, G., et al. (1999). Targeted disruption of the plastid RNA polymerase genes rpoA, B and C1: molecular biology, biochemistry and ultrastructure. *Plant J.* 18, 477–489. doi: 10.1046/j.1365-313X.1999.00473.x
- Dietz, K., and Pfannschmidt, T. (2011). Novel regulators in photosynthetic redox control of plant metabolism and gene expression. *Plant Physiol.* 155, 1477–1585. doi: 10.1104/pp.110.170043
- Finster, S., Eggert, E., Zoschke, R., Weihe, A., and Schmitz-Linneweber, C. (2013). Light-dependent, plastome-wide association of the plastid-encoded RNA polymerase with chloroplast DNA. *Plant J.* 76, 849–860. doi: 10.1111/tpj.12339
- Gao, Z. P., Yu, Q. B., Zhao, T. T., Ma, Q., Chen, G. X., and Yang, Z. N. (2011). A functional component of the transcriptionally active chromosome complex, *Arabidopsis* pTAC14, interacts with pTAC12/HEMERA and regulates plastid gene expression. *Plant Physiol.* 157, 1733–1745. doi: 10.1104/pp.111.184762
- Garcia, M., Myouga, F., Takechi, K., Sato, H., Nabeshima, K., Nagata, N., et al. (2008). An *Arabidopsis* homolog of the bacterial peptidoglycan synthesis enzyme MurE has an essential role in chloroplast development. *Plant J.* 53, 924–934. doi: 10.1111/j.1365-313X.2007.03379.x
- Gilkerson, J., Perez-Ruiz, J. M., Chory, J., and Callis, J. (2012). The plastid-localized pfkB-type carbohydrate kinases FRUCTOKINASE-LIKE 1 and 2 are essential for growth and development of *Arabidopsis* thaliana. *BMC Plant Biol.* 12:102. doi: 10.1186/1471-2229-12-102
- Hajdukiewicz, P. T., Allison, L. A., and Maliga, P. (1997). The two RNA polymerases encoded by the nuclear and the plastid compartments transcribe distinct groups of genes in tobacco plastids. *EMBO J.* 16, 4041–4048. doi: 10.1093/emboj/16.13.4041
- Hess, W. R., and Borner, T. (1999). Organellar RNA polymerases of higher plants. *Int. Rev. Cytol.* 190, 1–59. doi: 10.1016/S0074-7696(08)62145-2
- Hricová, A., Quesada, V., and Micó, J. L. (2006). The SCABRA3 nuclear gene encodes the plastid RpoTp RNA polymerase, which is required for chloroplast biogenesis and mesophyll cell proliferation in *Arabidopsis*. *Plant Physiol.* 141, 942–956. doi: 10.1104/pp.106.080069
- Huang, C., Yu, Q. B., Lv, R. H., Yin, Q. Q., Chen, G. Y., Xu, L., et al. (2013). The reduced plastid-encoded polymerase-dependent plastid gene expression leads to the delayed greening of the *Arabidopsis fln2* mutant. *PLoS ONE* 8:e73092. doi: 10.1371/journal.pone.0073092
- Ishizaki, Y., Tsunoyama, Y., Hatano, K., Ando, K., Kato, K., Shinmyo, A., et al. (2005). A nuclear-encoded sigma factor, *Arabidopsis* SIG6, recognizes sigma-70 type chloroplast promoters and regulates early chloroplast development in cotyledons. *Plant J.* 42, 133–144. doi: 10.1111/j.1365-313X.2005.02362.x
- Jeon, Y., Jung, H. J., Kang, H., Park, Y. I., Lee, S. H., and Pai, H. S. (2012). S1 domain-containing STF modulates plastid transcription and chloroplast biogenesis in *Nicotiana benthamiana*. *New Phytol.* 193, 349–363. doi: 10.1111/j.1469-8137.2011.03941.x
- Kindgren, P., Kremnev, D., Blanco, N. E., de Dios Barajas López, J., Fernández, A. P., Tellgren-Roth, C., et al. (2012). The plastid redox insensitive 2 mutant of *Arabidopsis* is impaired in PEP activity and high light-dependent plastid redox signalling to the nucleus. *Plant J.* 70, 279–291. doi: 10.1111/j.1365-313X.2011.04865.x
- Kleine, T. (2012). *Arabidopsis* thaliana mTERF proteins: evolution and functional classification. *Front Plant Sci.* 3:233. doi: 10.3389/fpls.2012.00233
- Krause, K., and Krupinska, K. (2000). Molecular and functional properties of highly purified transcriptionally active chromosomes from *spinach* chloroplasts. *Physiol. Plant* 109, 188–195. doi: 10.1034/j.1399-3054.2000.100211.x
- Leister, D. (2003). Chloroplast research in the genomic age. *Trends Genet.* 19, 47–57. doi: 10.1016/S0168-9525(02)00003-3
- Lerbs-Mache, S. (2011). Function of plastid sigma factors in higher plants: regulation of gene expression or just preservation of constitutive transcription? *Plant Mol. Biol.* 76, 235–249. doi: 10.1007/s11103-010-9714-4
- Li, H. M., and Chiu, C. C. (2010). Protein transport into chloroplasts. *Annu. Rev. Plant Biol.* 61, 157–180. doi: 10.1146/annurev-arplant-042809-112222
- Liere, K., Weihe, A., and Börner, T. (2011). The transcription machineries of plant mitochondria and chloroplasts: composition, function, and regulation. *J. Plant Physiol.* 168, 1345–1360. doi: 10.1016/j.jplph.2011.01.005
- Lopez-Juez, E., and Pyke, K. A. (2005). Plastids unleashed: their development and their integration in plant development. *Int. J. Dev. Biol.* 49, 557–577. doi: 10.1387/ijdb.051997el
- Loschereder, H., Homann, A., Ogrzewalla, K., and Link, G. (2004). Proteomics-based sequence analysis of plant gene expression—the chloroplast transcription apparatus. *Phytochemistry* 65, 1785–1793. doi: 10.1016/j.phytochem.2004.04.034
- Maréchal, A., Parent, J. S., Véronneau-Lafortune, F., Joyeux, A., Lang, B. F., and Brisson, N. (2009). Whirly proteins maintain plastid genome stability in *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A.* 106, 14693–14698. doi: 10.1073/pnas.0901710106
- McFadden, G. I., and van Dooren, G. G. (2004). Evolution: red algal genome affirms a common origin of all plastids. *Curr. Biol.* 14, R514–R516. doi: 10.1016/j.cub.2004.06.041
- Melonk, J., Matros, A., Trosch, M., Mock, H. P., and Krupinska, K. (2012). The core of chloroplast nucleoids contains architectural SWIB domain proteins. *Plant Cell* 24, 3060–3073. doi: 10.1105/tpc.112.099721
- Myouga, F., Hosoda, C., Umezawa, T., Iizumi, H., Kuromori, T., Motohashi, R., et al. (2008). A heterocomplex of iron superoxide dismutases defends chloroplast nucleoids against oxidative stress and is essential for chloroplast development in *Arabidopsis*. *Plant Cell* 20, 3148–3162. doi: 10.1105/tpc.108.061341
- Ogawa, T., Nishimura, K., Aoki, T., Takase, H., Tomizawa, K., Ashida, H., et al. (2009). A phosphofructokinase B-type carbohydrate kinase family protein, NARA5, for massive expressions of plastid-encoded photosynthetic genes in *Arabidopsis*. *Plant Physiol.* 151, 114–128. doi: 10.1104/pp.109.139683
- Ogrzewalla, K., Piotrowski, M., Reinbothe, S., and Link, G. (2002). The plastid transcription kinase from mustard (*Sinapis alba* L.). A nuclear-encoded CK2-type chloroplast enzyme with redox-sensitive function. *Eur. J. Biochem.* 269, 3329–3337. doi: 10.1046/j.1432-1033.2002.03017\_269\_13.x
- 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. doi: 10.1105/tpc.105.036392

- Pfalz, J., and Pfannschmidt, T. (2013). Essential nucleoid proteins in early chloroplast development. *Trends Plant Sci.* 18, 186–194. doi: 10.1016/j.tplants.2012.11.003
- Pfannschmidt, T., and Link, G. (1994). Separation of two classes of plastid DNA-dependent RNA polymerases that are differentially expressed in mustard (*Sinapis alba* L.) seedlings. *Plant Mol. Biol.* 25, 69–81.
- Pfannschmidt, T., and Link, G. (1997). The A and B forms of plastid DNA-dependent RNA polymerase from mustard (*Sinapis alba* L.) transcribe the same genes in a different developmental context. *Mol. Gen. Genet.* 257, 35–44. doi: 10.1007/s004380050621
- Pfannschmidt, T., Ogrzewalla, K., Baginsky, S., Sickmann, A., Meyer, H. E., and Link, G. (2000). The multisubunit chloroplast RNA polymerase A from mustard (*Sinapis alba* L.). Integration of a prokaryotic core into a larger complex with organelle-specific functions. *Eur. J. Biochem.* 267, 253–261. doi: 10.1046/j.1432-1327.2000.00991.x
- Powikowska, M., Khrouchchova, A., Martens, H. J., Zygadlo-Nielsen, A., Melonek, J., Schulz, A., et al. (2014). SVR4 (suppressor of variegation 4) and SVR4-like: two proteins with a role in proper organization of the chloroplast genetic machinery. *Physiol. Plant* 150, 477–492. doi: 10.1111/pl.12108
- Prikryl, J., Watkins, K. P., Friso, G., van Wijk, K. J., and Barkan, A. (2008). A member of the Whirly family is a multifunctional RNA- and DNA-binding protein that is essential for chloroplast biogenesis. *Nucleic Acids. Res.* 36, 5152–5165. doi: 10.1093/nar/gkn492
- Pyo, Y. J., Kwon, K. C., Kim, A., and Cho, M. H. (2013). Seedling Lethal1, a pentatricopeptide repeat protein lacking an E/E+ or DYW domain in *Arabidopsis*, is involved in plastid gene expression and early chloroplast development. *Plant Physiol.* 163, 1844–1858. doi: 10.1104/pp.113.227199
- Qiao, J., Li, J., Chu, W., and Luo, M. (2013). PRDA1, a novel chloroplast nucleoid protein, is required for early chloroplast development and is involved in the regulation of plastid gene expression in *Arabidopsis*. *Plant Cell Physiol.* 54, 2071–2084. doi: 10.1093/pcp/pct148
- Qiao, J. W., Ma, C. L., Wimmelbacher, M., Bornke, F. M., and Luo, M. Z. (2011). Two novel proteins, MRL7 and its paralog MRL-7, have essential but functionally distinct roles in chloroplast development and are involved in plastid gene expression regulation in *Arabidopsis*. *Plant Cell Physiol.* 52, 1017–1030. doi: 10.1093/pcp/pqr054
- Reiland, S., Messerli, G., Baerenfaller, K., Gerrits, B., Endler, A., Grossmann, J., et al. (2009). Large-scale *Arabidopsis* phosphoproteome profiling reveals novel chloroplast kinase substrates and phosphorylation networks. *Plant Physiol.* 150, 889–903. doi: 10.1104/pp.109.138677
- Schröter, Y., Steiner, S., Matthäi, K., and Pfannschmidt, T. (2010). Analysis of oligomeric protein complexes in the chloroplast sub-proteome of nucleic acid-binding proteins from mustard reveals potential redox regulators of plastid gene expression. *Proteomics* 10, 2191–2204. doi: 10.1002/pmic.200900678
- Schweer, J. (2010). Plant sigma factors come of age: flexible transcription factor network for regulated plastid gene expression. *J. Endocyt. Cell Res.* 20, 1–12.
- Schweer, J., Türkeri, H., Kolpack, A., and Link, G. (2010b). Role and regulation of plastid sigma factors and their functional interactors during chloroplast transcription—recent lessons from *Arabidopsis thaliana*. *Eur. J. Cell Biol.* 89, 940–946. doi: 10.1016/j.ejcb.2010.06.016
- Schweer, J., Türkeri, H., Link, B., and Link, G. (2010a). AtSIG6, a plastid sigma factor from *Arabidopsis*, reveals functional impact of cpCK2 phosphorylation. *Plant J.* 62, 192–202. doi: 10.1111/j.1365-313X.2010.04138.x
- Steiner, S., Dietzel, L., Schröter, Y., Fey, V., Wagner, R., and Pfannschmidt, T. (2009). The role of phosphorylation in redox regulation of photosynthesis genes *psaA* and *psbA* during photosynthetic acclimation of mustard. *Mol. Plant* 2, 416–429. doi: 10.1093/mp/ssp007
- Steiner, S., Schröter, Y., Pfalz, J., and Pfannschmidt, T. (2011). Identification of essential subunits in the plastid-encoded RNA polymerase complex reveals building blocks for proper plastid development. *Plant Physiol.* 157, 1043–1055. doi: 10.1104/pp.111.184515
- Sugiura, M. (1992). The chloroplast genome. *Plant Mol. Biol.* 19, 149–168. doi: 10.1007/BF00015612
- Suzuki, J. Y., Ytterberg, A. J., Beardslee, T. A., Allison, L. A., Wijk, K. J., and Maliga, P. (2004). Affinity purification of the tobacco plastid RNA polymerase and *in vitro* reconstitution of the holoenzyme. *Plant J.* 40, 164–172. doi: 10.1111/j.1365-313X.2004.02195.x
- Swiatecka-Hagenbruch, M., Emanuel, C., Hedtke, B., Liere, K., and Börner, T. (2008). Impaired function of the phage-type RNA polymerase RpoTp in transcription of chloroplast genes is compensated by a second phage-type RNA polymerase. *Nucleic Acids. Res.* 36, 785–792. doi: 10.1093/nar/gkm111
- Tseng, C. C., Sung, T. Y., Li, Y. C., Hsu, S. J., Lin, C. L., and Hsieh, M. H. (2010). Editing of *accD* and *ndhF* chloroplast transcripts is partially affected in the *Arabidopsis* vanilla cream1 mutant. *Plant Mol. Biol.* 73, 309–323. doi: 10.1007/s11103-010-9616-5
- Türkeri, H., Schweer, J., and Link, G. (2012). Phylogenetic and functional features of the plastid transcription kinase cpCK2 from *Arabidopsis* signify a role of cysteinyl SH-groups in regulatory phosphorylation of plastid sigma factors. *FEBS J.* 279, 395–409. doi: 10.1111/j.1742-4658.2011.08433.x
- Williams-Carrier, R., Zoschke, R., Belcher, S., Pfalz, J., and Barkan, A. (2014). A major role for the plastid-encoded RNA polymerase complex in the expression of plastid transfer RNAs. *Plant Physiol.* 164, 239–248. doi: 10.1104/pp.113.228726
- Wu, H., and Zhang, L. (2010). The PPR protein PDM1 is involved in the processing of rpoA pre-mRNA in *Arabidopsis thaliana*. *Chin. Sci. Bull.* 55, 3485–3489. doi: 10.1007/s11434-010-4040-4
- Xiong, J. Y., Lai, C. X., Qu, Z., Yang, X. Y., Qin, X. H., and Liu, G. Q. (2009). Recruitment of AtWHY1 and AtWHY3 by a distal element upstream of the kinesin gene AtKP1 to mediate transcriptional repression. *Plant Mol. Biol.* 71, 437–449. doi: 10.1007/s11103-009-9533-7
- Yagi, Y., Ishizaki, Y., Nakahira, Y., Tozawa, Y., and Shiina, T. (2012). Eukaryotic-type plastid nucleoid protein pTAC3 is essential for transcription by the bacterial-type plastid RNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.* 109, 7541–7546. doi: 10.1073/pnas.1119403109
- Yin, Q. Q., Cui, Y. L., Zhang, G. R., Zhang, H. D., Wang, X. M., and Yang, Z. N. (2012). The *Arabidopsis* pentatricopeptide repeat protein PDM1 is associated with the intergenic sequence of S11-rpoA for rpoA monocistronic RNA cleavage. *Chinese Sci. Bull.* 57, 3452–3459. doi: 10.1007/s11434-012-5278-9
- Yu, Q. B., Jiang, Y., Chong, K., and Yang, Z. N. (2009). AtECB2, a pentatricopeptide repeat protein, is required for chloroplast transcript accD RNA editing and early chloroplast biogenesis in *Arabidopsis thaliana*. *Plant J.* 59, 1011–1023. doi: 10.1111/j.1365-313X.2009.03930.x
- Yu, Q. B., Lu, Y., Ma, Q., Zhao, T. T., Huang, C., Zhao, H. F., et al. (2013). TAC7, an essential component of the plastid transcriptionally active chromosome complex, interacts with FLN1, TAC10, TAC12 and TAC14 to regulate chloroplast gene expression in *Arabidopsis thaliana*. *Physiol. Plant.* 148, 408–421. doi: 10.1111/j.1399-3054.2012.01718.x
- Yu, Q. B., Ma, Q., Kong, M. M., Zhao, T. T., Zhang, X. L., Zhou, Q., et al. (2014). AtECB1/MRL7, a thioredoxin-like fold protein with disulfide reductase activity, regulates chloroplast gene expression and chloroplast biogenesis in *Arabidopsis thaliana*. *Mol. Plant* 7, 206–217. doi: 10.1093/mp/sst092
- Zhong, L., Zhou, W., Wang, H., Ding, S., Lu, Q., Wen, X., et al. (2013). Chloroplast small heat shock protein HSP21 interacts with plastid nucleoid protein pTAC5 and is essential for chloroplast development in *Arabidopsis* under heat stress. *Plant Cell* 25, 2925–2943. doi: 10.1105/tpc.113.111229
- Zhou, W., Cheng, Y., Yap, A., Chateigner-Boutin, A. L., Delannoy, E., Hammami, K., et al. (2009). The *Arabidopsis* gene YS1 encoding a DYW protein is required for editing of rpoB transcripts and the rapid development of chloroplasts during early growth. *Plant J.* 58, 82–96. doi: 10.1111/j.1365-313X.2008.03766.x
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Received:** 31 March 2014; **accepted:** 14 June 2014; **published online:** 03 July 2014.
- Citation:** Yu Q-B, Huang C and Yang Z-N (2014) Nuclear-encoded factors associated with the chloroplast transcription machinery of higher plants. *Front. Plant Sci.* 5:316. doi: 10.3389/fpls.2014.00316
- This article was submitted to Plant Physiology, a section of the journal *Frontiers in Plant Science*.
- Copyright © 2014 Yu, Huang and Yang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.**



# Recent advances in the study of chloroplast gene expression and its evolution

**Yusuke Yagi<sup>1</sup> and Takashi Shiina<sup>2\*</sup>**

<sup>1</sup> Faculty of Agriculture, Kyushu University, Fukuoka, Japan

<sup>2</sup> Graduate School of Life and Environmental Sciences, Kyoto Prefectural University, Kyoto, Japan

**Edited by:**

Thomas Pfannschmidt, Joseph Fourier University, France

**Reviewed by:**

Frederik Börnke, Leibniz-Institute for Vegetable and Ornamental Crops, Germany

Hannet Roschzttardtz, University of Wisconsin-Madison, USA

**\*Correspondence:**

Takashi Shiina, Graduate School of Life and Environmental Sciences, Kyoto Prefectural University, Kyoto 606-8522, Japan

e-mail: shiina@kpu.ac.jp

Chloroplasts are semiautonomous organelles which possess their own genome and gene expression system. However, extant chloroplasts contain only limited coding information, and are dependent on a large number of nucleus-encoded proteins. During plant evolution, chloroplasts have lost most of the prokaryotic DNA-binding proteins and transcription regulators that were present in the original endosymbiont. Thus, chloroplasts have a unique hybrid transcription system composed of the remaining prokaryotic components, such as a prokaryotic RNA polymerase as well as nucleus-encoded eukaryotic components. Recent proteomic and transcriptomic analyses have provided insights into chloroplast transcription systems and their evolution. Here, we review chloroplast-specific transcription systems, focusing on the multiple RNA polymerases, eukaryotic transcription regulators in chloroplasts, chloroplast promoters, and the dynamics of chloroplast nucleoids.

**Keywords:** chloroplast, transcription, PEP, NEP, pTAC, nucleoid

## INTRODUCTION

Chloroplasts are believed to have arisen from an endosymbiotic event between a photosynthetic cyanobacterium and the ancestral eukaryotic cell. Although chloroplasts of modern plants and algae have retained the genome of the symbiont, that genome has markedly shrunk over endosymbiotic evolution. Many chloroplast-encoded genes were lost or transferred to the nucleus soon after endosymbiosis. Thus, chloroplast genomes of extant land plants have only 50 protein-coding genes involved in photosynthesis, gene expression, lipid metabolism and other processes, 30 tRNA genes and full sets of rRNA genes. In spite of their small genomes (0.15 Mbp in land plant chloroplasts versus 3 Mbp in cyanobacteria), chloroplast gene expression is regulated by more complex systems compared to the simple prokaryotic regulatory system. Chloroplast gene expression is mediated by two distinct types of RNA polymerase (RNAP) and is highly dependent on post-transcriptional regulation, such as the processing of polycistronic transcripts, intron splicing and RNA editing. Moreover, recent RNA-seq analyses of chloroplast transcripts identified unexpected diversifications of RNA molecules, such as non-coding and antisense RNAs (Hotto et al., 2011; Zhelyazkova et al., 2012). However, the genes encoded in chloroplast genomes are insufficient to regulate their complicated gene expression, and so the chloroplast gene expression machinery includes various nucleus-encoded regulatory components.

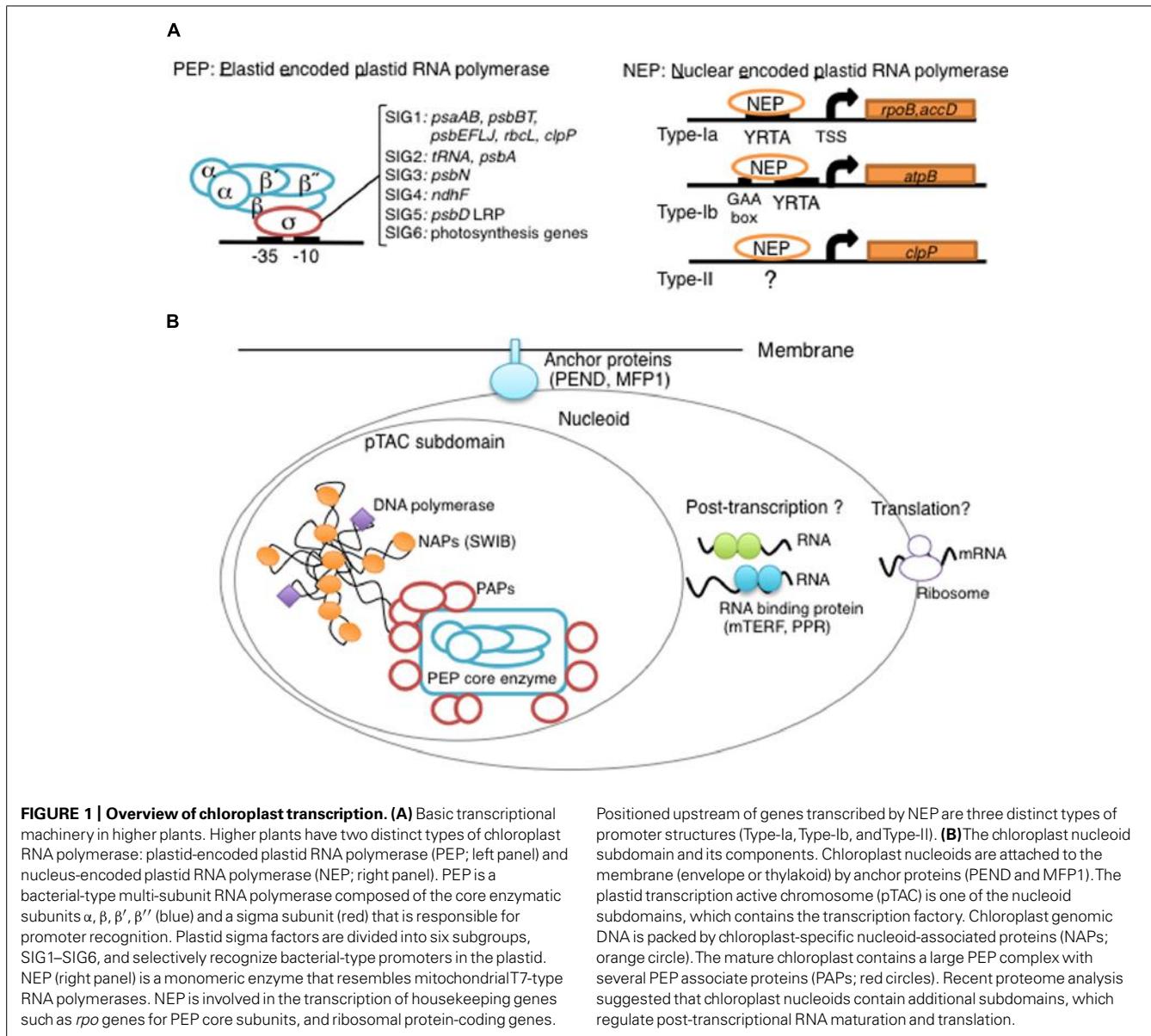
Although basic chloroplast gene expression is mediated by prokaryotic machineries derived from the ancestral cyanobacterium, chloroplasts lost their homologs of bacterial regulatory elements such as transcription factors (TFs) and nucleoid proteins at an early stage of their evolution. Genomics and proteomics analyses of chloroplast proteins in *Arabidopsis thaliana* have suggested that 60% of the chloroplast proteome may have

been newly acquired from the nuclear genome of host cells after the endosymbiotic event (Abdallah et al., 2000). Indeed, recent analyses of the chloroplast nucleoid proteins identified many non-bacterial components that play critical roles in chloroplast gene expression including transcription, post-transcriptional RNA processing, and translation. Here, we summarize the current knowledge regarding the chloroplast gene expression system.

## TWO BASIC CHLOROPLAST TRANSCRIPTION MACHINERIES WITH DIFFERENT EVOLUTIONARY ORIGIN

Chloroplast gene expression is largely dependent on prokaryotic machineries derived from the ancestral cyanobacterium. The bacterial multi-subunit RNAP is composed of a core Rpo complex, which has the catalytic enzyme activity, and a sigma factor, which recognizes promoter sequences (Ishihama, 2000). Chloroplasts contain the bacterial-type RNAP, called plastid-encoded plastid RNAP (PEP), which shares functional similarity with the bacterial RNAP (Igloi and Kossel, 1992; **Figure 1A**) However, all genes for chloroplast sigma factors have been transferred to the nuclear genome, whereas genes for core subunits are typically retained in the chloroplast genome as *rpoA*, *rpoB*, *rpoC1*, and *rpoC2*.

Early work demonstrated that almost all photosynthesis-related transcripts are significantly reduced in PEP-deficient plants, such as ribosome-deficient mutants of barley (*Hordeum vulgare*), *iojap* mutants of maize and tobacco mutants with disrupted *rpo* genes generated by gene targeting using chloroplast transformation (Han et al., 1992; Hess et al., 1993, 1994; Allison et al., 1996; De Santis-Maclossek et al., 1999), whereas a set of housekeeping genes are still active in these mutants. The inhibitor sensitivity of this transcription activity is similar to that of phage T7 RNAP, but not to that of bacterial RNAP (Kapoor et al., 1997; Sakai et al., 1998). In *Arabidopsis*, three phage-type RNAP genes



were identified and their subcellular localization was determined (*RpoTp*: chloroplasts, *RpoTm*: mitochondria, *RpoTmp*: chloroplast and mitochondria; Hettke et al., 1997, 2000; Weihe and Borner, 1999; Figure 1A). *RpoTmp* and *RpoTp* likely represent nuclear encoded RNAP (NEP) enzyme in chloroplasts [reviewed in (Liere et al., 2011)], while *RpoTmp* has been identified in dicotyledonous plants such as *Arabidopsis* and tobacco but not in monocotyledonous plant genomes (Chang et al., 1999; Ikeda and Gray, 1999; Emanuel et al., 2004).

Only one *RpoT* gene has been identified in green algae, such as *Chlamydomonas reinhardtii*, *Ostreococcus tauri*, and *Thalassiosira pseudonana*, which likely encodes mitochondrial RNAP (Maier et al., 2008). Similarly, the genome of the lycophyte *Selaginella moellendorffii* contains only one *RpoT* gene, the product of which has been shown to target mitochondria (Yin et al., 2009). On the other hand, the moss *Physcomitrella patens* has three *RpoT*

genes. However, all GFP-fused moss *RpoT*s were detected exclusively in mitochondria, suggesting that the moss *RpoT* genes also encode mitochondrial RNAP (Kabeya et al., 2002; Richter et al., 2002, 2013). Moreover, phylogenetic analysis of plant *RpoT* genes suggests that NEP appeared through the gene duplication of mitochondrial RNAP after the separation of angiosperms from gymnosperms (Yin et al., 2010).

## SELECTIVE CHLOROPLAST TRANSCRIPTION BY PEP AND NEP

Chloroplast genes can be categorized into three subgroups, classes I–III: class I photosynthesis-related genes are mainly transcribed by PEP; Class II includes many housekeeping genes (*clpP* and the *rrn* operon) that are transcribed by both PEP and NEP; class III genes (*accD* and the *rpoB* operon) are exclusively transcribed by NEP (Allison et al., 1996; Hajdukiewicz et al., 1997).

PEP recognizes standard chloroplast promoters resembling the bacterial  $\sigma^{70}$  type promoters with  $-10$  and  $-35$  consensus elements (Gatenby et al., 1981; Gruissem and Zurawski, 1985; Strittmatter et al., 1985; Shiina et al., 2005; **Figure 1A**). A genome-wide mapping of transcription start sites (TSSs) by RNA sequencing in barley green chloroplasts demonstrated that 89% of the mapped TSSs have a conserved  $-10$  element (TATAAT) at three to nine nucleotides upstream, while the  $-35$  element was mapped upstream of the  $-10$  element in only 70% of the TSSs (Zhelyazkova et al., 2012). These results suggest that most genes are transcribed from  $\sigma^{70}$ -type promoters by PEP in green leaves.

Higher plants have multiple sigma factors that are expected to confer promoter specificity upon the PEP core complex (Shiina et al., 2005; Lerbs-Mache, 2011). Molecular genetic analyses revealed that SIG2 is responsible for the transcription of a group of tRNA genes, but not photosynthesis genes (Kanamaru et al., 2001), while SIG6 is essential for the transcription of a wide range of photosynthesis-related genes at an early stage of chloroplast development (Ishizaki et al., 2005). It seems likely that SIG2 and SIG6 work in cooperation during light-dependent chloroplast development (Hanaoka et al., 2003; Ishizaki et al., 2005). In addition, SIG3 and SIG4 have been shown to specifically target *psbN* and *ndhF* genes in *Arabidopsis* (Favery et al., 2005; Zghidi et al., 2007). Recently, ChIP analysis of SIG1 revealed the target genes (*psaAB*, *psbBT*, *psbEFLJ*, *rbcL*, and *clpP*; Hanaoka et al., 2012). SIG5 is a unique sigma factor whose expression is rapidly induced by various environmental stresses such as a high osmolarity, or salinity, a low temperature as well as high-light stress (Tsunoyama et al., 2002; Nagashima et al., 2004). SIG5 likely recognizes specific promoters, including the *psbD* light-responsive promoter (LRP), and mediates stress-induced transcription in chloroplasts (Nagashima et al., 2004; Tsunoyama et al., 2004). Taken together, it is likely that each chloroplasts sigma factor is responsible for the transcription of a distinct set of genes, and plays specific roles in transcriptional regulation in response to developmental and/or environmental cues.

Phylogenetic analysis revealed that chloroplast sigma factors are related to essential group 1 and non-essential group 2 sigma factors in bacteria. *Chlamydomonas*, a single-celled green alga, possesses a single sigma factor that is related to SIG2 in land plants, suggesting the absence of multiple sigma factor-mediated transcriptional regulation in chloroplasts. Endosymbiosis of ancestral cyanobacteria in plant cells may have reduced the need for transcriptional regulation in chloroplasts and caused the reduction of the number of sigma factors in green algae. On the other hand, in liverwort (*M. polymorpha* L.) and moss (*P. patens*), three sigma factors related to SIG1, SIG2, and SIG5 are encoded in the nucleus (Shiina et al., 2009; Ueda et al., 2013). The multiple sigma factors in bryophytes may show a promoter preference and play roles in tissue-specific and stress-responsive transcriptional regulation in chloroplasts (Hara et al., 2001; Ichikawa et al., 2004; Kanazawa et al., 2013; Ueda et al., 2013).

Most NEP promoters (*rpoB*, *rpoA*, and *accD*) share a core sequence, the YRTA motif (type-Ia; Liere and Maliga, 1999; Weihe and Borner, 1999; Hirata et al., 2004; **Figure 1A**). The YRTA motif is similar to motifs found in promoters of plant mitochondria

(Binder and Brennicke, 2003; Kuhn et al., 2005). In addition, GAA-box has been identified upstream of the YRTA motif in a subclass of NEP promoters (type-Ib; Kapoor and Sugiura, 1999). In contrast to these standard NEP promoters, type-II NEP promoters mapped upstream of the dicot *clpP* gene lack the YRTA motif and are dependent on downstream sequences of the TSS (Weihe and Borner, 1999). Furthermore, it has been shown that the *rrn* operon and certain tRNAs are transcribed from other non-consensus-type NEP promoters [Reviewed by (Liere et al., 2011)].

Although the class I genes have been clarified as being exclusively transcribed by PEP, the genome-wide mapping of TSSs in barley revealed that most genes including photosynthesis genes have both PEP and NEP promoters. It seems likely that NEP supports transcription of photosynthesis genes at the early stage of seedling greening (Zhelyazkova et al., 2012). Interestingly, 73% of NEP-dependent TSSs possess the YRTA motif typical for type-Ia and -Ib NEP promoters, whereas GAA-boxes have been barely mapped upstream of the barley NEP promoters. These results suggest that type-Ia, but not type-Ib NEP promoters play a major role in transcription by NEP in barley chloroplasts. In contrast, type-II NEP promoters, which are dependent on downstream sequences of the TSSs, were identified in barley as well as tobacco.

## THE LARGE TRANSCRIPTION COMPLEX IN HIGHER PLANT CHLOROPLASTS

Two types of PEP-containing preparation have been biochemically isolated in mustard and *Arabidopsis*: soluble RNAP (sRNAP) and plastid transcriptionally active chromosome (pTAC) attached to chloroplast membranes (Hess and Borner, 1999). Transcription by sRNAP is dependent on exogenously added template DNA, whereas the pTAC can initiate transcription from the endogenous chloroplast DNA (Igloi and Kossel, 1992; Krause et al., 2000). Interestingly, protein compositions of highly purified sRNAP fractions are dependent on chloroplast development (Pfannschmidt and Link, 1994). The sRNAP of etioplasts in dark-grown leaves is a naked RNAP without additional subunits similar to the *E. coli* RNAP core complex. Etioplasts convert to photosynthetically active chloroplasts in the presence of light. During chloroplast development in mustard, the RNAP develops a more complex form that contains 13 additional polypeptides (Pfannschmidt and Link, 1994). It seems likely that the simple sRNAP in etioplasts converts to a more complex sRNAP in chloroplasts by recruiting additional components during chloroplast development.

Proteomic analyses of pTAC fractions isolated from mature chloroplasts of *Arabidopsis* and mustard have identified 35 polypeptides including 18 novel proteins termed pTAC1-pTAC18, in addition to PEP core subunits, DNA polymerase, DNA gyrase, Fe-dependent superoxide dismutases (Fe-S-ODs), phosphofructokinase-B type enzymes (PFKB1 and PFKB2), thioredoxin, and three ribosomal proteins (Pfalz et al., 2006). DNA- and/or RNA-binding domains, protein–protein interaction domains, or epitopes with other reported cellular functions have been identified in some of pTAC proteins. Most *Arabidopsis* knockout mutants of pTAC proteins exhibit seedling-lethal symptoms or chlorophyll-deficient phenotypes. PEP-dependent transcription is significantly impaired in the pTAC mutants,

whereas NEP-dependent transcription is up-regulated. These phenotypes and chloroplast gene expression patterns are reminiscent of those of *rpo* mutants (Allison et al., 1996; Hajdukiewicz et al., 1997), suggesting a critical role for pTAC proteins in PEP transcription.

Affinity purification of the tobacco PEP (Suzuki et al., 2004) and more recent analysis of subunits of the PEP complex in mustard (Steiner et al., 2011) and tobacco complex identified at least 10 PEP-associated proteins (PAPs). Recently, chromatin immunoprecipitation assays were performed with one of the typical PAPs, pTAC3/PAP1. The results revealed that pTAC3/PAP1 associates with the PEP complex in all three steps of the transcription cycle including initiation, elongation and termination, suggesting that pTAC3/PAP1 is an essential component of the chloroplast PEP complex (Yagi et al., 2012). Several studies on protein–protein interactions among PAPs have been reported [reviewed in (Pfalz and Pfannschmidt, 2013)]. Almost all PAP genes, except for *Trx-z*, are conserved among all land plants, but not in the green alga *Chlamydomonas*. It seems likely that terrestrial plants may have acquired non-cyanobacterial novel PEP components during land plant evolution to regulate plastid transcription (Pfalz and Pfannschmidt, 2013).

It has been suggested that a series of checkpoints control the establishment of the chloroplast transcription machinery (Steiner et al., 2011; Pfalz and Pfannschmidt, 2013). In imbibed seeds, predominant NEP is responsible for transcription of housekeeping genes. NEP also transcribes chloroplast-encoded *rpo* genes for PEP core subunits to produce a basic PEP-B complex (NEP–PEP cascade). PEP-B is responsible for the major activity in etioplasts and in an early stage of greening. This step may be the first checkpoint. Subsequently, PEP-B associates with PAPs and converts them into a larger PEP complex (PEP-A) during light-dependent chloroplast development. PEP-A formation is strictly dependent on light. Indeed, it has been reported that expression of the *pTAC3/PAP1* gene is induced by light during the greening process (Yagi et al., 2012). PAP mutants mostly show the aberrant development of chloroplasts and transcription of chloroplast-encoded genes, suggesting their essential roles in PEP-A. Furthermore, recent genome-wide analysis of the chloroplast transcriptome revealed reduced expressions of numerous chloroplast tRNAs in several PAPs mutants (pTAC2, pTAC12, MurE, PRIN2), suggesting that PAPs play a major role in tRNA transcription in chloroplasts (Williams-Carrier et al., 2014). Thus PAPs are also responsible for protein translation in chloroplasts. Therefore, the assembly of PAPs in the PEP-A complex may be the second checkpoint in the establishment of the chloroplast transcription machinery. To prevent uncontrolled chloroplast development under adverse conditions, these checkpoints likely play critical roles in the control of chloroplast gene transcription.

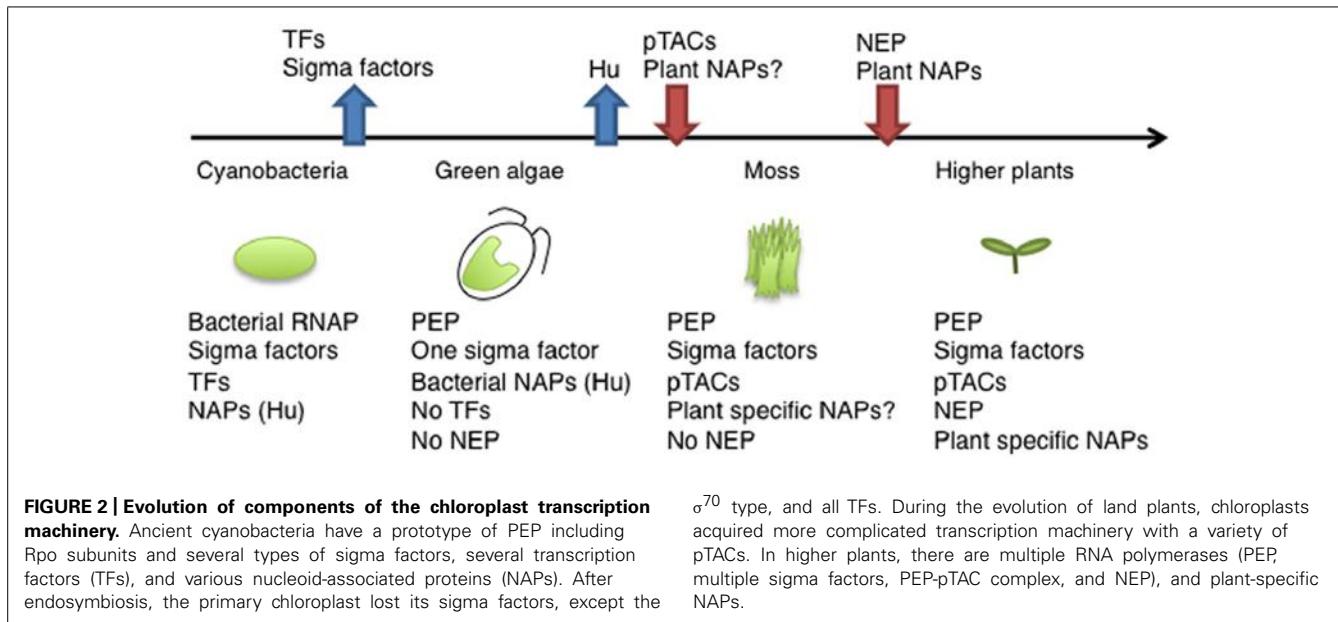
## THE PLASTID NUCLEOIDS: DYNAMICS AND UNIQUE COMPONENTS

The plastid DNA exists as large protein-DNA complexes named the plastid nucleoid. Plastid nucleoids contain an average of 10–20 copies of the plastid DNA (Kuroiwa, 1991), and their size, shape, and distribution vary depending on the plastid type (Miyamura

et al., 1986; Sato et al., 1997). Each chloroplast contains ~20 nucleoids that are randomly located on the thylakoid membranes. Immature proplastids in seeds contain only one nucleoid that is located at the center of the organelle. The plastid nucleoids divide into a few small dots and redistribute to the inner envelope membranes during early chloroplast development. At a later stage of chloroplast development, nucleoids are relocated to the thylakoid membranes. It has been suggested that plastid nucleoid organization and dynamics are involved in the regulation of plastid function, gene expression and differentiation. Two DNA-binding proteins, PEND and MFP1, are likely responsible for the association of nucleoids with chloroplast membranes (Sato et al., 1998; Jeong et al., 2003; **Figure 1B**).

In *E. coli*, chromosome DNA packaging patterns affect gene expression, and are regulated by nucleoid-associated proteins (NAPs) such as HU, H-NS, and FIS [reviewed by (Dillon and Dorman, 2010)]. Among bacterial NAPs, HU is one of the major DNA-binding proteins and is involved in chromosome DNA packaging. HU-like proteins (HLPs) are conserved in cyanobacteria, the red alga *Cyanidioschyzon merolae* (Kobayashi et al., 2002), and the green alga *Chlamydomonas* (Karcher et al., 2009). The HLP in *Chlamydomonas* has roles in nucleoid maintenance and gene expression, indicating conserved roles of HU during chloroplast evolution (Karcher et al., 2009). However, land plants including mosses and flowering plants have not only lost the HU genes, but also other prokaryotic DNA-binding proteins (**Figure 2**). Nevertheless, atomic force microscopy observations revealed that plastid nucleoids are highly organized and form a beads-on-a-string structure similar to that observed in bacterial nucleoids, suggesting that another host cell-derived DNA-binding protein took over the functions of HU (Melonek et al., 2012). Recently, eukaryotic SWIB (SWI/SNF complex B) domain containing proteins have been identified from the proteome of a further-enriched pTAC fraction (TAC-II) of spinach chloroplasts (Melonek et al., 2012). SWIB4 that has a histone H1 motif, can functionally complement an *E. coli* mutant lacking the histone-like nucleoid structuring protein H-NS, indicating that SWIB4 is the most likely counterpart of the bacterial NAPs in chloroplasts. EM observation of isolated pTAC identified chromatin-like beaded structures with several protruding DNA loops, suggesting that pTACs represent a subdomain of the chloroplast nucleoid (Yoshida et al., 1978; Briat et al., 1982). These findings suggest that pTAC forms a central core of the plastid nucleoid and a transcription factory (**Figure 1B**).

The proteomes of highly enriched nucleoid fractions have been characterized in maize proplastids and mature chloroplasts (Majeran et al., 2012). As expected, the chloroplast nucleoids contain all PEP core Rpos and PAPs, and almost all other pTAC proteins. Furthermore, additional proteins involved in post-transcriptional processes, such as pentatricopeptide repeat proteins (PPR proteins), mitochondrial transcription factor (mTERF)-domain proteins, 70S ribosomes and ribosome assembly factors have been identified in the proteome of the chloroplast nucleoids, suggesting that several post-transcriptional events including RNA processing, splicing and editing, and translation, occur in nucleoids, and that these processes are co-regulated with transcription (**Figure 1B**). Human mitochondrial nucleoids have



been shown to form layered structures, the central core involved in replication and transcription, and the peripheral region where translation and complex assembly may occur (Bogenhagen et al., 2008). By analogy, the further characterization of plastid nucleoids will provide insights into the structural specialization of plastid nucleoids; DNA maintenance and transcription in a core domain and various aspects of RNA metabolism in several subdomains.

## PERSPECTIVE

Recent proteomic and transcriptomic researches and the development of novel ChIP and imaging technologies have advanced the understanding of the molecular basis of RNAP complexes and nucleoid architecture. In land plants, neither the nuclear nor chloroplast genome encodes prokaryotic transcription factors and nucleoid proteins, whereas chloroplasts retain prokaryotic-type RNAP (Figure 2). In fact, land plants have a number of novel host cell-derived transcription regulators and DNA-binding proteins that are involved in the regulation of chloroplast transcription. Thus, it seems likely that chloroplast transcription is mediated by a hybrid system of prokaryotic and eukaryotic origin. Further molecular characterization of pTACs and plastid nucleoid proteins would provide novel insights into the unique plastid gene expression system and as yet known mechanisms of plastid differentiation.

## ACKNOWLEDGMENTS

We would like to thank Y. Ishizaki for critical reading of the manuscript. This work was supported by JSPS and MEXT Grants-in-Aid for Scientific Research (24657036, 25291065, 25120723) and a grant from the Mitsubishi Foundation to Takashi Shiina, and a Grant-in-Aid for JSPS Fellows to Yusuke Yagi.

## REFERENCES

- Abdallah, F., Salamini, F., and Leister, D. (2000). A prediction of the size and evolutionary origin of the proteome of chloroplasts of *Arabidopsis*. *Trends Plant Sci.* 5, 141–142. doi: 10.1016/S1360-1385(00)01574-0
- Allison, L. A., Simon, L. D., and Maliga, P. (1996). Deletion of rpoB reveals a second distinct transcription system in plastids of higher plants. *EMBO J.* 15, 28 02–2809.
- Binder, S., and Brennicke, A. (2003). Gene expression in plant mitochondria: transcriptional and post-transcriptional control. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 358, 181–189. doi: 10.1098/rstb.2002.1179
- Bogenhagen, D. F., Rousseau, D., and Burke, S. (2008). The layered structure of human mitochondrial DNA nucleoids. *J. Biol. Chem.* 283, 3665–3675. doi: 10.1074/jbc.M708444200
- Briat, J. F., Gigot, C., Laulhere, J. P., and Mache, R. (1982). Visualization of a spinach plastid transcriptionally active DNA-protein complex in a highly condensed structure. *Plant Physiol.* 69, 1205–1211. doi: 10.1104/pp.69.5.1205
- Chang, C. C., Sheen, J., Bligny, M., Niwa, Y., Lerbs-Mache, S., and Stern, D. B. (1999). Functional analysis of two maize cDNAs encoding T7-like RNA polymerases. *Plant Cell* 11, 911–926. doi: 10.1105/tpc.11.5.911
- De Santis-Maclossek, G., Kofer, W., Bock, A., Schöch, S., Maier, R. M., Wanner, G., et al. (1999). Targeted disruption of the plastid RNA polymerase genes rpoA, B and C1: molecular biology, biochemistry and ultrastructure. *Plant J.* 18, 477–489. doi: 10.1046/j.1365-313X.1999.00473.x
- Dillon, S. C., and Dorman, C. J. (2010). Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. *Nat. Rev. Microbiol.* 8, 185–195. doi: 10.1038/nrmicro2261
- Emanuel, C., Weihe, A., Grainer, A., Hess, W. R., and Borner, T. (2004). Chloroplast development affects expression of phage-type RNA polymerases in barley leaves. *Plant J.* 38, 460–472. doi: 10.1111/j.0960-7412.2004.02060.x
- Favory, J. J., Kobayashi, M., Tanaka, K., Peltier, G., Kreis, M., Valay, J. G., et al. (2005). Specific function of a plastid sigma factor for *ndhf* gene transcription. *Nucleic Acids Res.* 33, 5991–5999. doi: 10.1093/nar/gki908
- Gatenby, A. A., Castleton, J. A., and Saul, M. W. (1981). Expression in *E. coli* of maize and wheat chloroplast genes for large subunit of ribulose bisphosphate carboxylase. *Nature* 291, 117–121. doi: 10.1038/291117a0
- Gruissem, W., and Zurawski, G. (1985). Analysis of promoter regions for the spinach chloroplast *rbcL*, *atpB* and *psbA* genes. *EMBO J.* 4, 33 75–3383.
- Hajdukiewicz, P. T., Allison, L. A., and Maliga, P. (1997). The two RNA polymerases encoded by the nuclear and the plastid compartments transcribe distinct groups of genes in tobacco plastids. *EMBO J.* 16, 4041–4048. doi: 10.1093/emboj/16.13.4041
- Han, C. D., Coe, E. H. Jr., and Martienssen, R. A. (1992). Molecular cloning and characterization of *iojap* (*ij*), a pattern striping gene of maize. *EMBO J.* 11, 4037–4046.

- Hanaoka, M., Kanamaru, K., Takahashi, H., and Tanaka, K. (2003). Molecular genetic analysis of chloroplast gene promoters dependent on SIG2, a nucleus-encoded sigma factor for the plastid-encoded RNA polymerase, in *Arabidopsis thaliana*. *Nucleic Acids Res.* 31, 7090–7098. doi: 10.1093/nar/gkg935
- Hanaoka, M., Kato, M., Anma, M., and Tanaka, K. (2012). SIG1, a sigma factor for the chloroplast RNA polymerase, differently associates with multiple DNA regions in the chloroplast chromosomes in vivo. *Int. J. Mol. Sci.* 13, 12182–12194. doi: 10.3390/ijms131012182
- Hara, K., Sugita, M., and Aoki, S. (2001). Cloning and characterization of the cDNA for a plastid sigma factor from the moss *Physcomitrella patens*. *Biochim. Biophys. Acta* 1517, 302–306. doi: 10.1016/S0167-4781(00)00250-5
- Hedtke, B., Börner, T., and Weihe, A. (1997). Mitochondrial and chloroplast phage-type RNA polymerases in *Arabidopsis*. *Science* 277, 809–811 doi: 10.1126/science.277.5327.809
- Hedtke, B., Börner, T., and Weihe, A. (2000). One RNA polymerase serving two genomes. *EMBO Rep.* 1, 435–440. doi: 10.1093/embo-reports/kvd086
- Hess, W. R., and Borner, T. (1999). Organellar RNA polymerases of higher plants. *Int. Rev. Cytol.* 190, 1–59. doi: 10.1016/S0074-7696(08)62145-2
- Hess, W. R., Muller, A., Nagy, F., and Börner, T. (1994). Ribosome-deficient plastids affect transcription of light-induced nuclear genes: genetic evidence for a plastid-derived signal. *Mol. Gen. Genet.* 242, 305–312. doi: 10.1007/BF00280420
- Hess, W. R., Prombona, A., Fieder, B., Sutrarnanian, A. R., and Börner, T. (1993). Chloroplast rps15 and the rpoB/C1/C2 gene cluster are strongly transcribed in ribosome-deficient plastids: evidence for a functioning non-chloroplast-encoded RNA polymerase. *EMBO J.* 12, 563–571.
- Hirata, N., Yonekura, D., Yanagisawa, S., and Iba, K. (2004). Possible involvement of the 5'-flanking region and the 5'UTR of plastid accD gene in NEP-dependent transcription. *Plant Cell Physiol.* 45, 176–186. doi: 10.1093/pcp/pch021
- Hotto, A. M., Schmitz, R. J., Fei, Z., Ecker, J. R., and Stern, D. B. (2011). Unexpected diversity of chloroplast noncoding RNAs as revealed by deep sequencing of the *Arabidopsis* transcriptome. *G3 (Bethesda)* 1, 559–570. doi: 10.1534/g3.111.000752
- 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. doi: 10.1104/pp.104.053033
- Igloi, G. L., and Kossel, H. (1992). The transcriptional apparatus of chloroplasts. *Crit. Rev. Plant Sci.* 10, 525–558. doi: 10.1080/07352689209382326
- Ikeda, T. M., and Gray, M. W. (1999). Identification and characterization of T3/T7 bacteriophage-like RNA polymerase sequences in wheat. *Plant Mol. Biol.* 40, 567–578. doi: 10.1023/A:1006203928189
- Ishihama, A. (2000). Functional modulation of *Escherichia coli* RNA polymerase. *Annu. Rev. Microbiol.* 54, 499–518. doi: 10.1146/annurev.micro.54.1.499
- Ishizaki, Y., Tsunoyama, Y., Hatano, K., Ando, K., Kato, K., Shinmyo, A., et al. (2005). A nuclear-encoded sigma factor, *Arabidopsis* SIG6, recognizes sigma-70 type chloroplast promoters and regulates early chloroplast development in cotyledons. *Plant J.* 42, 133–144. doi: 10.1111/j.1365-313X.2005.02362.x
- Jeong, S. Y., Rose, A., and Meier, I. (2003). MFP1 is a thylakoid-associated, nucleoid-binding protein with a coiled-coil structure. *Nucleic Acids Res.* 31, 5175–5185. doi: 10.1093/nar/gkg693
- Kabeya, Y., Hashimoto, K., and Sato, N. (2002). Identification and characterization of two phage-type RNA polymerase cDNAs in the moss *Physcomitrella patens*: implication of recent evolution of nuclear-encoded RNA polymerase of plastids in plants. *Plant Cell Physiol.* 43, 245–255. doi: 10.1093/pcp/pcf041
- Kanamaru, K., Nagashima, A., Fujiwara, M., Shimada, H., Shirano, Y., Nakabayashi, K., et al. (2001). An *Arabidopsis* sigma factor (SIG2)-dependent expression of plastid-encoded tRNAs in chloroplasts. *Plant Cell Physiol.* 42, 1034–1043. doi: 10.1093/pcp/pce155
- Kanazawa, T., Ishizaki, K., Kohchi, T., Hanaoka, M., and Tanaka, K. (2013). Characterization of four nuclear-encoded plastid RNA polymerase sigma factor genes in the liverwort *Marchantia polymorpha*: blue-light- and multiple stress-responsive sig5 was acquired early in the emergence of terrestrial plants. *Plant Cell Physiol.* 54, 1736–1748. doi: 10.1093/pcp/pct119
- Kapoor, S., and Sugiura, M. (1999). Identification of two essential sequence elements in the nonconsensus type II PatpB-290 plastid promoter by using plastid transcription extracts from cultured tobacco BY-2 cells. *Plant Cell* 11, 1799–1810. doi: 10.1105/tpc.11.9.1799
- Kapoor, S., Suzuki, J. Y., and Sugiura, M. (1997). Identification and functional significance of a new class of non-consensus-type plastid promoters. *Plant J.* 11, 327–337. doi: 10.1046/j.1365-313X.1997.11020327.x
- Karcher, D., Koster, D., Schadach, A., Klevesath, A., and Bock, R. (2009). The *Chlamydomonas* chloroplast HLP protein is required for nucleoid organization and genome maintenance. *Mol. Plant* 2, 1223–1232. doi: 10.1093/mp/ssp083
- Kobayashi, T., Takahara, M., Miyagishima, S. Y., Kuroiwa, H., Sasaki, N., Ohta, N., et al. (2002). Detection and localization of a chloroplast-encoded HU-like protein that organizes chloroplast nucleoids. *Plant Cell* 14, 1579–1589. doi: 10.1105/tpc.002717
- Krause, K., Maier, R. M., Kofer, W., Krupinska, K., and Herrmann, R. G. (2000). Disruption of plastid-encoded RNA polymerase genes in tobacco: expression of only a distinct set of genes is not based on selective transcription of the plastid chromosome. *Mol. Gen. Genet.* 263, 1022–1030. doi: 10.1007/PL00008690
- Kuhn, K., Weihe, A., and Borner, T. (2005). Multiple promoters are a common feature of mitochondrial genes in *Arabidopsis*. *Nucleic Acids Res.* 33, 337–346. doi: 10.1093/nar/gki179
- Kuroiwa, T. (1991). The replication, differentiation, and inheritance of plastids with emphasis on the concept of organelle nuclei. *Int. Rev. Cytol.* 128, 1–58. doi: 10.1016/S0074-7696(08)60496-9
- Lerbs-Mache, S. (2011). Function of plastid sigma factors in higher plants: regulation of gene expression or just preservation of constitutive transcription? *Plant Mol. Biol.* 76, 235–249. doi: 10.1007/s11103-010-9714-4
- 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.* 18, 249–257. doi: 10.1093/emboj/18.1.249
- Liere, K., Weihe, A., and Borner, T. (2011). The transcription machineries of plant mitochondria and chloroplasts: composition, function, and regulation. *J. Plant Physiol.* 168, 1345–1360. doi: 10.1016/j.jplph.2011.01.005
- Maier, U. G., Bozarth, A., Funk, H. T., Zauner, S., Rensing, S. A., Schmitz-Linneweber, C., et al. (2008). Complex chloroplast RNA metabolism: just debugging the genetic programme? *BMC Biol.* 6:36. doi: 10.1186/1741-707-6-36
- Majeran, W., Friso, G., Asakura, Y., Qu, X., Huang, M., Ponnala, L., et al. (2012). Nucleoid-enriched proteomes in developing plastids and chloroplasts from maize leaves: a new conceptual framework for nucleoid functions. *Plant Physiol.* 158, 156–189. doi: 10.1104/pp.111.188474
- Melonk, J., Matros, A., Trosch, M., Mock, H. P., and Krupinska, K. (2012). The core of chloroplast nucleoids contains architectural SWIB domain proteins. *Plant Cell* 24, 3060–3073. doi: 10.1105/tpc.112.099721
- Miyamura, S., Nagata, T., and Kuroiwa, T. (1986). Quantitative fluorescence microscopy on dynamic changes of plastid nucleoids during wheat development. *Protoplasma* 133, 66–72. doi: 10.1007/BF01293188
- Nagashima, A., Hanaoka, M., Shikanai, T., Fujiwara, M., Kanamaru, K., Takahashi, H., et al. (2004). The multiple-stress responsive plastid sigma factor, SIG5, directs activation of the psbd blue light-responsive promoter (BLRP) in *Arabidopsis thaliana*. *Plant Cell Physiol.* 45, 357–368. doi: 10.1093/pcp/pch050
- Pfalz, J., Liere, K., Kandlbinder, A., Dietz, K. J., and Oelmuller, R. (2006). pTAC2, -6, and -12 are components of the transcriptionally active plastid chromosome some that are required for plastid gene expression. *Plant Cell* 18, 176–197. doi: 10.1105/tpc.105.036392
- Pfalz, J., and Pfannschmidt, T. (2013). Essential nucleoid proteins in early chloroplast development. *Trends Plant Sci.* 18, 186–194. doi: 10.1016/j.tplants.2012.11.003
- Pfannschmidt, T., and Link, G. (1994). Separation of two classes of plastid DNA-dependent RNA polymerases that are differentially expressed in mustard (*Sinapis alba* L.) seedlings. *Plant Mol. Biol.* 25, 69–81. doi: 10.1007/BF00024199
- Richter, U., Kiessling, J., Hedtke, B., Decker, E., Reski, R., Borner, T., et al. (2002). Two RpoT genes of *Physcomitrella patens* encode phage-type RNA polymerases with dual targeting to mitochondria and plastids. *Gene* 290, 95–105. doi: 10.1016/S0378-1119(02)00583-8
- Richter, U., Richter, B., Weihe, A., and Borner, T. (2013). A third mitochondrial RNA polymerase in the moss *Physcomitrella patens*. *Curr. Genet.* 60, 25–34. doi: 10.1007/s00294-013-0405-y

- Sakai, A., Saito, C., Inada, N., and Kuroiwa, T. (1998). Transcriptional activities of the chloroplast-nuclei and proplastid-nuclei isolated from tobacco exhibit different sensitivities to tagetitoxin: implication of the presence of distinct RNA polymerases. *Plant Cell Physiol.* 39, 928–934. doi: 10.1093/oxfordjournals.pcp.a029456
- Sato, N., Misumi, O., Shinada, Y., Sasaki, M., and Yoine, M. (1997). Dynamics of localization and protein composition of plastid nucleoids in light-grown pea seedlings. *Protoplasma* 200, 163–173. doi: 10.1007/BF01283292
- Sato, N., Ohshima, K., Watanabe, A., Ohta, N., Nishiyama, Y., Joyard, J., et al. (1998). Molecular characterization of the PEND protein, a novel bZIP protein present in the envelope membrane that is the site of nucleoid replication in developing plastids. *Plant Cell* 10, 859–872. doi: 10.1105/tpc.10.5.859
- Shiina, T., Ishizaki, Y., Yagi, Y., and Nakahira, Y. (2009). Function and evolution of plastid sigma factors. *Plant Biotechnol.* 26, 57–66. doi: 10.5511/plantbiotechnology.26.57
- 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. doi: 10.1016/S0074-7696(05)44001-2
- Steiner, S., Schroter, Y., Pfalz, J., and Pfannschmidt, T. (2011). Identification of essential subunits in the plastid-encoded RNA polymerase complex reveals building blocks for proper plastid development. *Plant Physiol.* 157, 1043–1055. doi: 10.1104/pp.111.184515
- Strittmatter, G., Gozdzicka-Jozefiak, A., and Kossel, H. (1985). Identification of an rRNA operon promoter from Zea mays chloroplasts which excludes the proximal tRNA<sup>Val</sup>GAC from the primary transcript. *EMBO J.* 4, 599–604.
- Suzuki, J. Y., Ytterberg, A. J., Beardslee, T. A., Allison, L. A., Wijk, K. J., and Maliga, P. (2004). Affinity purification of the tobacco plastid RNA polymerase and in vitro reconstitution of the holoenzyme. *Plant J.* 40, 164–172. doi: 10.1111/j.1365-313X.2004.02195.x
- Tsunoyama, Y., Ishizaki, Y., Morikawa, K., Kobori, M., Nakahira, Y., Takeba, G., et al. (2004). Blue light-induced transcription of plastid-encoded psbD gene is mediated by a nuclear-encoded transcription initiation factor, AtSig5. *Proc. Natl. Acad. Sci. U.S.A.* 101, 3304–3309. doi: 10.1073/pnas.0308362101
- 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. doi: 10.1016/S0014-5793(02)02538-3
- Ueda, M., Takami, T., Peng, L., Ishizaki, K., Kohchi, T., Shikanai, T., et al. (2013). Subfunctionalization of sigma factors during the evolution of land plants based on mutant analysis of liverwort (*Marchantia polymorpha* L.) MpSIG1. *Genome Biol. Evol.* 5, 1836–1848. doi: 10.1093/gbe/evt137
- Weihe, A., and Borner, T. (1999). Transcription and the architecture of promoters in chloroplasts. *Trends Plant Sci.* 4, 169–170. doi: 10.1016/S1360-1385(99)01407-7
- Williams-Carrier, R., Zoschke, R., Belcher, S., Pfalz, J., and Barkan, A. (2014). A major role for the plastid-encoded RNA polymerase complex in the expression of plastid tRNAs. *Plant Physiol.* 164, 239–248. doi: 10.1104/pp.113.01104
- Yagi, Y., Ishizaki, Y., Nakahira, Y., Tozawa, Y., and Shiina, T. (2012). Eukaryotic-type plastid nucleoid protein pTAC3 is essential for transcription by the bacterial-type plastid RNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.* 109, 7541–7546. doi: 10.1073/pnas.1119403109
- Yin, C., Richter, U., Börner, T., and Weihe, A. (2009). Evolution of phage-type RNA polymerases in higher plants: characterization of the single phage-type RNA polymerase gene from *Selaginella moellendorffii*. *J. Mol. Evol.* 68, 528–538. doi: 10.1007/s00239-009-9229-2
- Yin, C., Richter, U., Börner, T., and Weihe, A. (2010). Evolution of plant phage-type RNA polymerases: the genome of the basal angiosperm *Nuphar advena* encodes two mitochondrial and one plastid phage-type RNA polymerases. *BMC Evol. Biol.* 10:379. doi: 10.1186/1471-2148-10-379
- Yoshida, Y., Laulhere, J. P., Rozier, C., and Mache, R. (1978). Visualization of folded chloroplast DNA from spinach. *Biol. Cell* 32, 187–190.
- 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. doi: 10.1093/nar/gk11067
- Zhelyazkova, P., Sharma, C. M., Forstner, K. U., Liere, K., Vogel, J., and Borner, T. (2012). The primary transcriptome of barley chloroplasts: numerous noncoding RNAs and the dominating role of the plastid-encoded RNA polymerase. *Plant Cell* 24, 123–136. doi: 10.1105/tpc.111.089441

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 30 December 2013; paper pending published: 31 January 2014; accepted: 06 February 2014; published online: 25 February 2014.

Citation: Yagi Y and Shiina T (2014) Recent advances in the study of chloroplast gene expression and its evolution. *Front. Plant Sci.* 5:61. doi: 10.3389/fpls.2014.00061

This article was submitted to Plant Physiology, a section of the journal *Frontiers in Plant Science*.

Copyright © 2014 Yagi and Shiina. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# AtSIG6 and other members of the sigma gene family jointly but differentially determine plastid target gene expression in *Arabidopsis thaliana*

Sylvia Bock, Jennifer Ortelt and Gerhard Link\*

Department of Biology and Biotechnology, University of Bochum, Bochum, Germany

**Edited by:**

Thomas Pfannschmidt, University Joseph Fourier Grenoble, France

**Reviewed by:**

David G. Oppenheimer, University of Florida, USA  
Stefan Gleissberg, Ohio University, USA

**\*Correspondence:**

Gerhard Link, Department of Biology and Biotechnology, University of Bochum, Universitaetsstr. 150, D-44780 Bochum, Germany  
e-mail: gerhard.link@rub.de

Plants contain a nuclear gene family for plastid sigma factors, i.e., proteins that associate with the “bacterial-type” organellar RNA polymerase and confer the ability for correct promoter binding and transcription initiation. Questions that are still unresolved relate to the “division of labor” among members of the sigma family, both in terms of their range of target genes and their temporal and spatial activity during development. Clues to the *in vivo* role of individual sigma genes have mainly come from studies of sigma knockout lines. Despite its obvious strengths, however, this strategy does not necessarily trace-down causal relationships between mutant phenotype and a single sigma gene, if other family members act in a redundant and/or compensatory manner. We made efforts to reduce the complexity by genetic crosses of *Arabidopsis* single mutants (with focus on a chlorophyll-deficient *sig6* line) to generate double knockout lines. The latter typically had a similar visible phenotype as the parental lines, but tended to be more strongly affected in the transcript patterns of both plastid and sigma genes. Because triple mutants were lethal under our growth conditions, we exploited a strategy of transformation of single and double mutants with RNAi constructs that contained sequences from the unconserved sigma region (UCR). These RNAi/knockout lines phenotypically resembled their parental lines, but were even more strongly affected in their plastid transcript patterns. Expression patterns of sigma genes revealed both similarities and differences compared to the parental lines, with transcripts at reduced or unchanged amounts and others that were found to be present in higher (perhaps compensatory) amounts. Together, our results reveal considerable flexibility of gene activity at the levels of both sigma and plastid gene expression. A (still viable) “basal state” seems to be reached, if 2–3 of the 6 *Arabidopsis* sigma genes are functionally compromised.

**Keywords:** chloroplast transcription, plant sigma factors, nuclear gene family, knockout mutants, RNA interference, plastid target gene expression

## INTRODUCTION

Despite the small number of genes in the chloroplast genome (Sugiura, 1992), plastid transcription is a surprisingly complex process. It involves two different RNA polymerases commonly named NEP (nuclear-encoded polymerase) and PEP (plastid-encoded polymerase) (Hedtke et al., 1997; Maliga, 1998). The latter is surrounded by multiple transcription factors (Shiina et al., 2005), important representatives of which are the nuclear-encoded sigma factors. Like their bacterial counterparts (Ishihama, 1988; Burgess and Anthony, 2001), these plant factors are thought to direct the (PEP) polymerase complex to its cognate promoters and ensure faithful transcription initiation. Again, as is the case in bacteria, the plastids of higher plants typically contain more than a single sigma factor species (e.g., a family comprising six proteins ATSIG1 - 6 in *Arabidopsis thaliana*) (Isono et al., 1997; Tanaka et al., 1997; Fujiwara et al., 2000; Shiina et al., 2009). This obvious analogy has therefore stimulated research addressing the role of individual members of the plant sigma factor family.

Work carried out with *Arabidopsis* knockout lines containing T-DNA insertions in single sigma genes has provided an initial picture, demonstrating the presence or absence of a recognizable mutant phenotype depending on the affected sigma gene as well as the developmental stage investigated (Tsunoyama et al., 2002; Hanaoka et al., 2003; Privat et al., 2003; Nagashima et al., 2004; Favory et al., 2005; Ishizaki et al., 2005; Loschelder et al., 2006; Schweer et al., 2006, 2009; Zghidi et al., 2007). A readily noticeable phenotype is evident for instance in the case of AtSIG6, where mutant lines tend to have strong chlorophyll deficiency and altered plastid target gene expression patterns, yet only in seedlings but not, e.g., in plants during the subsequent rosette leaf stages (Ishizaki et al., 2005; Loschelder et al., 2006; Schweer et al., 2006, 2009).

Unlike the situation in bacteria (Ishihama, 1988), none of the plastid sigma factors seems to have a “primary” essential role in a sense that its loss would confer a lethal or seriously compromised phenotype (Ortelt and Link, 2014). What then might be

the reasons for the variable (non-lethal) phenotypes noticeable in plant sigma knockout lines? A perhaps most direct explanation would be that the plastid factors function in a partially overlapping manner, yet in a highly flexible way at different developmental stages, in different organs, and under variable environmental conditions. Clues supporting this idea come, e.g., from the modular architecture of the plant sigma factors, each of which has a C-terminal conserved region (CR) responsible for basal sigma activity and a N-terminal unconserved region (UCR) of regulatory function (Ortelt and Link, 2014).

Nevertheless, it seems likely that the current functional description of the underlying network is not yet complete. For instance, knocking out one single sigma gene may or may not have the consequence that a second (or third, fourth etc.) member of the factor family is functionally recruited in a specific developmental context. To reduce the complexity of the *Arabidopsis* sigma family, we took advantage of single and double mutant lines and also adopted RNAi (RNA interference) techniques (Fire et al., 1998) for that purpose. We then analyzed the gene expression situation (at RNA level) both for plastid target genes as well as for the members of sigma gene family themselves. We reasoned that such work could add novel information to help explain the flexibility of the plant sigma factor network and its phenotypic consequences.

## MATERIALS AND METHODS

### ARABIDOPSIS MUTANT AND RNAi LINES, GROWTH CONDITIONS, SCREENING

Single mutant lines *sig1-2* ("sig1"), *sig3-2* ("sig3"), and *sig6-2* ("sig6") in the Col-0 (ecotype Columbia) background of *Arabidopsis thaliana* were obtained from the GABI-Kat collection of T-DNA insertion lines ([www.gabi-kat.de](http://www.gabi-kat.de)) (Rosso et al., 2003). Double knockouts *sig1 sig6* and *sig6 sig3* were generated from the single mutants by genetic crosses. Both single and double knockouts were transformed with sigma UCR sequences cloned into RNAi vector pHELLSGATE12 (Wesley et al., 2001) ([www.csiro.au](http://www.csiro.au)) using the Gateway system (Life Technologies). PCR-amplified sigma cDNA representing the full-size UCR (555 bp after the start codon for AtSIG1, 768 bp for AtSIG3, and 696 bp for AtSIG6) was inserted into the donor vector pDONR/zeo (Life Technologies) and mobilized to the destination vector as described in the Gateway manual ([www.invitrogen.com](http://www.invitrogen.com)). It was then introduced into *Arabidopsis* by Agrobacterium-mediated floral dip transformation (Clough and Bent, 1998) and progeny were screened using antibiotic selection, PCR and Southern blot analyses as described (Loschelder et al., 2006). The criteria for successful generation of double mutants, i.e., absence of an amplified PCR product using primers that flank the T-DNA insertion and presence of a product using a primer pair across the junction between T-DNA and sigma gene sequence, were tested for each candidate line ([signal.salk.edu/tdnaprimer.2.html](http://signal.salk.edu/tdnaprimer.2.html)). Only those lines that fulfilled these requirements were further propagated and subsequent experiments were carried out with at least three independent T3 lines for each construct. Seeds were sown on MS medium (Murashige and Skoog, 1962) containing 0.4% (w/v) gelrite and 1% (w/v) sucrose, stratified at 4°C for 2 d, and then transferred to 24°C for germination and growth under short-day conditions (8 h light/16 h

dark, 60 mmol m<sup>-2</sup> s<sup>-1</sup>). Seedlings were harvested at day 6 or 10, or growth was continued until day 14, at which time plantlets were transferred to sterile soil for another 6 d until day 20. Tissue samples were immediately frozen in liquid nitrogen and stored at -85°C until use.

### PHYSIOLOGICAL PARAMETERS (CHLOROPHYLL CONTENT, ROOT LENGTH MEASUREMENTS)

Chlorophyll content was measured using 8 replicates of 10 seedlings or young plantlets at each time-point (6 d, 10 d, 20 d after sowing) and measurements were repeated twice using independently grown plant material. Following weighing of seedlings, they were ground in 80% (v/v) acetone. The extract was centrifuged at 10,000 g for 10 min and the supernatant was used for photometric chlorophyll determination at 663 and 645 nm. For root lengths measurements, seedlings were grown as described above, yet using vertical positioning of Petri dishes. Length determination was carried out using ImageJ (<http://imagej.nih.gov/ij>) followed by graphical presentation. Values were means of three independent replicates each obtained from 20 seedlings.

### RT-qPCR DETECTION OF TRANSCRIPTS

Total seedling RNA (2 µg) was prepared and reverse-transcribed into cDNA as described by Loschelder et al. (2006), yet using oligo-(dT)-primers and the Moloney Murine Leukemia Virus Reverse Transcriptase (M-MLV; Promega). Quantitative real-time RT-PCR (RT-qPCR) was in an Illumina Eco System ([www.illumina.com](http://www.illumina.com)) with KAPA SybrFast qPCR MasterMix Universal (Peqlab) in a volume of 20 µl. Upon amplification (50°C for 2 min, 95°C for 10 min, followed by 40 cycles at 95°C for 10 s, 60°C for 30 s, and 72°C for 15 s), melting curve analysis was carried out, using the *Actin2* C<sub>t</sub> value as a reference for normalization. For each primer pair, the real-time experiments were carried out with at least three biological and technical replicate samples. Primers (**Table 1**) were selected on the basis of minimal sequence homology among the members of the *Arabidopsis* sigma gene family (Schweer et al., 2009).

### NORTHERN BLOT HYBRIDIZATION

Total RNA was isolated as described by Chomczynski and Sacchi (1987). RNA (2 µg per lane) was fractionated, blotted, and hybridized, using gene-specific RNA probes (**Table 1**). The latter were generated by cloning of PCR-amplified gene segments into pGEM-T Easy (Promega), followed by *in vitro* transcription of constructs using T7 RNA polymerase (Promega) as described (Loschelder et al., 2006; Schweer et al., 2006). DIG (digoxigenin) labeling conditions and immunological detection using anti-digoxigenin antibody (Roche) were as detailed in the DIG user's manual ([www.roche-applied-science.com](http://www.roche-applied-science.com)).

## RESULTS

### CHARACTERIZATION OF SIGMA-DEFICIENT SINGLE AND DOUBLE MUTANTS

To raise information on the fine-tuning and complex regulation within the sigma gene family from *Arabidopsis*, we analyzed single and double sigma mutants. Starting material were the three single mutant lines *sig1-2*, *sig3-2*, and *sig6-2* defective

**Table 1 | Primers used for construction of RNAi lines, northern hybridization, and RT-qPCR.**

| Gene                          | Forward primer (5' → 3')  | Reverse primer (5' → 3')  |
|-------------------------------|---|---|
| <b>RNAi CONSTRUCTS</b>        |   |   |
| <i>SIG1</i>                   | GGGGACAAGTTGTACAAAAAA<br>GCAGGCTCATGGCTACTGC<br>AGCTGTTATAGGA     | GGGGACCACTTGTACAAGAAAG<br>CTGGGTCTACTTCCACTAGAAA<br>CATCAGAAAC  |
| <i>SIG2</i>                   | GGGGACAAGTTGTACAAAAAA<br>AGCAGGCTCCATGTCTTCTGT<br>CTTCTTCCTCAG    | GGGGACCACTTGTACAAGAAAG<br>CTGGGTCTAATTATGATCAACTT<br>CCTGCGCAAC |
| <i>SIG4</i>                   | GGGGACAAGTTGTACAAAAAA<br>AGCAGGCTCCATGGCGACGAC<br>GATTCCCCTACTACA | GGGGACCACTTGTACAAGAAAG<br>CTGGGTCTATCCAACAACAGGA<br>ACACTAATTGT |
| <b>NORTHERN HYBRIDIZATION</b> |   |   |
| <i>atpB</i>                   | GGGAAACCGTTGATAATT  | AACGCTCAATTTCGTGCT  |
| <i>rbcL</i>                   | TCGGTGGAGGAACCTTAGGC  | TGCAAGATCACGTCCTCAT   |
| <i>RbcS</i>                   | TGGCTCCTCATGCTCTCTTC  | GCTGAGGCAGGTCACAAC  |
| <i>rrn16</i>                  | CGGTATCTGGGAATAAGC  | GAAATCCCTCTGCCCTAC  |
| <i>trnE</i>                   | CTACCCCCAGGGGAAGTC  | AGGACTTGGTGTCTGCTACC  |
| <b>RT-qPCR</b>                |   |   |
| <i>SIG1</i>                   | TTTCTGCATGGTGGTTGA  | GTGGCACAGACAATCGAGA   |
| <i>SIG2</i>                   | TGTGCCCTAAACACAACAA   | GCTTGCACATCAAGCATA  |
| <i>SIG3</i>                   | GTTGGGGGCTGCTGAGTTAT  | TTCGGCGTAGTATCCACCAA  |
| <i>SIG4</i>                   | CGTCTCCTCTGCTCCCTACA  | TAGTGTACCGCAAACCGAGA  |
| <i>SIG5</i>                   | AAGCGATGATCCAGCA  | GAACACGAAGGAGCGGAAT   |
| <i>SIG6</i>                   | CTGGAGAGGAGGCAGTTGA   | CCGGCAATTCTGTTAGATG   |
| <i>Actin2</i>                 | CGCCATCCAAGCTGTTCT  | TCACGTCCAGCAAGGTCA  |

in *AtSIG1*, *AtSIG3*, and *AtSIG6*, respectively (Figure 1A). These lines were then used in crosses to generate double mutants and selfed progeny lines were tested by genomic PCR for presence of the T-DNA insertion. As shown in Figure 1B, the amplification products detectable with wildtype DNA (using the primer pairs detailed in Table 1) were absent in either the single or double mutant lines, indicating a loss of sigma gene function in all these lines. This was confirmed by using primer pairs across the junction between T-DNA and sigma gene sequence, which generated amplification products (data not shown).

The single mutant lines *sig1-2* and *sig3-2* each have a visible phenotype resembling that of the wildtype (Figure 1C, panels 2 and 3). In contrast, the two double mutants *sig1 sig6* and *sig6 sig3*, each resulting from crosses with the *sig6* mutant line (panels 4), both reveal yellowish to white cotyledons, with only minimal light-green leaf primordia recognizable at day 10 (panels 5 and 6). Previous work (Ishizaki et al., 2005; Loschelder et al., 2006; Schweer et al., 2006) had shown that mutant alleles of the Arabidopsis sigma gene *AtSIG6* account for reduced chlorophyll content during seedling development. As is evident from Figure 1D, the *sig6-2* line (in comparison with the wildtype) reveals approximately 80% loss of chlorophyll at the 6 d and 10 d stages, respectively, but only 20% loss in 20 day old rosette plants. Chlorophyll quantification of the double mutants likewise shows highly reduced amounts, which (in *sig1 sig6*) are similar to or (in *sig6 sig3*) are even more pronounced than those of the *sig6* parental line at the 6 d and 10 d time-points. At day 20, there is an even further reduction by 15% in *sig1 sig6* and by almost 30% in *sig6 sig3* beyond the *sig6-2* level.

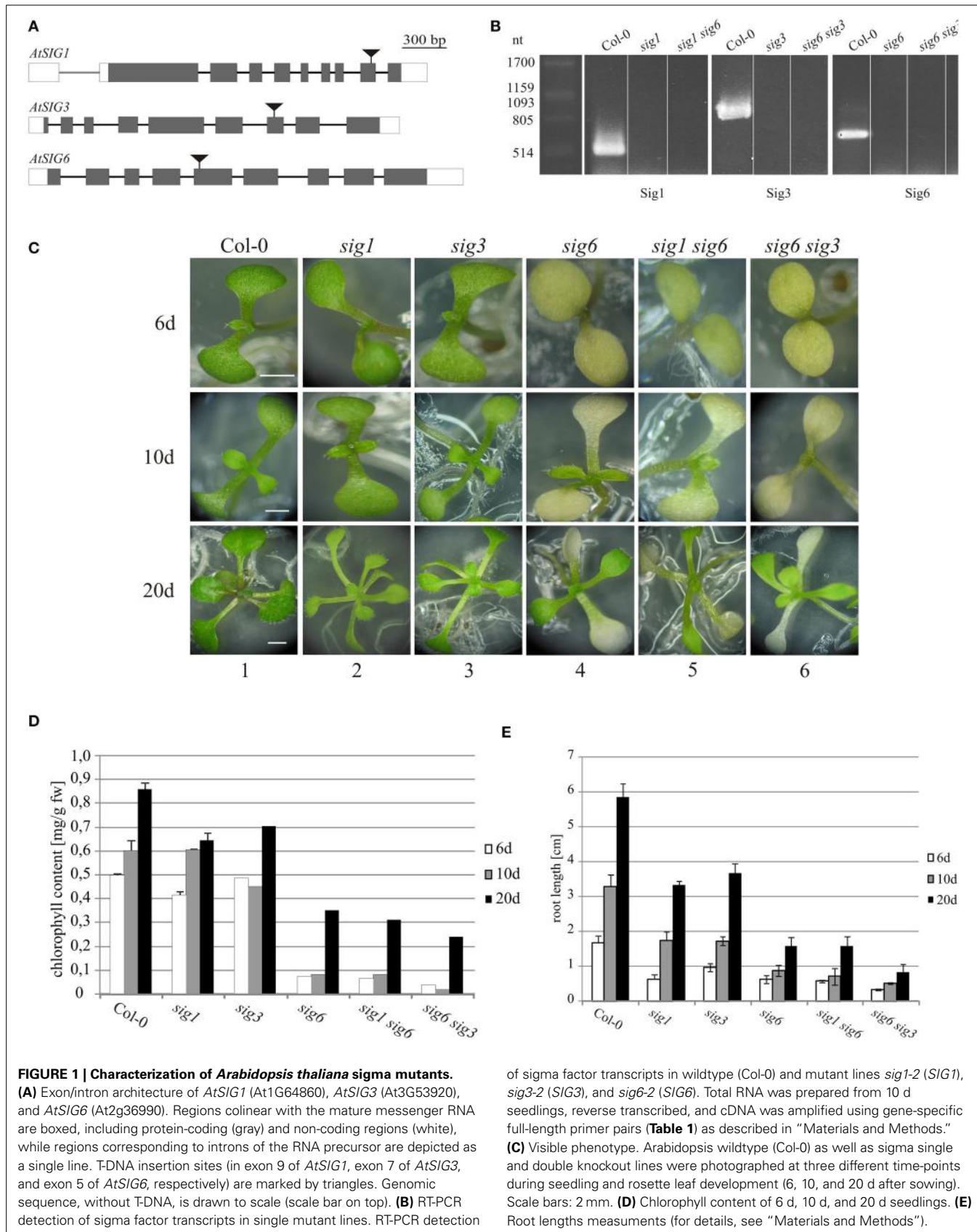
Root length (Figure 1E) is likewise reduced in the single mutants compared to wildtype, with a stronger effect noticeable for *sig6* than for *sig1-2* and *sig3-2*. The double mutant lines again show similar (*sig1 sig6*) or even greater (*sig6 sig3*) length reduction compared to *sig6-2*.

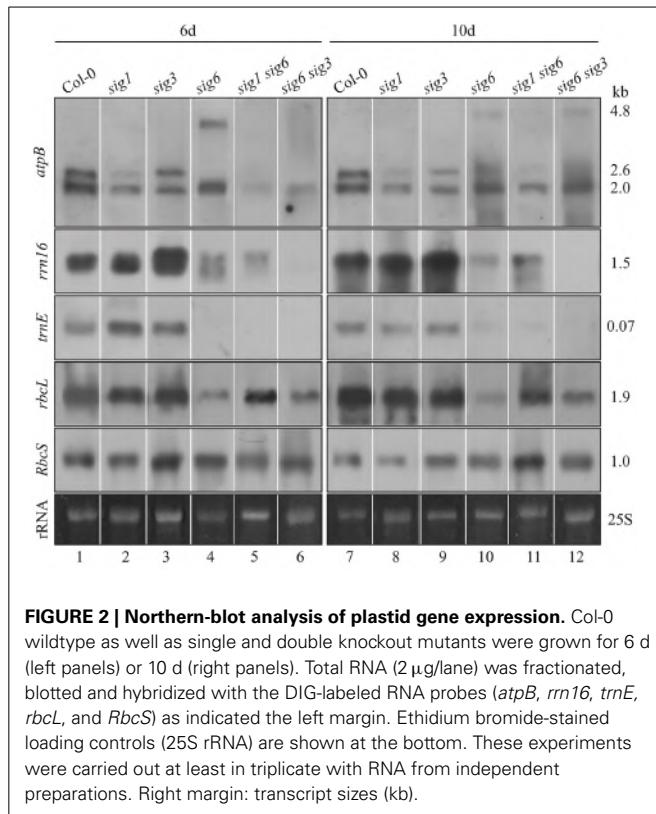
## PLASTID GENE EXPRESSION AT RNA LEVEL IN WILDTYPE, SINGLE AND DOUBLE KNOCKOUT MUTANTS

To assess consequences of sigma gene inactivation on plastid transcript patterns, we carried out northern blot analyses using total RNA from 6- and 10-day old single and double mutant lines. Since many plastid genes give rise to distinct (single or multiple) transcripts of relatively high abundance, this technique was selected to rapidly reveal stage-specific RNA expression patterns of organellar target genes of sigma-dependent transcription. Representatives of all three major classes of plastid genes, i.e., those for proteins (*atpB* and *rbcL*, the genes for the  $\beta$  subunit of the organellar ATP synthase and the large subunit of ribulose-1,5-bisphosphate carboxylase-oxygenase or abbreviated Rubisco, respectively), ribosomal RNAs (*rrn16*, the gene for 16S rRNA) and transfer RNAs (*trnE*, the gene for glutamic acid-specific tRNA), were included in this analysis (for map position and sequence see [www.ncbi.nlm.nih.gov/nuccore/7525012](http://www.ncbi.nlm.nih.gov/nuccore/7525012)).

Using plastid RNA probes *rrn16*, *trnE*, and *rbcL* (Figure 2), no appreciable deviation from the Col-0 wildtype pattern (lane 1 and 7) was noticeable at 6 or 10 days for either the *sig3-2* or *sig1-2* mutant lines (lane 2, 3 and 8, 9), except for a relative increase in intensity of the *rrn16* signal at 6 d in the *sig3-2* line (lane 3). The *atpB* probe (top panel) revealed another difference, i.e., preferential loss of the 2.6 kb (PEP-dependent) transcript (Schweer et al., 2006) in the *sig1-2* line (lane 2 and 8). Unlike both the *sig3-2* and *sig1-2* lines, the *sig6-2* line shows overall less intense hybridization bands compared to wildtype (lane 4 and 10), although to a variable extent depending on the gene investigated. This is most evident for *atpB* (Loschelder et al., 2006), with a loss of the 2.6 kb (PEP-dependent) transcript and concomitant appearance of the 4.8 kb (NEP-dependent) “SOS” transcript (Schweer et al., 2006), particularly at the 6 day stage. Another gene showing dramatic down-regulation of transcript intensity in the *sig6* line is *trnE*, whereas the *rbcL* (Loschelder et al., 2006) and *rrn16* gene expression appears somewhat less affected and that of a nuclear (*RbcS*) control gene remains substantially unaffected (lane 4 and 10).

The double mutant lines *sig1 sig6* and *sig6 sig3* (lane 5, 6 and 11, 12) show a reduction in transcript intensities, which however only partly reflects the patterns of their parental single mutant lines. Following hybridization with the *atpB* probe (top panels), both double mutants show highly reduced signal intensity at the position of the 2.6 kb transcript (lane 2, 5, 6 and 8, 11, 12). In addition, also the 2.0 kb (NEP-dependent) signal is reduced at day 6 but reappears at day 10. The transcript detected with the *rrn16* probe (second panels) is present in reduced amounts in the *sig1 sig6* mutant (lane 5, 11) and is virtually absent in *sig6 sig3* (lane 6 and 12). The *trnE* signal (third panels) is not detectable in both double mutants at either time-point. In the *sig6 sig3* mutant (lane 6 and 12), the intensity of the *rbcL* hybridization signal (forth



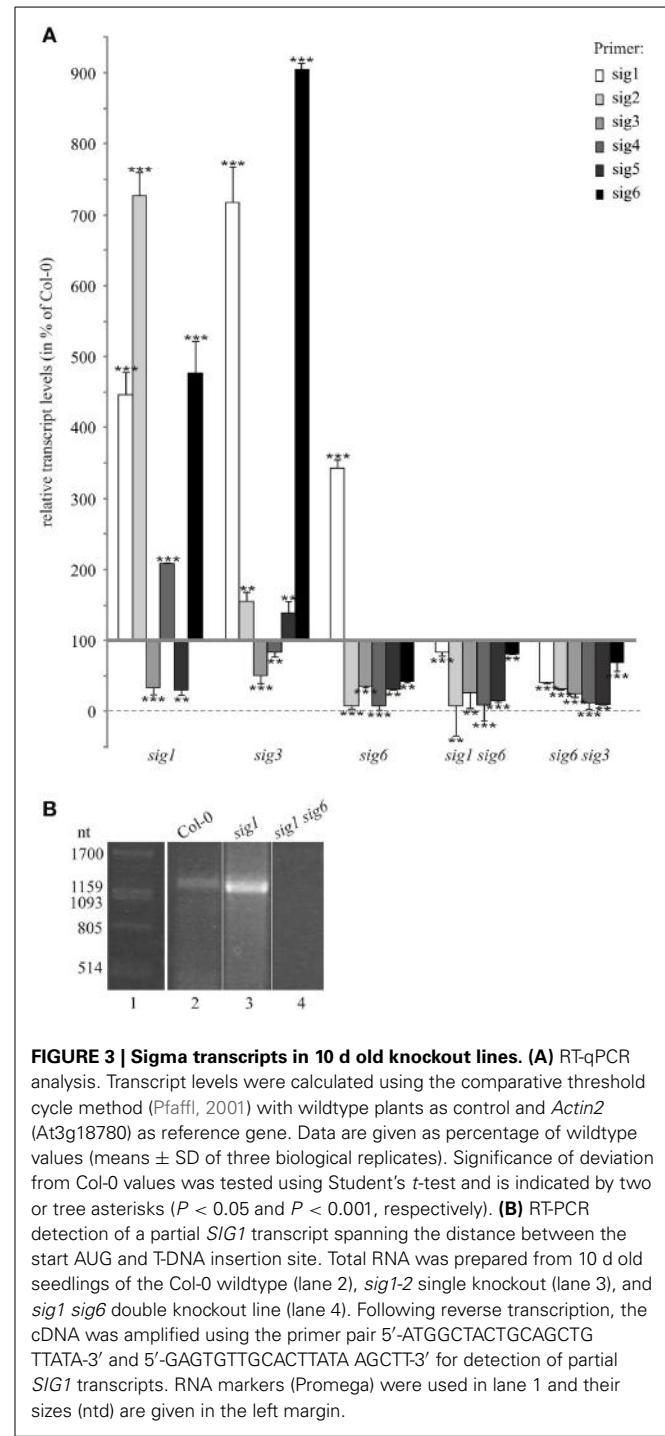


**FIGURE 2 | Northern-blot analysis of plastid gene expression.** Col-0 wildtype as well as single and double knockout mutants were grown for 6 d (left panels) or 10 d (right panels). Total RNA (2 µg/lane) was fractionated, blotted and hybridized with the DIG-labeled RNA probes (*atpB*, *rrn16*, *trnE*, *rbcL*, and *RbcS*) as indicated the left margin. Ethidium bromide-stained loading controls (25S rRNA) are shown at the bottom. These experiments were carried out at least in triplicate with RNA from independent preparations. Right margin: transcript sizes (kb).

panels) resembles that of the parental *sig6*-2 single mutant (lane 4, 10) and is diminished compared to the *sig3*-2 signal (lane 3 and 9). The *sig1 sig6* signal (lane 5 and 11), however, is increased relative to that for *sig6*-2 (lane 4 and 10) at both time points. It is comparable to that for *sig1*-2 (lane 2) at 6 d (lane 5) but shows some relative decrease at 10 d (lane 11 vs. lane 8). Finally, the nuclear (*RbcS*) control reveals more uniform transcript intensity in all tested lines at 6 d (lane 1–6) than at 10 d (lane 7–12). At the latter time-point, the most notable effect is the relative increase in signal intensity for the *sig1 sig6* line (lane 11) compared to the parental mutants and even the wildtype. This may reflect the known plastid to nuclear signaling (Woodson et al., 2012) in response to altered sigma-dependent chloroplast transcription (see “Discussion”).

### SIGMA GENE EXPRESSION NETWORK IN SIGMA SINGLE AND DOUBLE KNOCKOUT LINES

To assess correlations between plastid target gene expression and sigma gene activity, transcript levels for all members of the sigma family were determined. To detect and quantitate these low-abundant transcripts, real-time qPCR rather than northern-blot hybridization was used. Whole-cell RNA preparations from 10 day-old homozygous single and double knockout lines as well as from the Col-0 wildtype were reverse-transcribed and subjected to real-time qPCR. Data were normalized to *Actin2* RNA expression levels (see “Materials and Methods”). As the RNA patterns of plastid target genes (Figure 2) indicated a close similarity of double mutant lines primarily with the *sig6*-2 single mutant line, we tested these lines for correlation of target gene expression



**FIGURE 3 | Sigma transcripts in 10 d old knockout lines.** (A) RT-qPCR analysis. Transcript levels were calculated using the comparative threshold cycle method (Pfaffl, 2001) with wildtype plants as control and *Actin2* (At3g18780) as reference gene. Data are given as percentage of wildtype values (means ± SD of three biological replicates). Significance of deviation from Col-0 values was tested using Student's *t*-test and is indicated by two or tree asterisks ( $P < 0.05$  and  $P < 0.001$ , respectively). (B) RT-PCR detection of a partial *SIG1* transcript spanning the distance between the start AUG and T-DNA insertion site. Total RNA was prepared from 10 d old seedlings of the Col-0 wildtype (lane 2), *sig1*-2 single knockout (lane 3), and *sig1 sig6* double knockout line (lane 4). Following reverse transcription, the cDNA was amplified using the primer pair 5'-ATGGCTACTGCAGCTG TTATA-3' and 5'-GAGTGTGCACTTATA AGCTT-3' for detection of partial *SIG1* transcripts. RNA markers (Promega) were used in lane 1 and their sizes (ntd) are given in the left margin.

with sigma gene expression patterns themselves. As shown in Figure 3A for the *sig6*-2 single mutant line, most sigma transcripts are down-regulated compared to wildtype, while the *SIG1* transcript seems to be strongly enhanced. In both double mutant lines, however, all six sigma transcripts including that of *SIG1* are reduced. Hence, except for *SIG1*, the expression phenotype of the sigma gene family in each double knockout substantially reflects that of the *sig6*-2 parental line.

To test for possible similarity with other parental mutants, we also investigated *sig1-2* and *sig3-2* lines. In terms of their real-time transcript patterns (**Figure 3A**), these two lines each can be clearly set apart from *sig6-2* as well as from the two double mutants. In the *sig3-2* single knockout line, the only strong down-regulation is noticeable for transcripts detected by the *SIG3* primers, which is consistent with a major or full loss of *SIG3*-related transcripts in this mutant. While *SIG4* transcript levels are almost at wild-type level, those of *SIG2* and *SIG5* are moderately elevated and transcript levels of *SIG1* and *SIG6* are even strongly enhanced.

The *sig1-2* knockout line reveals transcript levels that are decreased in the case of *SIG3* and *SIG5*, but are increased for *SIG2*, *SIG4*, and *SIG6*. In addition, despite the T-DNA insertion interrupting the *SIG1* gene (**Figure 1A**) and the concomitant loss of a full-size *SIG1* transcript (**Figure 1B**), the RT-qPCR data for the *sig1-2* line (**Figure 3A**) suggested the existence of significant amounts of *SIG1*-related transcripts. The latter may represent partial (non-functional) transcripts upstream of the T-DNA insertion site. To test this, RT-PCR was carried out with primers that flank the distance from the ATG of the first exon to a site directly in front of the T-DNA insertion (**Figure 3B**). Both the wildtype (lane 2) and the *sig1-2* single mutant (lane 3) showed a PCR signal consistent with a transcript spanning the entire region defined by this primer pair. In contrast, such signal was not detectable in the double mutant line (*sig1 sig6*) (lane 4).

#### CHARACTERIZATION OF RNAi-MODIFIED SIGMA KNOCKOUT LINES

Attempts to further reduce the complexity of the sigma family were initially hampered by our inability to generate sigma triple-mutant lines, possibly because of their lethality. We therefore chose RNAi in combination with the existing single and double sigma mutants as an alternative strategy. “Combined” RNAi/knockout lines were created by transformation of the *sig6-2* and *sig6 sig3* mutants using constructs that contain the complete sequence for the unconserved sigma region (UCR) in pHELLSGATE12 (Wesley et al., 2001). Lines that could be stably maintained included those with the UCR sequences of *SIG2* or *SIG4* in the *sig6-2* single mutant background as well as those with the UCR sequences of *SIG1* or *SIG4* in the *sig6 sig3* double mutant background.

At each time-point (6, 10, and 20 days after sowing) in **Figure 4A**, the combined RNAi/knockout line *sig6::UCRSIG4* (panels 3) has a pigment-deficient visible phenotype which resembles that of the parental *sig6* line (panels 1). The same is true for *sig6::UCRSIG2* (panels 2) as well as for combined lines *sig6 sig3::UCRSIG1* (panels 5) and *sig6 sig3::UCRSIG4* (panels 6), and their parental double mutant line *sig6 sig3* (panels 4) (see also **Figure 1C**). The pigment-deficient visible appearance of all investigated RNAi/knockout lines is reflected by their chlorophyll content (**Figure 4C**). Likewise, root length measurements (**Figure 4D**) establish growth deficiency similar to that noticeable for the parental single (*sig6-2*) and double mutant lines (*sig6 sig3*) (see **Figures 1D,E**).

As shown in **Figure 4B**, transcripts of plastid target genes were found to be even more compromised than those of the (single and double mutant) parental lines (see also **Figure 2**), although differentially and to a variable extent. For instance, except for the loss

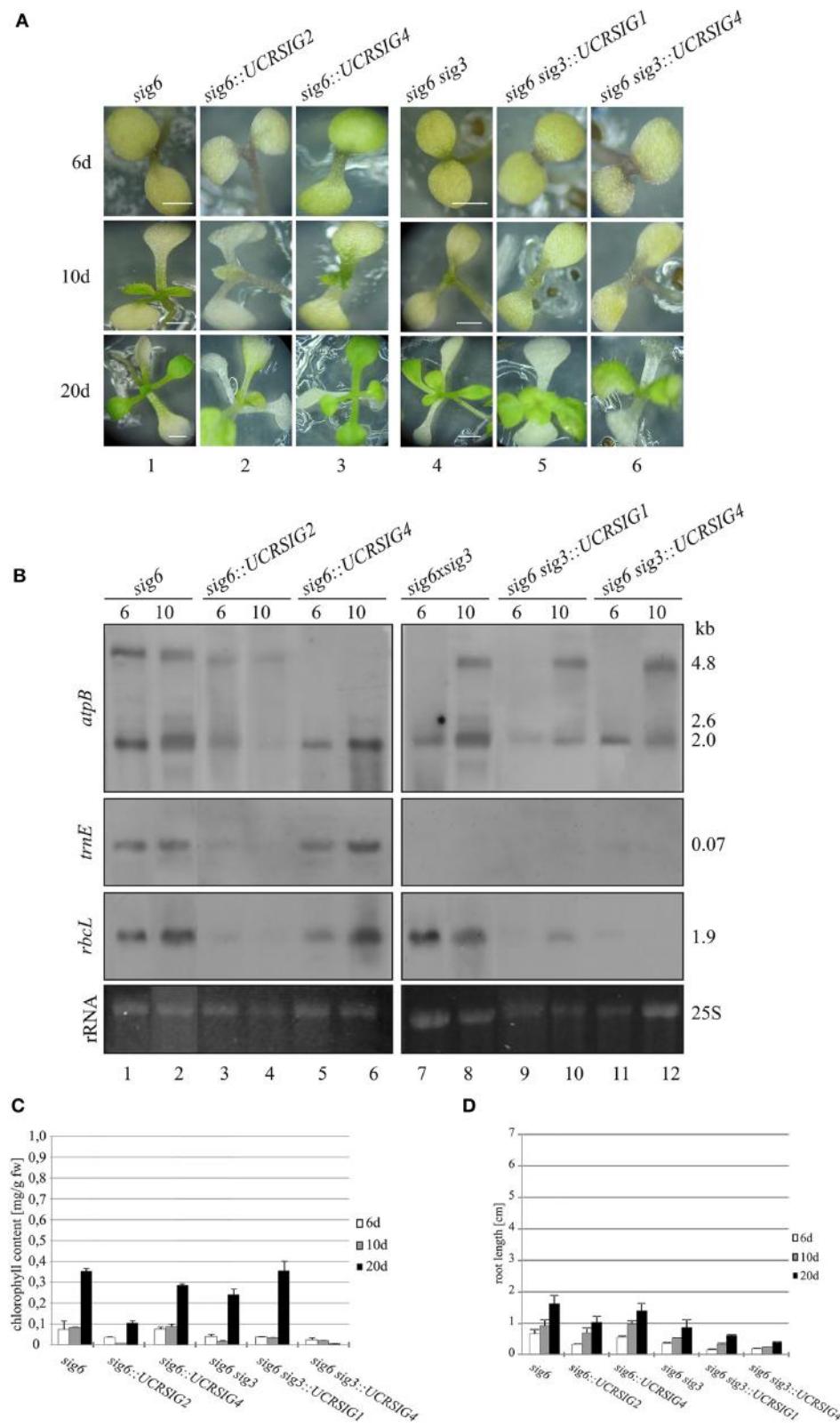
of the 4.8 kb *atpB* transcript, the relative intensity of hybridization signals of *sig6::UCRSIG4* (panels 5 and 6) is comparable to that of *sig6-2* (panels 1 and 2) both at 6 d and 10 d, respectively. In contrast, *sig6::UCRSIG2* (panels 3 and 4) shows a weak but discernible signal (especially at 6 d) at the position of the 4.8 kb *atpB* transcript, while all other signals are virtually absent or highly reduced at both time-points in comparison with *sig6* (panels 1 and 2). This also includes the 2.0 kb band at the position of the NEP-dependent *atpB* transcript (Schweer et al., 2006), whereas the PEP-dependent 2.6 kb *atpB* transcript is absent both in *sig6* itself and in all *sig6*-derived lines. The transcript patterns of *sig6 sig3::UCRSIG1* (lanes 9 and 10) and *sig6 sig3::UCRSIG4* (panels 11 and 12) are similar to, but are more strongly affected than, that of their parental line *sig6 sig3* (panels 7 and 8). While all three lines show the 4.8 kb *atpB* transcript (at day 10 but not day 6), none of them reveals the 2.6 kb (PEP-dependent) *atpB* transcript and the 2.0 kb (NEP-dependent) transcript seems to be diminished in the combined RNAi/double knockout lines (panels 9–12). The *trnE* transcript is absent in all three lines (panels 7–12), and the *rbcL* transcript is highly reduced in the “combined” lines (lanes 9–12) compared to the parental double knockout (panels 7 and 8).

The same RNAi/knockout lines were also tested for their sigma transcript patterns in comparison with those of the *sig6* and *sig6 sig3* parental lines (**Figure 5**). In the case of the *sig6::UCRSIG4* line, the steady-state transcript concentrations of all sigma genes are further reduced compared to those of *sig6* itself. In contrast, *sig6::UCRSIG2* shows a more diverse pattern, with a moderate further reduction of the *SIG2*, *SIG3*, and *SIG6* transcripts compared to the *sig6-2* line, substantially unchanged levels of the *SIG1* and *SIG5* transcripts, and strongly increased concentration of the *SIG4* transcript (**Figure 5A**). The RNAi/double knockout lines *sig6 sig3::UCRSIG1* and *sig6 sig3::UCRSIG4* (**Figure 5B**) both reveal moderate to strong further reduction of the *SIG2–SIG6* transcripts but enhanced levels of the *SIG1* transcript compared to the parental *sig6 sig3* line. Their transcript patterns thus seem more similar to one another than to those shown in **Figure 5A**.

#### DISCUSSION

This work was carried out to help reach a fuller understanding of the plant sigma genes, i.e., the small nuclear gene family for chloroplast transcription factors resembling the bacterial sigma transcription initiation factors. We sought to analyze phenotypic and molecular consequences of altered patterns of sigma gene activity in *Arabidopsis thaliana*. Such changes can be “homeostatic,” i.e., balanced and compensatory, which can be anticipated during developmental and/or physiological transitions. In more extreme situations, however, the sigma network can be thought to become “disrupted,” i.e., (irreversibly) imbalanced and rendered non-functional, as might be expected in lethal or heavily compromised mutants and/or under strong environmental stress. To narrow down the limits of homeostatic vs. imbalanced states of the sigma gene family, we have chosen strategies to differentially affect the sigma network, including single and double sigma mutants in combination with RNAi.

The parental single mutants, *sig1-2*, *sig3-2*, and *sig6-2*, were chosen for reasons of their specific phenotypic characteristics. *SIG6* single mutant lines reveal a clear-cut chlorophyll-deficient



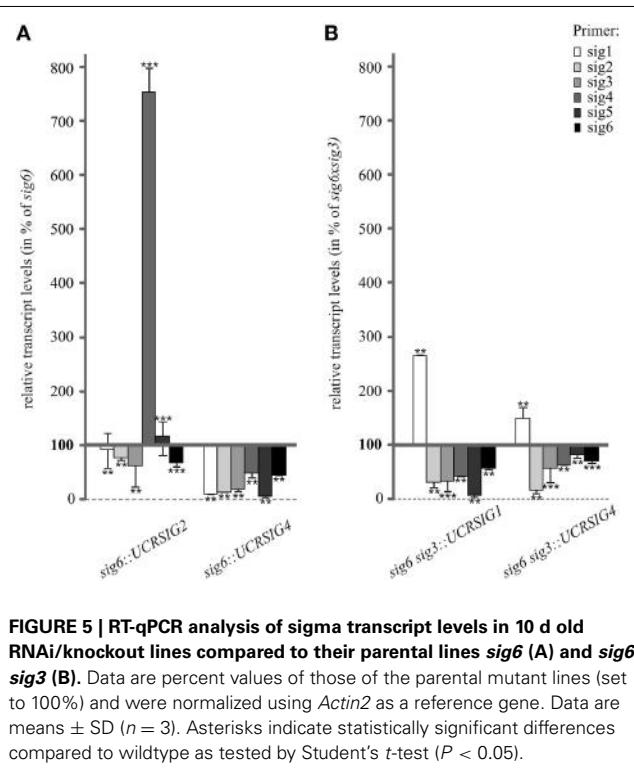
**FIGURE 4 | “Combined” RNAi/knockout lines.** The single knockout line *sig6* and the double knockout line *sig6 sig3* each were transformed with the pHELLSGATE12 vector carrying *SIG1*, *SIG2*, or *SIG4* cDNA segments that

represent the full-size UCR of sigma factors SIG1, SIG2, or SIG4, respectively (**Table 1**). **(A)** Phenotype of the RNAi/knockout lines and their parental lines (Continued)

**FIGURE 4 | Continued**

during development (6, 10, and 20 d after sowing). Scale bars: 2 mm. **(B)** Northern blot analysis of plastid gene expression in RNAi/knockout lines and their parental lines. Total RNA (2 µg/lane) from 6 and 10 d old seedlings was fractionated, blotted and hybridized with *atpB*, *trnE*, or *rbcL* probes as indicated in the left margin. Ethidium bromide-stained loading controls (25S rRNA) are shown at the bottom. These experiments were carried out at least in triplicate with RNA from independent preparations. Right margin: transcript sizes (kb). **(C)** Chlorophyll content of

seedlings and young plantlets 6 d, 10 d, or 20 d after sowing. Measurements involved 8 replicates of 10 samples each from three series of independently grown plant material. Weighed samples were ground in 80% (v/v) acetone and photometric chlorophyll determination at 663 and 645 nm was carried out. **(D)** Root lengths. Following growth of seedlings and young plantlets on vertically positioned Petri dishes, length measurements were carried out using ImageJ (<http://imagej.nih.gov/ij>). Values were means of three independent replicates using samples representing 20 seedlings or plantlets each.



**FIGURE 5 | RT-qPCR analysis of sigma transcript levels in 10 d old RNAi/knockout lines compared to their parental lines *sig6* (A) and *sig6 sig3* (B).** Data are percent values of those of the parental mutant lines (set to 100%) and were normalized using *Actin2* as a reference gene. Data are means  $\pm$  SD ( $n = 3$ ). Asterisks indicate statistically significant differences compared to wildtype as tested by Student's *t*-test ( $P < 0.05$ ).

and developmental-stage-specific (seedling) phenotype, indicating a specialized and/or functionally dominating role of this factor (Ishizaki et al., 2005; Loschelder et al., 2006; Schweer et al., 2006, 2009). In contrast, *sig3* lines do not show pronounced pigment deficiency and the corresponding factor SIG3 has been assigned a functionally redundant role, perhaps as a possible safeguard in case of loss of other sigma factor(s) (Schweer, 2010; Lerbs-Mache, 2011). Although *sig1* knockout mutants of *Arabidopsis* have not yet been presented in terms of their gene expression characteristics, such mutants are of considerable interest in view of recent findings that sigma factor 1 (SIG1) is subject to phosphorylation control (Shimizu et al., 2010), as is known for sigma factor 6 (Schweer et al., 2010a,b; Türkeri et al., 2012).

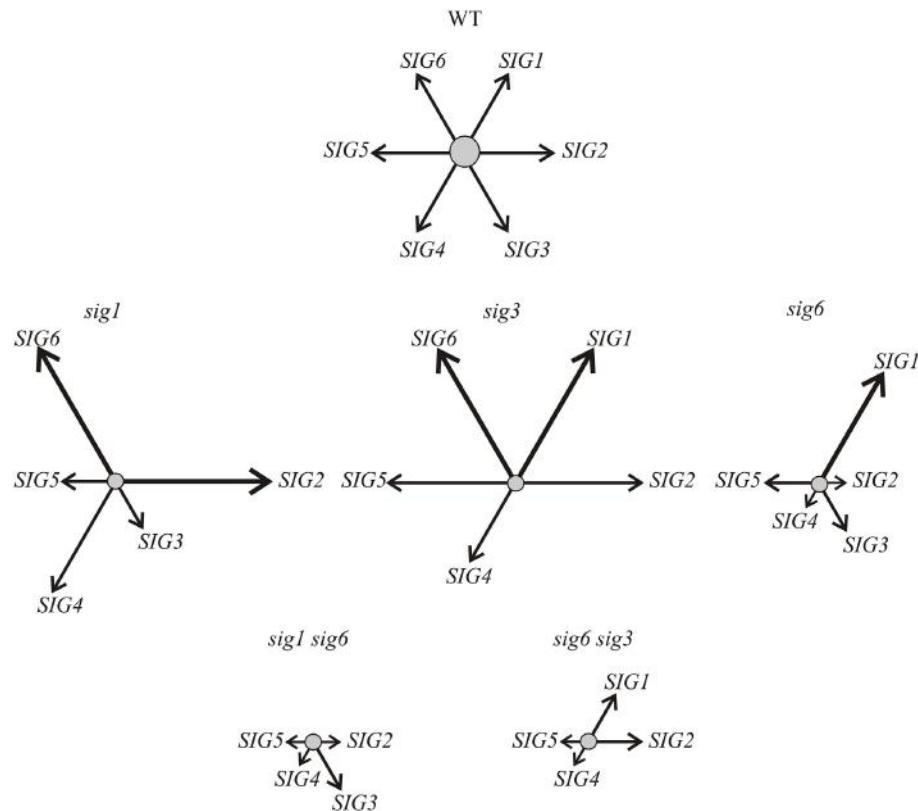
To reduce the number of functional sigma genes, the *sig6*-2 knockout line (Loschelder et al., 2006; Schweer et al., 2006, 2009) was crossed with either of two other single mutant lines, *sig1*-2 and *sig3*-2, giving rise to the double mutants *sig1 sig6* and *sig6 sig3*. The latter reveal a growth-retarded and highly chlorophyll-deficient phenotype at seedling stage, even exceeding that of the parental *sig6*-2 knockout. Plastid target gene

expression at RNA level was strongly compromised in both double mutants. Assessment of sigma gene expression itself using RT-qPCR showed both losses but also increases in transcript frequency for individual members of the gene family in the single mutants. In contrast, however, a global decrease of sigma transcripts was noticeable in the double mutants. Hence, due to functional redundancy and compensation of sigma family members, a homeostatic balance seems still to prevail in the single knockout lines, while the balance may be strongly shifted or completely lost in the double mutants.

Attempts to further reduce the complexity of the sigma family by construction of triple mutants were unsuccessful, likely because of lethality of the progeny from these crosses. As an alternative, we therefore combined knockout mutant with RNAi technology. Using various sigma-specific unconserved regions (UCRs) in the pHELLSGATE12 RNAi vector, single (*sig6*) and double (*sig6 sig3*) mutant lines were transformed by constructs based on this vector. Resulting progeny lines generated by transformation of *sig6* showed a chlorophyll-deficient phenotype similar to or even stronger than that of the parental knockout line. Those generated from *sig6 sig3* did not reveal a further enhanced phenotype compared to the parental double knockout line, suggesting that a “basal” (minimal) state which cannot be further reduced without loss of viability may have been reached already in the latter.

An argument against this notion, however, comes from results of the target gene expression studies, showing that the *rbcL* transcript is readily detectable in *sig6 sig3* but is highly reduced or absent in the “combined” RNAi/double knockout lines (Figure 4B). Furthermore, the RT-qPCR analysis (Figure 5B) shows enhanced transcript levels of the *SIG1* transcript in *sig6 sig3::UCRSIG1* and *sig6 sig3::UCRSIG4* as compared to the parental line *sig6 sig3*, which may reflect a still balanced functional state of the sigma gene family in these lines. In any case, it is notable that down-regulation of the expression of a single sigma gene can both negatively or positively affect that of another family member. For instance, the expression of *SIG1* and *SIG6* seems to be regulated in an opposite manner in single knockout lines (Figure 6), indicating functional redundancy and mutual compensation.

A perhaps unexpected finding is the partial or even dramatic loss of the 2.0 kb (NEP-dependent) *atpB* transcript (Schweer et al., 2006) in all RNAi/knockout lines (Figure 4B). Sigma factor SIG6 was previously implicated in retrograde signaling from the chloroplast to nucleus (Woodson et al., 2012), i.e., a mechanism that can affect the expression of nuclear genes in response to altered sigma factor function and chloroplast transcription



**FIGURE 6 | Scheme symbolizing the transcript pattern of the sigma family in *Arabidopsis* mutant lines as revealed by RT-qPCR (Figure 3A).** For simplicity, the balanced network of sigma factor transcripts (*SIG1*–*SIG6*) in the wildtype is indicated by equal length and strength of the arrows arranged in a hexagonal pattern (top). As indicated by the altered length and strength of arrows, the contribution of individual sigma transcripts varies considerably in the single (middle portion) and double mutant lines (bottom). Compensation of a completely lacking and/or non-functional transcript by overexpression of different member(s)

(Pfannschmidt, 2010). It can be envisaged that the nuclear-encoded plastid polymerase might be regulated via this route, which in turn could explain the loss of the 2.0 kb NEP-dependent *atpB* transcript (Hanaoka et al., 2005).

Concerted regulation of both the PEP and NEP transcription systems via functional alterations of one or several sigma factors can be considered as efficient and flexible mechanism to achieve interorganellar integration. For instance, it might be interesting to investigate if the increased *RbcS* transcript level seen in the *sig1 sig6* double mutant (Figure 2) is primarily due to an altered sigma network in this mutant and/or involves NEP-dependent regulation. Clearly, differential and compensatory expression of sigma genes as studied here is only one of several control levels. Posttranslational modification such as phosphorylation (Schweer et al., 2010a,b; Shimizu et al., 2010) as well as interactions of sigma factors with other regulatory proteins are equally important (Morikawa et al., 2002; Chi et al., 2010), as is the topology of the plastid transcriptome (Yagi et al., 2012; Zhelyazkova et al., 2012). In any case, our current work points to a causal relationship between the expression status of

the gene family is noticeable in the single mutants (shifted, but probably still balanced, network). In the double mutants, this “valve” may no longer be available, as is suggested by the down-regulation of all remaining transcripts (imbalanced scrambled network). It should be noted, however, that the majority of the “combined” RNAi/knockout lines (Figure 5) still reveal increased (compensatory?) transcript levels of a single sigma gene (not shown here), suggesting that a (still viable) “basal state” may not have been reached yet in the double knockout mutants (bottom row).

the sigma gene family and responses at the level of target gene expression.

Finally, it should be recalled that all RNAi lines described here were generated using the “constitutive” pHELLSGATE12 silencing vector (Wesley et al., 2001). A somewhat similar picture also emerges from initial recent work with chemically inducible dexamethasone (DEX)-responsive RNAi lines based on the pOpOff2 vector (Wielopolska et al., 2005), providing proof of principle for DEX-responses that are visible at both target gene and sigma expression patterns (data not shown). Usage of such inducible “knock-down” system, also including, e.g., virus-induced gene silencing (VIGS) (Ratcliff et al., 1997), can be expected to open up new avenues in studies on temporal and spatial activities of the gene-containing plant cell organelles. This way, it should become possible to successfully analyze stages throughout the entire *Arabidopsis* life cycle.

## AUTHOR CONTRIBUTIONS

Sylvia Bock carried out planning, performance and presentation of most experiments. Jennifer Ortelt assisted in all aspects of

experimental analyses and manuscript preparation. Gerhard Link provided advice and assistance throughout this work.

## ACKNOWLEDGMENTS

We are indebted to Prof. Bernd Weisshaar, University of Bielefeld, and the GABI-Kat team at the Max-Planck-Institute fuer Zuechtungsforschung, Cologne, for the supply of the sigma factor mutant lines, to Prof. Peter Michael Waterhouse and colleagues at the CSIRO, Canberra, for the RNAi vectors, and to Prof. Minou Nowrouzian for guidance on RT-qPCR. We gratefully acknowledge the excellent technical assistance of Brigitte Link. This work was funded by the Deutsche Forschungsgemeinschaft (LI261/21-1).

## REFERENCES

- Burgess, R. R., and Anthony, L. (2001). How sigma docks RNA polymerase and what sigma does. *Curr. Opin. Microbiol.* 4, 126–131. doi: 10.1016/S1369-5274(00)00177-6
- Chi, W., Mao, J., Li, Q., Ji, D., Zou, M., Lu, C., et al. (2010). Interaction of the pentatricopeptide-repeat protein DELAYED GREENING 1 with sigma factor SIG6 in the regulation of chloroplast gene expression in *Arabidopsis* cotyledons. *Plant J.* 64, 14–25. doi: 10.1111/j.1365-313X.2010.04304.x
- Chomczynski, P., and Sacchi, N. (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* 162, 156–159. doi: 10.1016/0003-2697(87)90021-2
- Clough, S. J., and Bent, A. F. (1998). Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* 16, 735–744. doi: 10.1046/j.1365-313X.1998.00343.x
- Favery, J. J., Kobayashi, M., Tanaka, K., Peltier, G., Kreis, M., Valay, J. G., et al. (2005). Specific function of a plastid sigma factor for *ndhF* gene transcription. *Nucleic Acids Res.* 33, 5991–5999. doi: 10.1093/nar/gki908
- Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E., and Mello, C. C. (1998). Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 391, 806–811. doi: 10.1038/35888
- Fujiwara, M., Nagashima, A., Kanamaru, K., Tanaka, K., and Takahashi, H. (2000). Three new nuclear genes, *sigD*, *sigE* and *sigF*, encoding putative plastid RNA polymerase sigma factors in *Arabidopsis thaliana*. *FEBS Lett.* 481, 47–52. doi: 10.1016/S0014-5793(00)01965-7
- Hanaoka, M., Kanamaru, K., Fujiwara, M., Takahashi, H., and Tanaka, K. (2005). Glutamyl-tRNA mediates a switch in RNA polymerase use during chloroplast biogenesis. *EMBO Rep.* 6, 545–550. doi: 10.1038/sj.embo.7400411
- Hanaoka, M., Kanamaru, K., Takahashi, H., and Tanaka, K. (2003). Molecular genetic analysis of chloroplast gene promoters dependent on SIG2, a nucleus-encoded sigma factor for the plastid-encoded RNA polymerase, in *Arabidopsis thaliana*. *Nucleic Acids Res.* 31, 7090–7098. doi: 10.1093/nar/gkg935
- Hedtke, B., Börner, T., and Weihe, A. (1997). Mitochondrial and chloroplast phage-type RNA polymerases in *Arabidopsis*. *Science* 277, 809–811. doi: 10.1126/science.277.5327.809
- Ishihama, A. (1988). Promoter selectivity of prokaryotic RNA polymerases. *Trends Genet.* 4, 282–286. doi: 10.1016/0168-9525(88)90170-9
- Ishizaki, Y., Tsunoyama, Y., Hatano, K., Ando, K., Kato, K., Shinmyo, A., et al. (2005). A nuclear-encoded sigma factor, *Arabidopsis SIG6*, recognizes sigma-70 type chloroplast promoters and regulates early chloroplast development in cotyledons. *Plant J.* 42, 133–144. doi: 10.1111/j.1365-313X.2005.02362.x
- Isono, K., Shimizu, M., Yoshimoto, K., Niwa, Y., Satoh, K., Yokota, A., et al. (1997). Leaf-specifically expressed genes for polypeptides destined for chloroplasts with domains of sigma 70 factors of bacterial RNA polymerases in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U.S.A.* 94, 14948–14953. doi: 10.1073/pnas.94.26.14948
- Lerbs-Mache, S. (2011). Function of plastid sigma factors in higher plants: regulation of gene expression or just preservation of constitutive transcription? *Plant Mol. Biol.* 76, 235–249. doi: 10.1007/s11103-010-9714-4
- Loschelder, H., Schweer, J., Link, B., and Link, G. (2006). Dual temporal role of plastid sigma factor 6 in *arabidopsis* development. *Plant Physiol.* 142, 642–650. doi: 10.1104/pp.106.085878
- Maliga, P. (1998). Two plastid RNA polymerases of higher plants: an evolving story. *Trends Plant Sci.* 3, 4–6. doi: 10.1016/S1360-1385(97)87981-2
- Morikawa, K., Shiina, T., Murakami, S., and Toyoshima, Y. (2002). Novel nuclear-encoded proteins interacting with a plastid sigma factor, Sig1, in *Arabidopsis thaliana*. *FEBS Lett.* 514, 300–304. doi: 10.1016/S0014-5793(02)02388-8
- Murashige, T., and Skoog, F. (1962). A revised medium for rapid growth and bioassays with tobacco tissue cultures. *Plant Physiol.* 15, 473–497. doi: 10.1111/j.1399-3054.1962.tb08052.x
- Nagashima, A., Hanaoka, M., Motohashi, R., Seki, M., Shinozaki, K., Kanamaru, K., et al. (2004). DNA microarray analysis of plastid gene expression in an *Arabidopsis* mutant deficient in a plastid transcription factor, SIG2. *Biosci. Biotechnol. Biochem.* 68, 694–704. doi: 10.1271/bbb.68.694
- Ortelt, J., and Link, G. (2014). Plastid gene transcription: promoters and RNA polymerases. *Methods Mol. Biol.* 1132, 47–72. doi: 10.1007/978-1-62703-995-6\_3
- Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* 29:e45. doi: 10.1093/nar/29.9.e45
- Pfannschmidt, T. (2010). Plastidial retrograde signalling—a true plastid factor or just metabolite signatures? *Trends Plant Sci.* 15, 427–435. doi: 10.1016/j.tplants.2010.05.009
- Privat, I., Hakimi, M. A., Buhot, L., Favory, J. J., and Lerbs-Mache, S. (2003). Characterization of *Arabidopsis* plastid sigma-like transcription factors SIG1, SIG2 and SIG3. *Plant Mol. Biol.* 51, 385–399. doi: 10.1023/A:1022095017355
- Ratcliff, F., Harrison, B., and Baulcombe, D. (1997). A similarity between viral defense and gene silencing in plants. *Science* 276, 1558–1560. doi: 10.1126/science.276.5318.1558
- Rosso, M. G., Li, Y., Strizhov, N., Reiss, B., Dekker, K., and Weisshaar, B. (2003). An *Arabidopsis thaliana* T-DNA mutagenized population (GABI-Kat) for flanking sequence tag-based reverse genetics. *Plant Mol. Biol.* 53, 247–259. doi: 10.1023/B:PLAN.0000009297.37235.4a
- Schweer, J. (2010). Plant sigma factors come of age: flexible transcription factor network for regulated plastid gene expression. *Endocytobiosis Cell Res.* 20, 1–20. Available online at: [http://zs.thulb.uni-jena.de/receive/jportal\\_jparticle\\_00165740](http://zs.thulb.uni-jena.de/receive/jportal_jparticle_00165740)
- Schweer, J., Geimer, S., Meurer, J., and Link, G. (2009). *Arabidopsis* mutants carrying chimeric sigma factor genes reveal regulatory determinants for plastid gene expression. *Plant Cell Physiol.* 50, 1382–1386. doi: 10.1093/pcp/pcp069
- Schweer, J., Loschelder, H., and Link, G. (2006). A promoter switch that can rescue a plant sigma factor mutant. *FEBS Lett.* 580, 6617–6622. doi: 10.1016/j.febslet.2006.11.010
- Schweer, J., Türkeri, H., Link, B., and Link, G. (2010a). AtSIG6, a plastid sigma factor from *Arabidopsis*, reveals functional impact of cpCK2 phosphorylation. *Plant J.* 62, 192–202. doi: 10.1111/j.1365-313X.2010.04138.x
- Schweer, J., Türkeri, H., Kolpack, A., and Link, G. (2010b). Role and regulation of plastid sigma factors and their functional interactors during chloroplast transcription – recent lessons from *Arabidopsis thaliana*. *Eur. J. Cell Biol.* 89, 940–946. doi: 10.1016/j.ejcb.2010.06.016
- Shiina, T., Ishizaki, Y., Yagi, Y., and Nakahira, Y. (2009). Function and evolution of plastid sigma factors. *Plant Biotechnol.* 26, 57–66. doi: 10.5511/plantbiotechnology.26.57
- 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. doi: 10.1016/S0074-7696(05)44001-2
- Shimizu, M., Kato, H., Ogawa, T., Kurachi, A., Nakagawa, Y., and Kobayashi, H. (2010). Sigma factor phosphorylation in the photosynthetic control of photosystem stoichiometry. *Proc. Natl. Acad. Sci. U.S.A.* 107, 10760–10764. doi: 10.1073/pnas.0911692107
- Sugiura, M. (1992). The chloroplast genome. *Plant Mol. Biol.* 19, 149–168. doi: 10.1007/BF00015612
- Tanaka, K., Tozawa, Y., Mochizuki, N., Shinozaki, K., Nagatani, A., Wakasa, K., et al. (1997). Characterization of three cDNA species encoding plastid RNA polymerase sigma factors in *Arabidopsis thaliana*: evidence for the sigma factor heterogeneity in higher plant plastids. *FEBS Lett.* 413, 309–313. doi: 10.1016/S0014-5793(97)00906-X
- 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. doi: 10.1016/S0014-5793(02)02538-3
- Türkeri, H., Schweer, J., and Link, G. (2012). Phylogenetic and functional features of the plastid transcription kinase cpCK2 from *Arabidopsis* signify a role

- of cysteinyl SH-groups in regulatory phosphorylation of plastid sigma factors. *FEBS J.* 279, 395–409. doi: 10.1111/j.1742-4658.2011.08433.x
- Wesley, S. V., Helliwell, C. A., Smith, N. A., Wang, M., Rouse, D. T., Liu, Q., et al. (2001). Construct design for efficient, effective and high-throughput gene silencing in plants. *Plant J.* 27, 581–590. doi: 10.1046/j.1365-313X.2001.01105.x
- Wielopolska, A., Townley, H., Moore, I., Waterhouse, P., and Helliwell, C. (2005). A high-throughput inducible RNAi vector for plants. *Plant Biotechnol. J.* 3, 583–590. doi: 10.1111/j.1467-7652.2005.00149.x
- Woodson, J. D., Perez-Ruiz, J. M., Schmitz, R. J., Ecker, J. R., and Chory, J. (2012). Sigma factor-mediated plastid retrograde signals control nuclear gene expression. *Plant J.* 73, 1–13. doi: 10.1111/tpj.12011
- Yagi, Y., Ishizaki, Y., Nakahira, Y., Tozawa, Y., and Shiina, T. (2012). Eukaryotic-type plastid nucleoid protein pTAC3 is essential for transcription by the bacterial-type plastid RNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.* 109, 7541–7546. doi: 10.1073/pnas.1119403109
- 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. doi: 10.1093/nar/gkl1067
- Zhelyazkova, P., Sharma, C. M., Förstner, K. U., Liere, K., Vogel, J., and Börner, T. (2012). The Primary transcriptome of barley chloroplasts: numerous noncoding RNAs and the dominating role of the plastid-encoded RNA polymerase. *Plant Cell* 24, 123–136. doi: 10.1105/tpc.111.089441

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Received: 24 July 2014; accepted: 09 November 2014; published online: 25 November 2014.*

*Citation: Bock S, Ortelt J and Link G (2014) AtSIG6 and other members of the sigma gene family jointly but differentially determine plastid target gene expression in *Arabidopsis thaliana*. *Front. Plant Sci.* 5:667. doi: 10.3389/fpls.2014.00667*

*This article was submitted to Plant Physiology, a section of the journal Frontiers in Plant Science.*

*Copyright © 2014 Bock, Ortelt and Link. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*



# Plastid encoded RNA polymerase activity and expression of photosynthesis genes required for embryo and seed development in *Arabidopsis*

Dmitry Kremnev and Åsa Strand\*

Umeå Plant Science Centre, Department of Plant Physiology, Umeå University, Umeå, Sweden

**Edited by:**

Thomas Pfannschmidt, University Joseph Fourier Grenoble, France

**Reviewed by:**

Tatjana Kleine,  
Ludwig-Maximilians-Universität München, Germany  
Silva Mache, Laboratoire de Physiologie Cellulaire Végétale, France

**\*Correspondence:**

Åsa Strand, Umeå Plant Science Centre, Department of Plant Physiology, Umeå University, S-901 87 Umeå, Sweden  
e-mail: asa.strand@umu.se

Chloroplast biogenesis and function is essential for proper plant embryo and seed development but the molecular mechanisms underlying the role of plastids during embryogenesis are poorly understood. Expression of plastid encoded genes is dependent on two different transcription machineries; a plastid-encoded bacterial-type RNA polymerase (PEP) and a nuclear-encoded phage-type RNA polymerase (NEP), which recognize distinct types of promoters. However, the division of labor between PEP and NEP during plastid development and in mature chloroplasts is unclear. We show here that PLASTID REDOX INSENSITIVE 2 (PRIN2) and CHLOROPLAST STEM-LOOP BINDING PROTEIN 41 kDa (CSP41b), two proteins identified in plastid nucleoid preparations, are essential for proper plant embryo development. Using Co-IP assays and native PAGE we have shown a direct physical interaction between PRIN2 and CSP41b. Moreover, PRIN2 and CSP41b form a distinct protein complex *in vitro* that binds DNA. The *prin2.2* and *csp41b-2* single mutants displayed pale phenotypes, abnormal chloroplasts with reduced transcript levels of photosynthesis genes and defects in embryo development. The respective *csp41b-2prin2.2* homo/heterozygote double mutants produced abnormal white colored ovules and shrunken seeds. Thus, the *csp41b-2prin2.2* double mutant is embryo lethal. *In silico* analysis of available array data showed that a large number of genes traditionally classified as PEP dependent genes are transcribed during early embryo development from the pre-globular stage to the mature-green-stage. Taken together, our results suggest that PEP activity and consequently the switch from NEP to PEP activity, is essential during embryo development and that the PRIN2-CSP41b DNA binding protein complex possibly is important for full PEP activity during this process.

**Keywords:** chloroplast, PEP, NEP, embryo development, photosynthesis, PRIN2, CSP41b

## INTRODUCTION

The chloroplasts house the photosynthetic light reactions where sunlight is converted into chemical energy. Plastids are also the location of a number of vital metabolic pathways, including primary carbon metabolism and the biosynthesis of fatty acids, amino acids, and tetrapyrroles. Chloroplast function is required throughout the life cycle of the plant and compromised activity can result in embryo lethality. Abortion of developing embryos is known to occur when amino acid, nucleotide or fatty acid biosynthesis is impaired, or when import of chloroplast proteins and translation are disrupted (McElver et al., 2001; Tzafrir et al., 2003; Hsu et al., 2010; Bryant et al., 2011). In contrast, disrupting components of the photosynthetic apparatus leads to reduced pigmentation and changed physiology rather than embryo lethality (Tzafrir et al., 2003; Bryant et al., 2011). Chloroplasts are detected as early as at the globular stage of the embryo (Tejos et al., 2010) and transcript profiling during embryo development showed a significant increase in the expression of nuclear genes encoding components involved in energy production, carbon fixation and photosynthesis already from the globular embryogenic stage (Spencer et al., 2007; Le et al.,

2010; Belmonte et al., 2013). Furthermore, it has been indicated that chloroplasts can provide the embryo with energy and O<sub>2</sub> required for biosynthesis and respiration (Rolletschek et al., 2003). Additionally, it was shown that *Brassica* embryos were able to fix CO<sub>2</sub> and contributing to embryo growth rate and biomass (Goffman et al., 2005).

Chloroplasts, like mitochondria, evolved from free-living prokaryotic organisms that entered the eukaryotic cell through endosymbiosis. The gradual conversion from endosymbiont to organelle during the course of evolution has clearly been accompanied by a dramatic reduction in genome size as the chloroplasts lost most of their genes to the nucleus. The genes remaining in the chloroplast genome are related to photosynthesis or encode components of the plastid gene expression machinery (Wakasugi et al., 2001; Martin et al., 2002). The chloroplast genes of higher plants are transcribed by at least two types of RNA polymerases; the nuclear encoded plastid RNA polymerase (NEP), a T3-T7 bacteriophage type that predominantly mediates the transcription of the house keeping genes (Hedtke et al., 1997, 2000; Puthiyaveetil et al., 2010). The other type, plastid encoded RNA polymerase (PEP) is a bacterial-type

multi-subunit enzyme that predominantly mediates the transcription of photosynthesis-related genes (Allison et al., 1996; De Santis-Maclossek et al., 1999). Most chloroplast genes can be transcribed by both polymerases but they utilize different promoter elements (Hajdukiewicz et al., 1997; Pfannschmidt and Liere, 2005). The PEP enzyme recognizes the  $-10$  and  $-35$  *cis*-elements, similar to those found in bacterial promoters whereas the NEP enzyme recognizes the YRTA-motif, which can also be found upstream of several genes with PEP promoters indicating that these genes can be transcribed by both polymerases (Pfannschmidt and Liere, 2005). Transcription of the plastid encoded photosynthesis genes during chloroplast development and the activation of the photosynthetic reactions are accompanied by a switch from NEP to PEP activity (Hanaoka et al., 2005). However, the mechanisms underlying this change in major RNA polymerase activity and the division of labor between NEP and PEP in the chloroplast are unknown (Zhelyazkova et al., 2012).

A large number of proteins have been shown to be associated with the PEP complex and the components associated with PEP changes in response to developmental signals and changes in the environment (Pfalz and Pfannschmidt, 2013). The variation in the composition of the PEP complex suggests that regulation of plastid gene expression is both complex and sophisticated. PLASTID REDOX INSENSITIVE 2 (PRIN2) is a novel plant protein localized to the plastid nucleoids and plastid transcriptome analyses demonstrated that PRIN2 is required for full expression of genes transcribed by PEP (Kindgren et al., 2012). The role of the PEP associated proteins is unclear and whether these proteins contribute to the regulation of plastid gene expression by environmental and developmental cues remains to be determined. In order to shed light on the complex regulation of PEP activity and to understand the function of PRIN2 we pursued an assay to identify interacting protein partners of PRIN2. We report here a direct interaction between PRIN2 and CHLOROPLAST STEM-LOOP BINDING PROTEIN 41 kDa (CSP41b) using three different biochemical methods. CSP41b was described as an RNA binding protein identified in nucleoid preparation and was attributed numerous plastid functions, for example CSP41b was suggested to stimulate both transcription and translation in the chloroplast (Pfannschmidt et al., 2000; Yamaguchi et al., 2003; Suzuki et al., 2004; Hassidim et al., 2007; Bollenbach et al., 2009). We also demonstrate that PRIN2 and CSP41b form a distinct DNA binding protein complex *in vitro* and that the *csp41b-2prin2.2* double mutant is embryo lethal. Taken together, our results suggest that PEP activity and consequently the switch from NEP to PEP activity, is essential also during embryo development and that the PRIN2-CSP41b protein complex potentially is important for full PEP activity during this process.

## MATERIALS AND METHODS

### PLANT MATERIAL AND GROWTH CONDITIONS

Seedlings of *Arabidopsis thaliana* were grown on phytoagar plates containing  $1 \times$  Murashige and Skoog salt mixture supplemented with vitamins (Duchefa) and 2% sucrose. The T-DNA insertion lines: *prin2.2* (GK-772D07-024643) and *csp41b-2*

(SALK\_021748), were obtained from the European Arabidopsis Stock Centre. All genotypes are in Colombia ecotype. Seedlings and rosette plants were grown on soil at  $23^{\circ}\text{C}$  (16 h light  $150 \mu\text{mol}$  photons  $\text{m}^{-2} \text{s}^{-1}$ ) and  $18^{\circ}\text{C}$  (8 h dark).

### MORPHOLOGICAL ANALYSIS

Embryo isolation was done according to (Perry and Wang, 2003). Briefly, The siliques at the desired stage of DAP were dissected under a stereo microscope and the seeds were collected. Embryos were gently extruded from the seed coat by applying pressure on a glass plate covering them. Embryo/seed coat mixture was loaded a 25% Percoll gradient in isolation buffer (10 mM potassium phosphate, pH 7.0, 50 mM NaCl, 0.1 M sucrose) and centrifuged for 10 min at 800 g. The pellet was subjected to another round of purification using the 25% Percoll gradient. The pellet was re-suspended in isolation buffer and used for subsequent analysis. For transmission electron microscopy (TEM) pictures, 3-weeks-old plants grown on soil and the isolated seed embryos were prepared according to (Barajas-Lopez Jde et al., 2013).

### RNA ISOLATION, cDNA SYNTHESIS AND REAL-TIME PCR

Total RNA was isolated using Plant RNA Mini Kit (EZNA) and genomic DNA contamination was removed by DNase treatment (Fermentas). cDNA was synthesized using the iScript cDNA kit (Bio-Rad) according to the manufacturer's instructions and  $10 \times$  diluted. Real-Time PCR was performed using iQSYBR Green Supermix (Bio-Rad), with a final volume of  $10 \mu\text{L}$ . The PCR amplification was done using two-step protocol using the CFX96 Real-Time system (C1000 Thermal Cycler; Bio-Rad). All experiments were performed with three biological and three technical replicates, the relative gene expression was normalized to the expression of RCE1 (*At4g36800*) and PP2AA3 (*At1g13320*). Data analysis was done by CFX manager (Bio-Rad) and LinRegPCR software.

### EXPRESSION AND PURIFICATION OF RECOMBINANT PROTEINS

The coding sequences of PRIN2 and CSP41b were amplified with PCR. The PCR products were cloned using the NcoI-AccI sites into pET\_His1a vector. BL21 *Escherichia coli* cells were transformed with the expression constructs and induced for 6 h with 1 mM IPTG. Overexpressed proteins were affinity purified on  $\text{Ni}^{2+}$ -NTA agarose resin (Qiagen). The pET-His1a expression vector was kindly provided by Günter Stier, Umeå University, Sweden.

### EMSA

The 197 bp probe containing  $-196$  to  $+1$  PsaA promoter region was PCR amplified and labeled at the 3'-end with biotin-14-dCTP using biotin labeling kit (Invitrogen) according to the manufacturer's instructions. DNA-protein interactions were performed in 25  $\mu\text{L}$  reactions containing following reagents: 2.5  $\mu\text{L}$  of  $\times 10$  binding buffer (100 mM Tris HCl, 250 mM KCl, and 10 mM DTT), 1 mcg poly dIdC (Sigma-Aldrich), 2.5% glycerol, 0.05% Triton X-100, 5 mM MgCl<sub>2</sub>, 10 mM EDTA. The reaction mixture was incubated with DNA and protein at room temperature for 30min and was run on 6% native TBE-PAGE in

x0,5 TBE buffer at 100 V. DNA was transferred to nylon + membrane (Amersham) and was UV cross-linked to the membrane, incubated with Streptavidin-HRP and detected by Chemoluminescence Nucleic Acid Detection Module (Pierce) according to manufacturer's instructions (Shaikhali et al., 2012a,b).

#### Co-IP

To identify PRIN2 interacting partners, a 35S promoter linked to the full length PRIN2 coding sequence was cloned into the pGWB16\_myc expression construct. Col-0 plants were then transformed with 35S\_pGWB16myc\_PRIN2 using the floral-dip method (Clough and Bent, 1998). Two weeks old stable transformants overexpressing PRIN2 were used for chloroplast isolation in a two-step 50–25% Percoll gradient as described previously (Aronsson and Jarvis, 2002). Chloroplast proteins were extracted and incubated with 3 mg of anti-cMYC monoclonal antibody (Bio-Site) bound to the protein G-coated magnetic beads (Dynabeads Protein G Immunoprecipitation, Invitrogen) for 1 h at 4°C. All the washing steps were performed at 4°C according to the manufacturer's instructions. Immunoprecipitated PRIN2 protein complexes were eluted in SDS loading buffer containing 100 mM β-mercaptoethanol and proteins were separated on 6–12% SDS PAGE. Non-transformed Col-0 plants were used as negative control. For direct protein Co-IP assay, full length PRIN2 and CSP41b coding sequences were cloned into *NcoI*-*NcoI* sites of pRT104\_3myc and into *SacI*-*NotI* sites of pRT104\_3HA vector, respectively. Protoplasts obtained from *Arabidopsis* Ler-0 cell culture were transformed with pRT104\_PRIN2\_myc and pRT104\_CSP41b\_HA constructs as described previously by (Doelling and Pikaard, 1993).

#### ISOLATION OF THYLAKOID COMPLEXES AND BLUE NATIVE PAGE

Chloroplasts were isolated on a two-step 50–25% Percoll gradient from 4 weeks old rosette plants grown in short day as described previously (Aronsson and Jarvis, 2002). Thylakoid membranes were then purified (Hall et al., 2011) and protein complexes were solubilized in BN-solubilization buffer (30 mM HEPES, pH 7.4; 150 mM potassium acetate; 10% glycerole, 4% digitonin (SIGMA); 1% *b*-Dodecylmaltoside (SIGMA) for 40 min, 4°C. DM and digitonin in a mixture was shown to be suitable for better preservation of megacomplexes and at the same time good for solubilization (Järvi et al., 2011). 35 micro gram protein was loaded onto the 4–12% Bis-Tris Gel (NuPAGEH Novex 1.0 mm, Invitrogen) from each genotype.

## RESULTS

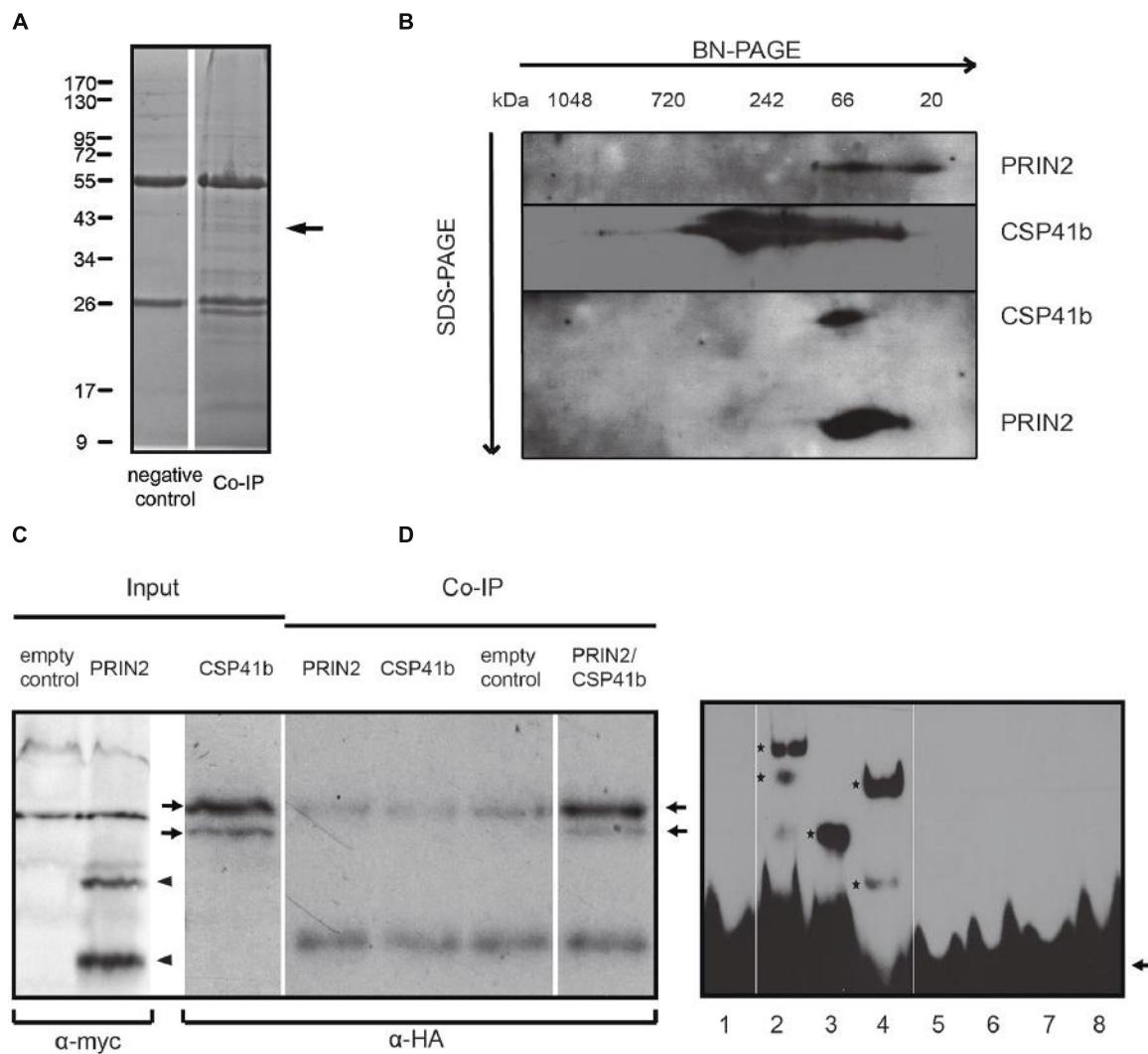
#### PRIN2 AND CSP41B FORM A DISTINCT PROTEIN COMPLEX THAT BINDS DNA

To understand the function of PRIN2 we wanted to identify proteins interacting with PRIN2 *in vivo*. To achieve this a co-IP approach was used. Full length PRIN2 protein fused to a cMyc-tag was expressed in *Arabidopsis* plants, intact chloroplasts were isolated and PRIN2-containing protein complexes were precipitated with anti-cMyc antibody, proteins were separated on SDS PAGE and distinct bands identified using mass spectrometry (Figure 1A). The Co-IP experiment was performed twice and as

many as 17 bands absent from the negative controls were in total cut from the gels. Most of the identified proteins were only found in one experiment and most likely represented unspecific interactions. However, one protein, CSP41b was identified in both experiments and several peptides with significant scores corresponding to CSP41b were identified in each sample (Table S1 in Supplementary Material). CSP41b is a conserved chloroplast protein and the *csb41b-2* mutant displayed impaired chloroplast transcription and plant development (Pfannschmidt et al., 2000; Suzuki et al., 2004; Bollenbach et al., 2009; Qi et al., 2012). The phenotype of the *csb41b-2* mutant was very similar to the phenotype of the *prin2* mutant and it is possible that CSP41b and PRIN2 are involved in the same process. Thus, the identified interaction between PRIN2 and CSP41b was chosen for further analysis.

First, to really confirm the interaction between CSP41b and PRIN2 purified recombinant proteins were separated on 2D electrophoresis (Figure 1B; Table S2). PRIN2 forms protein complexes ranging from 20–66 kDa while CSP41b is present in high molecular weight complexes ranging from 40–700 kDa. When PRIN2 and CSP41 are incubated together, PRIN2 seems to break the high molecular weight complexes of CSP41b and form a distinct protein complex with CSP41b suggesting heteromerization. To further confirm *in vivo* that PRIN2 and CSP41b directly interact with each other we used a third method by transiently expressing the full length proteins PRIN2 and CSP41b fused to cMyc- and HA-tags, respectively, in *Arabidopsis* protoplasts. Two bands of approximately 50 kDa in size correspond to CSP41b in the CSP41b-HA transformed protoplasts and these could be cytoplasmic/transite peptide-processed forms of the protein and/or proteolytic fragment of the protein. The PRIN2 protein exists in monomeric, dimeric and oligomeric forms (unpublished data) and migrates under these conditions as two bands of approximately 20 and 40 kDa, respectively, most likely corresponding to monomer and dimer. CSP41b was detected in the Co-IP fraction confirming interaction with PRIN2 (Figure 1C). Thus, using three independent methods CSP41b and PRIN2 were shown to directly interact.

Both PRIN2 and CSP41b were suggested to regulate transcription of PEP dependent chloroplast genes (Bollenbach et al., 2009; Kindgren et al., 2012). To investigate if PRIN2 and CSP41b interact upon DNA binding *in vitro*, the proteins were incubated with a 197 bp DNA probe containing –196 to +1 region of the *psaA* promoter from *Arabidopsis*. In the electrophoretic mobility shift assay (EMSA) both PRIN2 and CSP41b bound the labeled probe, PRIN2 formed at least two distinct complexes with DNA, while CSP41b formed only one DNA-protein complex (Figure 1D). PRIN2 and CSP41b are about 15 and 40 kDa, respectively, and the difference in the migration of the protein/DNA complexes of PRIN2 and CSP41b suggests that PRIN2 forms higher molecular weight oligomeric complexes and/or binds to several regions of the DNA probe. When PRIN2 and CSP41b were incubated together with DNA, a new band, intermediate in size to what was observed for the individual proteins, was detected that most likely corresponded to a *psaA*197-PRIN2/CSP41b heteromeric complex. The competition reactions with unlabeled *psaA*-198 bp confirm the DNA binding capacity of PRIN2 and CSP41b. A similar



**FIGURE 1 | PRIN2 interacts with CSP41b and forms a DNA-binding complex.** (A) Native PRIN2-cMyc protein complex from *Arabidopsis* chloroplasts was immunoprecipitated with anti-cMyc antibody and proteins were separated on the SDS-PAGE. Non-transformed *Col-0* chloroplasts were used as a negative control. Protein bands were cut from the gel and analyzed by mass spectrometry. An arrow indicates the position of the band where CSP41b was identified. (B) PRIN2 interaction with CSP41b detected by 2D Native PAGE/SDS PAGE gel. Recombinant PRIN2, CSP41b and 1:1 mixture of both proteins were incubated in the binding buffer and run on the 4–12% Native Bis-Tris Gel. Individual stripes were then cut and loaded on the 12% SDS PAGE gel. Proteins were detected with anti-His peroxidase. (C) Interaction between PRIN2 and CSP41b *in vivo* Co-IP assay. Full length PRIN2-cMyc and CSP41b-HA fusion proteins were expressed in *Arabidopsis* protoplasts and Co-IP was performed using anti-cMyc antibody linked to protein G-covered magnetic beads.

Immunoblotted samples from input and Co-IP probes were detected with either anti-cMyc chicken IgY fraction and rabbit HRP-linked anti-chicken IgY (H+L) or with anti-HA peroxidase. Arrowheads indicate PRIN2 protein and arrows show CSP41b protein bands detected by anti-HA antibody. (D) DNA binding of PRIN2 and CSP41b and their heteromerization in EMSA assay. Signal from *psaA*-198 bp biotin labeled probe was detected by chemiluminescence nucleic acid detection module. 3 µg of each purified protein was used in every reaction, the DNA/protein molar ratio was 1:100. DNA/protein complexes are marked with asterisks, free DNA probe with an arrow. Competition was done with unlabeled *psaA*-198 bp probe with 50-fold excess over the labeled probe. 3 µg of BSA was used as a negative control. (1) free probe, (2) probe + PRIN2, (3) probe + CSP41b, (4) probe + PRIN2 + CSP41b, (5) probe + PRIN2 + unlabeled probe, (6) probe + CSP41b + unlabeled probe, (7) probe + PRIN2 + CSP41b + unlabeled probe, (8) probe + BSA.

pattern of DNA-binding could also be observed when using the *psbA*-198 bp fragment as a probe, where PRIN2 and CSP41b proteins form heteromeric protein/DNA complexes (Supplementary Figure S1). Thus, both PRIN2 and CSP41b are able to interact with DNA in the EMSA assay and the formation of the heteromeric PRIN2/CSP41b complex also occurs upon DNA binding *in vitro*.

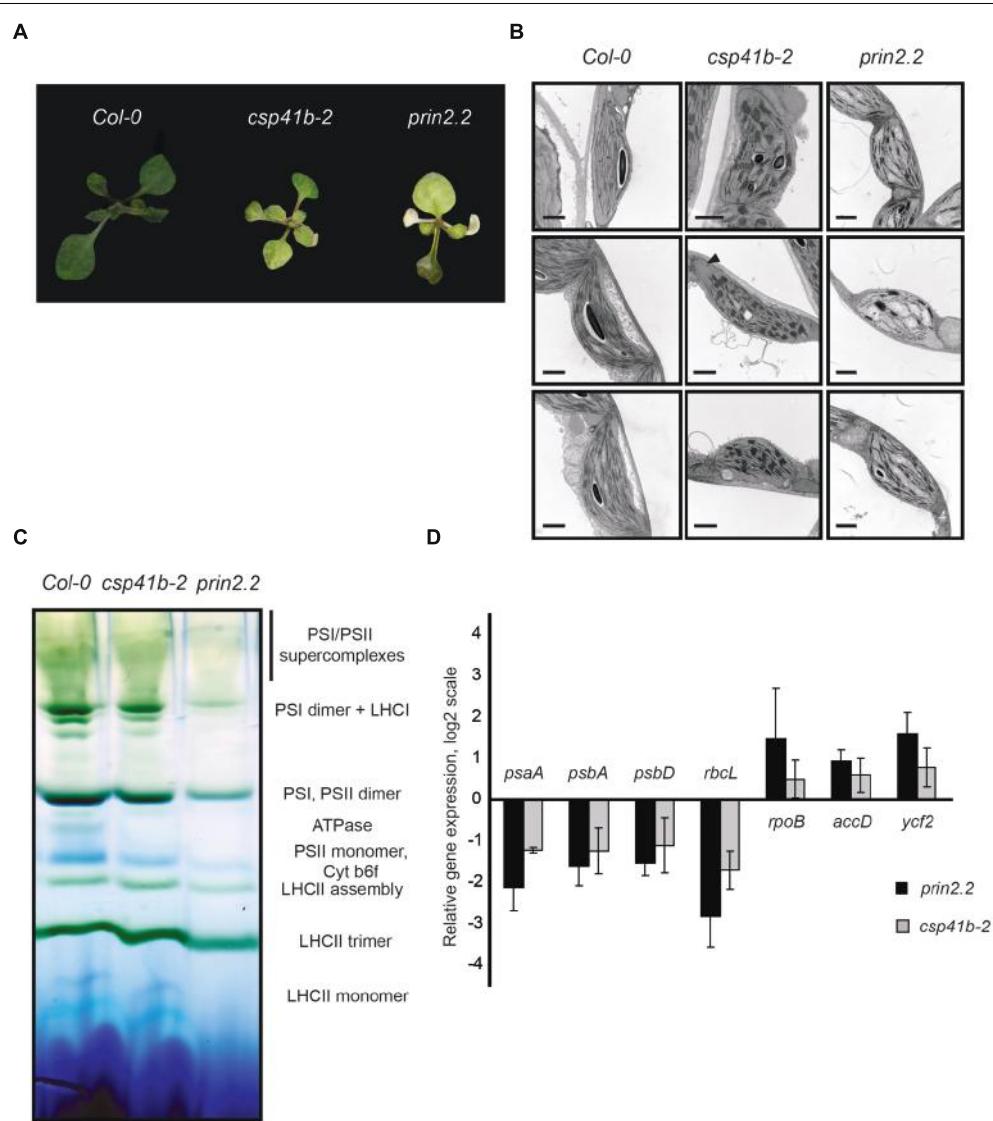
#### THE *prin2.2* AND *csp41b-2* MUTANTS SHOW DISTINCT PHENOTYPES AND IMPAIRED EXPRESSION OF CHLOROPLAST ENCODED GENES

The phenotypes of the *prin2.2* and *csp41b-2* mutants are similar to mutants of defined components of the PEP complex (De Santis-Maclossek et al., 1999; Krause et al., 2000; Legen et al., 2002; Pfalz et al., 2006). As has been demonstrated before *prin2.2* and *csp41b-2* showed a clear reduction in growth rate

and pale leaves compared to wild type (**Figure 2A**). The *prin2.2* plants have impaired chloroplast structure with reduced thylakoid membranes and grana stacks (**Figure 2B**). In contrast, the *csp41b-2* mutant showed an increased number of thylakoid membranes organized in grana stacks (**Figure 2B**). Moreover, large areas of the chloroplasts (depicted with an arrowhead) were devoid of any thylakoid membranes in the *csp41b-2* mutant. Correlated with aberrant chloroplast structure is the observed defect in the organization of the photosynthetic complexes in the *prin2.2* and *csp41b-2* mutants. The mutants showed reduced amounts of PSI, PSII, ATPase, and antenna complexes compared to wild type (**Figure 2C**). The defect in photosynthetic complex

stoichiometry and organization was especially strong for the *prin2.2* mutant.

Expression analyses were performed for genes categorized as genes primarily transcribed by PEP and NEP, respectively. The *psaA*, *psbA*, *psbD*, and *rbcL* genes belong to class I genes transcribed by the PEP polymerase. Consistent with previous studies (Bollenbach et al., 2009; Kindgren et al., 2012; Qi et al., 2012), both *prin2.2* and *csp41b-2* mutants showed decreased *psaA*, *psbA*, *psbD*, and *rbcL* expression levels compared to wild type. In contrast, the expression levels of the NEP genes, *accD*, *rpoB*, *ycf2* were elevated compared to wild type (**Figure 2D**). The chloroplast transcriptional machinery



**FIGURE 2 |**The *prin2* and *csp41b-2* show severe plant phenotypes. Plants were grown on soil under long day conditions (16 h light/8 h darkness) to characterize *prin2.2* and *csp41b-2* mutants. **(A)** Representative images of plants grown on soil for 3 weeks. **(B)** Representative TEM images of chloroplasts from the *Col-0*, *csp41b-2*, and *prin2.2* rosette plants. Bars represent 1 μm. **(C)** Blue Native PAGE of thylakoid membrane complexes isolated from *Col-0*, *csp41b-2*, and *prin2.2* chloroplasts. **(D)** Relative expression levels of chloroplast encoded genes in *prin2.2* and *csp41b-2* mutant plants. Expression levels were compared to the respective *Col-0* samples and calculated using Ubiquitin-protein ligase (At4g36800) as a reference gene. Data represents the mean from three independent biological replicates.

isolated from *Col-0*, *csp41b-2*, and *prin2.2* chloroplasts. **(D)** Relative expression levels of chloroplast encoded genes in *prin2.2* and *csp41b-2* mutant plants. Expression levels were compared to the respective *Col-0* samples and calculated using Ubiquitin-protein ligase (At4g36800) as a reference gene. Data represents the mean from three independent biological replicates.

that primarily utilizes PEP polymerase is clearly impaired in the *prin2.2* and *csp41b-2* mutants and when the two mutants were compared side by side the effect was stronger in the *prin2.2* mutant.

#### THE *prin2.2* AND *csp41b-2* MUTANTS SHOW DEFECTS IN EMBRYO DEVELOPMENT AND THE *csp41b-2 prin2.2* DOUBLE MUTANT IS EMBRYO LETHAL

In order to investigate the genetic interaction between PRIN2 and CSP41b we attempted to generate a *csp41b-2 prin2.2* double mutant (Supplementary Figure S2). However, the *csp41b-2 prin2.2* double mutant was embryo lethal and the *CSP41b-2 prin2.2/csp41b-2 prin2.2* mutant produced siliques where 18% (green:albino ovule = 240:51, Chi-square 8,67 for  $p < 0,05$ ) of all ovules appeared opaque (**Figure 3A**, arrows). Those impaired ovules finally turned into shrunken, dark colored seeds unable to germinate on MS media. A few ovules were also aborted at very early developmental stages (**Figure 3A**, arrowheads). The fact that theoretical 3:1 segregation green/albino is not supported by our statistics could be explained by the observed range in the stage at which the embryo development is arrested. Given the embryo lethality of the double mutant we investigated if there were any effects during embryo development in the *csp41b-2* and *prin2.2* single mutants. The *csp41b-2* embryos were undistinguishable from the wild type at the heart stage (**Figure 3B**). However, at the linear cotyledon and mature green (MG) stages the *csp41b-2* embryos were not as uniformly green as the wild type embryos. As has been shown before wild type embryos displayed a specific pattern of chlorophyll autofluorescence during embryogenesis (Tejos et al., 2010). This pattern was significantly altered at the linear cotyledon stage (LC), where the chloroplast containing tissue was mostly localized to the epidermal layers, in the *csp41b-2* embryos (**Figure 3C**). Distribution of chloroplast containing tissue was even more altered in the *prin2.2* embryos (**Figure 3C**). Consistent with these findings are the light microscopy pictures demonstrating that the *prin2.2* embryos were paler than wild type embryos at all developmental stages (**Figure 3B**).

To investigate the potential role of PRIN2 and CSP41b during embryo development, we used TEM to examine morphological differences in the embryos at the MG stage. Plastids of wild type embryos develop normally and have numerous thylakoids organized in grana stacks indicating that the plastids can be photosynthetically functional at this stage of embryo development (**Figure 4**). However, *csp41b-2* chloroplasts developed less thylakoid membranes and fewer grana stacks and exhibit quite often chloroplasts with large areas completely devoid of membranes (**Figure 4**). This specific defect in chloroplast structure of *csp41b-2* was also observed for the chloroplasts from 3-week-old plants (**Figure 2B**). The *prin2.2* showed an even stronger impairment in chloroplast development, the *prin2.2* plastids showed numerous vesicles and few thylakoid membranes and grana stacks. Thylakoid membranes were also often mis-oriented (**Figure 4**). Similarly to what was observed for the *csp41b-2* mutant the chloroplasts from the embryos of *prin2.2* showed similar defects to what was also seen in the adult plants (**Figure 2B**). Taken together our results indicate that chloroplast development in the embryo is impaired

in both *prin2.2* and *csp41b-2* single mutants and that the ovules are arrested at early developmental stages in the *csp41b-2 prin2.2* double mutant.

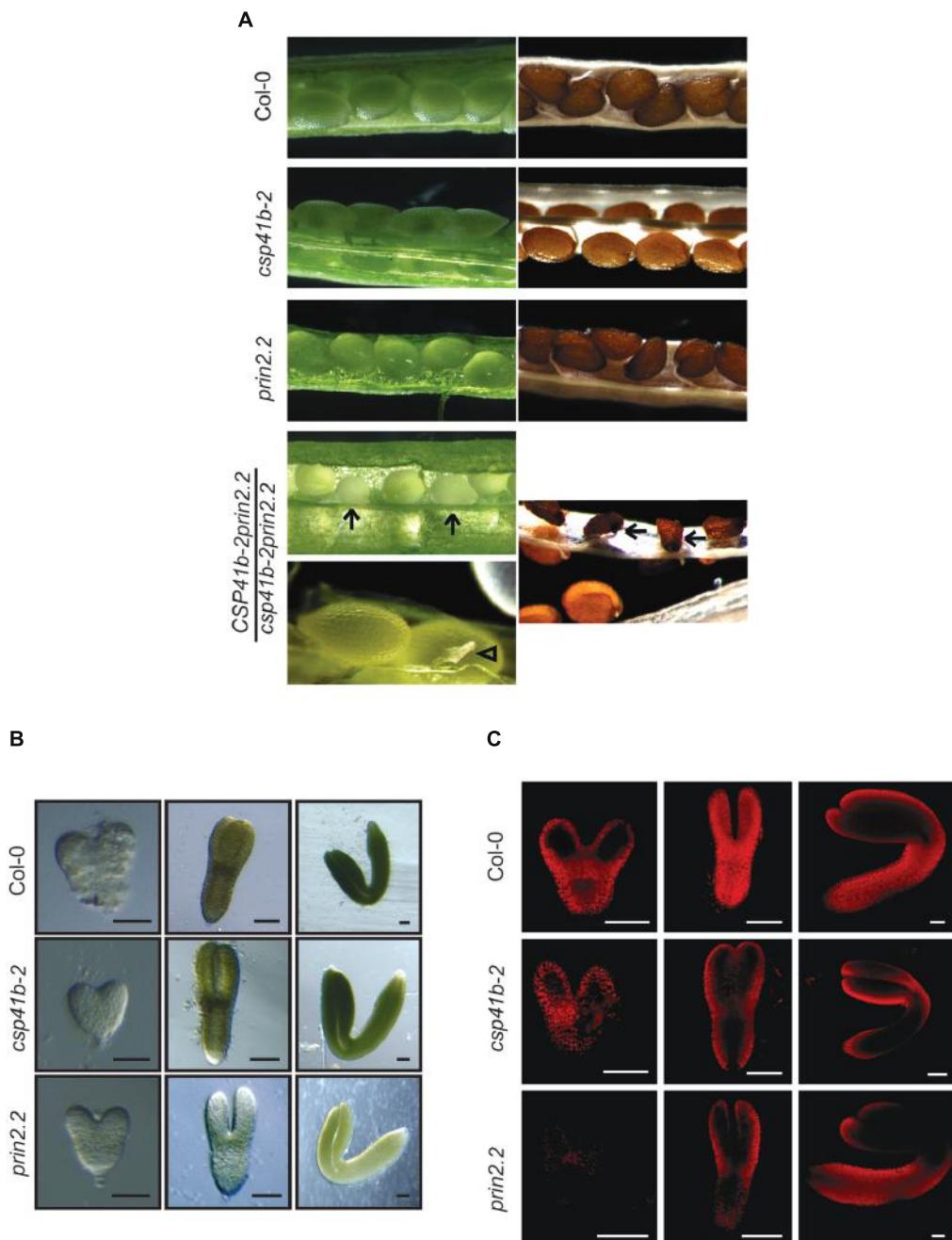
#### EXPRESSION OF PHOTOSYNTHESIS GENES ESSENTIAL DURING EMBRYO DEVELOPMENT

We performed *in silico* analysis of available array data where gene activity was profiled genome-wide in every organ, tissue, and cell type of *Arabidopsis* seeds from fertilization through maturity (Belmonte et al., 2013). We specifically investigated the expression of plastid encoded genes from preglobular (PG) to MG embryo stage. Overall transcription of chloroplast-encoded photosynthesis associated genes was activated from globular to MG embryo stage (**Figure 5A**). The highest fold change in expression level was observed for the genes encoding PSI and PSII core subunits, ATPase, and the ribosomal subunits (**Figure 5A**). Thus, from the analysis of the plastid transcriptome it is clear that photosynthesis associated components are highly expressed during embryo development and that therefore PEP mediated transcription most likely is activated during this process. Expression of the nuclear encoded components associated with the PEP complex was demonstrated to increase from PG to LC (**Figure 5B**). Especially the genes encoding the PTAC proteins showed a very strong upregulation during embryo development. In addition, transcript levels of *TRXZ* were significantly up-regulated during the transition from PG to LC stage. Also transcription of *PRIN2* and *CSP41b* were observed during the LC stage of embryo development (**Figure 5B**).

The observed embryo phenotype in the *prin2.2* and *csp41b-2* mutants encouraged us to study expression of *psaA*, *psbA*, and *psbD* at the MG embryo stage. Expression of *psaA*, *psbA*, and *psbD* was strongly down regulated both in *prin2.2* and *csp41b-2* mutants compared to wild type (**Figure 5B**; Supplementary Figure S3). Thus, these results suggest that also under embryo development PRIN2 and CSP41b are important for proper transcription of PEP genes. Taken together, these results suggest that the PEP component of the chloroplast transcription machinery is active during embryo development and that its activity is essential for proper embryo and seed development.

#### DISCUSSION

In leaf tissue, the initiation of chloroplast development in the light and the activation of the photosynthetic reactions are accompanied by repression of NEP activity and an increase of PEP-mediated plastid transcription (Hanaoka et al., 2005). However, the mechanisms underlying this change in major RNA polymerase activity and the division of labor between NEP and PEP in the chloroplast are unknown (Zhelyazkova et al., 2012). Expression of plastid-encoded photosynthetic components, thought to be mediated exclusively by PEP, was recently shown both in tobacco and barley to be driven by NEP in the absence of functional PEP, suggesting a less strict division of target genes between NEP and PEP (Legen et al., 2002; Lyubetsky et al., 2011; Zhelyazkova et al., 2012). NEP is particularly active in non-green tissues and in very young leaves (Emanuel et al., 2006). Similarly, the expression of the previously described



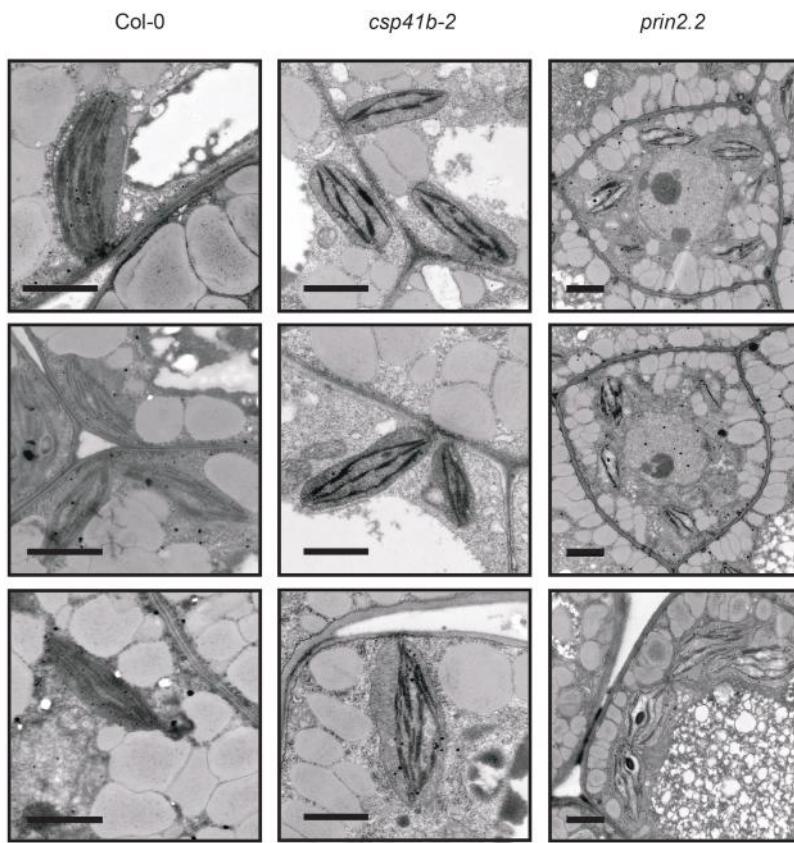
**FIGURE 3 |** The *csp41b-2prin2.2/csp41b-2prin2.2* double mutant is lethal. **(A)** Siliques from self-fertilized *CSP41b-2prin2.2/csp41b-2prin2.2* double mutant. Arrows indicate aborted seeds and arrowheads show aborted ovules. **(B)** Embryo development of Col-0, *csp41b-2*, and *prin2.2*. Representative images are shown for each genotype. Bars represent

50  $\mu$ m. **(C)** Chlorophyll autofluorescence in Col-0, *csp41b-2*, and *prin2.2* during specific stages of embryo development detected using confocal microscopy (Excitation wavelength 488 nm, Emission wavelength 687 nm). Representative images are shown for each genotype. Bars represent 50  $\mu$ m.

NEP-dependent genes was distinctly higher in developing compared with mature chloroplasts (Zoschke et al., 2007), indicating that NEP is very active during the early stages of chloroplast biogenesis. However, we have demonstrated that PEP activity is essential during embryo development in *Arabidopsis*. During embryo development NEP was unable to compensate for

impaired PEP activity in the *prin2.2* and *csp41b-2* mutants. Thus, the switch from NEP- to PEP-dependent transcription of the plastid-encoded genes is essential for embryo and seed development in *Arabidopsis*.

*In silico* analyses of available microarray data (Belmonte et al., 2013) showed a gradual increase in expression of genes encoding

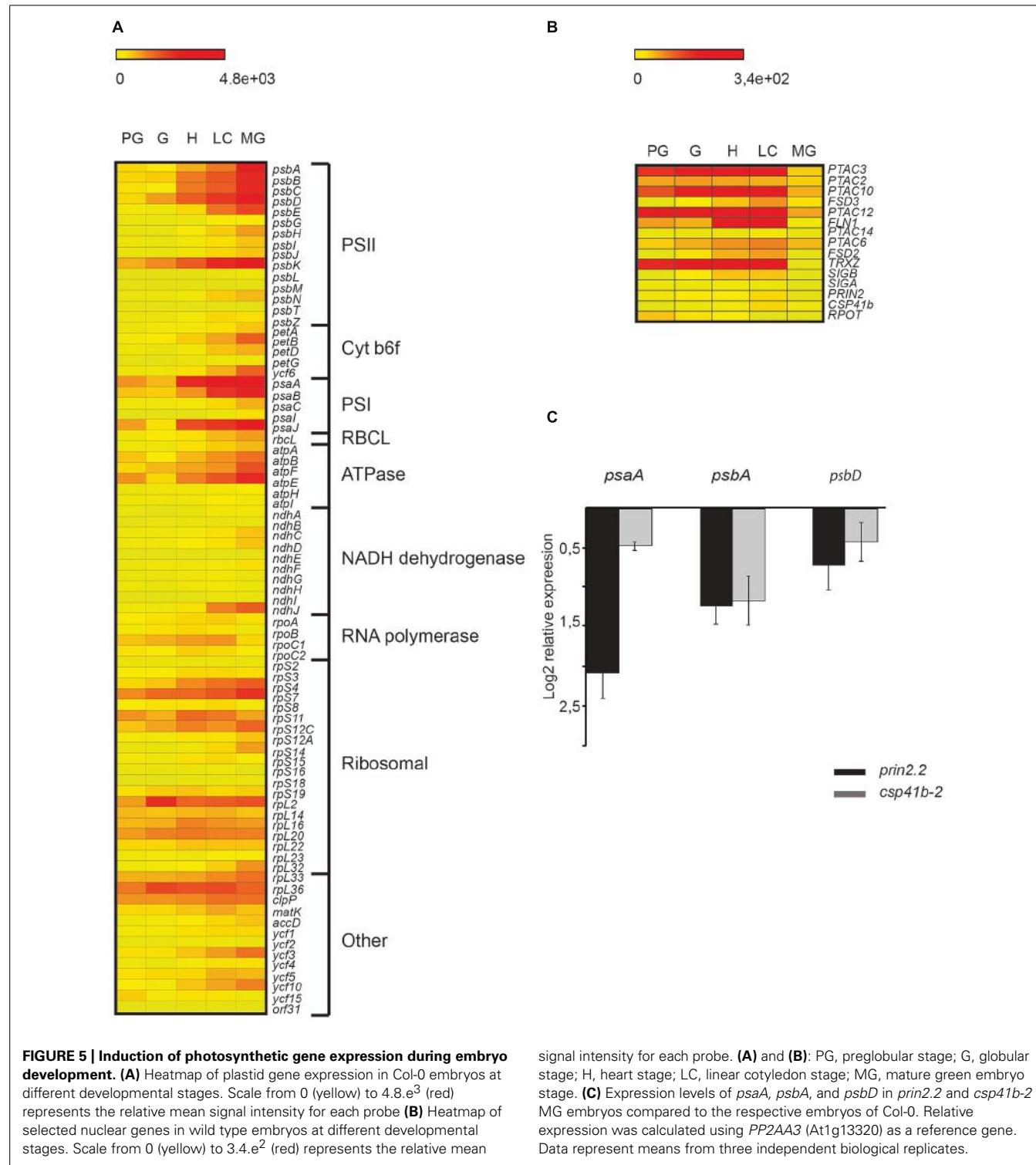


**FIGURE 4 | The *csp41b-2* and *2prin2.2* mutants display impaired chloroplast during embryo development.** Representative TEM images of chloroplasts from the Col-0, *csp41b-2*, and *prin2.2* mature green (MG) embryos. Bars represent 1  $\mu$ m.

components involved in photosynthesis and energy production from the PG to the MG stage of embryo development (Figure 5A). During embryogenesis, expression of *RPO*T encoding the plastid NEP polymerase was induced and the maximum expression level was reached as early as at the PG and globular stages (G; Figure 5B). The *RPO*T expression then diminished already at the heart stage (H; Figure 5B). The NEP promoter is the only promoter in the *rpoBC* operon, and expression of *rpoA*, *rpoB*, and *rpoC1/C2* followed the induction of *RPO*T. The core components of PEP were transcribed from the globular stage but in contrast to *RPO*T, expression was maintained all through the different embryo developmental stages (Figure 5A). The expression profile of the nuclear encoded sigma factors, especially SIGB (SIG2) was similar to the profiles of *rpoA*, *rpoB*, and *rpoC1/C2* (Figure 5B). Correlated with the induction of the genes encoding the core components of PEP was a very strong induction of the plastid encoded photosynthesis genes. Expression of *psbA,B,C,D,E*, and *psaA,B,J*, for example, increased approximately 10-fold when the MG embryo was compared to the PG suggesting that PEP transcription is activated during this developmental transition. The strong induction of photosynthesis related mRNAs during the PG to MG transition was also shown in a study where three periods of seed formation was investigated (Allorent et al., 2013). The embryo lethality of the

*csp41b-2prin2.2* double mutant further indicates that PEP driven transcription is required for effective transcription of photosynthesis genes to sustain the embryo with energy. It was also shown previously that genes associated with photosynthesis and carbon metabolism, are active in all embryo and endosperm sub-regions during early seed development (Belmonte et al., 2013) strongly suggesting there is a need for photosynthetic activity to contribute to embryo growth rate and biomass. Interestingly, at the MG embryo stage transcription of the nuclear encoded components of PEP was halted, suggesting that there is no need for further expression of chloroplast genes when the seeds enter the post-MG stage. In contrast, the chloroplast encoded photosynthesis genes showed a delayed response and maintained high expression also at the MG embryo stage. Possibly chloroplast transcription is subject to later anterograde regulation from the nucleus to repress expression.

True to its cyanobacterial origin, the core subunits of PEP are homologous to the cyanobacterial RNAP components (Martin et al., 2002). However, PEP also requires additional nuclear-encoded factors for its function (Pfannschmidt et al., 2000; Pfalz et al., 2006). As many as 40–60 proteins appear to be present in the TAC from chloroplasts. From Arabidopsis and mustard TACs 35 components were identified and 18 of those components, called pTACs, were novel proteins (Pfalz et al., 2006).



In addition, TAC and sRNAP preparations from pro-plastids, chloroplasts and etioplasts have different protein compositions, suggesting a multifaceted regulation of PEP activity (Reiss and Link, 1985; Pfannschmidt and Link, 1994; Suck et al., 1996). Very high expression levels were observed for genes encoding the pTAC components PTAC3, PTAC10, PTAC12, FLN1, PTAC6, and TRXZ

during the globular (G) to LC stages. Expression of *PRIN2* and *CSP41B* was also detected during the LC stage. The expression of all these additional PEP components suggests that PEP activity requires many different components already during embryo development and that regulation of plastid transcription is complex and sophisticated from the very early stages of plant development.

PRIN2 and CSP41b were both identified in nucleoid preparations (Majeran et al., 2012) and our results suggest that these two components and possibly the PRIN2-CSP41b complex are essential for PEP dependent transcription during early embryo development. Using three independent methods CSP41b and PRIN2 were shown to directly interact (Figure 1). PRIN2 have two conserved Cys residues that possibly are responsible for monomer/dimer/oligomer formation upon oxidation. Consistent with this hypothesis is the observation of two bands of 20 and 40 kDa that might correspond to PRIN2 monomer and dimer (Figure 1C). Interestingly, the CSP41b protein migrated as continuous multimeric complex ranging from 40 to ~700 kDa on the BN-PAGE. The presence of high molecular weight complexes was previously described for the native CSP41b protein in chloroplasts (Qi et al., 2012). Moreover, the CSP41b protein contains redox active Cys residues suggested to be putative targets of TRX (Stroher and Dietz, 2008) and the formation of high molecular weight complexes were also shown to be enhanced by oxidized conditions in the chloroplast stroma (Qi et al., 2012). When PRIN2 and CSP41b were incubated together they migrated on the gel as a distinct band (~66 kDa; Figure 1B) suggesting the formation of a defined heteromeric protein complex containing both proteins. PRIN2 and CSP41b were shown to independently bind to promoter fragments of *psaA* and *psbA* in EMSA assays. However, when PRIN2 and CSP41b were incubated together, a new band of intermediate size was observed (Figure 1D, Supplementary data Figure S1), suggesting an interaction between PRIN2 and CSP41b also upon DNA binding *in vitro*. CSP41b was previously shown in RIP-chip analysis to be an RNA binding protein with specificity toward photosynthesis-related transcripts mainly expressed by PEP (Qi et al., 2012). However, many RNA binding proteins have also been described to have DNA binding properties, including TFIIIA, p53, STAT1,  $\beta/\beta'$  subunits of RNAP and  $\sigma70$ , factors well known to regulate transcription (Sakonju et al., 1980; Clemens et al., 1993; Cassiday and Maher, 2002; Suswam et al., 2005). Another plant specific protein that binds both DNA and RNA is GUN1, a key component in retrograde communication between chloroplasts and nucleus (Koussevitzky et al., 2007). Possibly, the observed DNA binding of the PRIN2-CSP41b protein complex is required for full PEP activity as indicated by the embryo lethality of the *csp41b-2prin2.2* double mutant (Figure 3A).

Correlated with the embryo phenotype observed in *prin2.2* and *csp41b-2* single mutants was the impaired expression of chloroplast photosynthesis genes. The expression levels of *psaA*, *psbA*, and *psbD* were significantly lower in both mutants compared to wild type in MG embryos (Figure 5C), suggesting that the previously described defect in PEP-mediated gene expression in rosette plants is maintained in the mutants also during embryo development. Thus, our results establish a link between PEP activity and embryo development. Previously compromised translational and post-translational activities in the chloroplasts, such as mutations in elongation factor G and PPR proteins, have been shown to lead to embryo lethality (Ruppel and Hangarter, 2007; Khrouchchova et al., 2012; Sosso et al., 2012). Interestingly, in tobacco and barley neither of the knockouts completely lacking PEP activity exhibits an embryo lethal phenotype (Allison

et al., 1996; De Santis-MacIossek et al., 1999; Zhelyazkova et al., 2012). However, it should be emphasized that an essential role of plastid activity and embryo greening during embryogenesis has so far only been reported for oil-seed plants such as *Arabidopsis* and *Brassica* (He and Wu, 2009; Hsu et al., 2010). In *Arabidopsis*, PEP activity appears essential during embryo development and during this process NEP is unable to compensate for the impaired PEP activity. Recently maize mutants lacking several PEP-associated proteins such as PTACs and PRIN2 demonstrated deficiency of numerous plastid tRNAs. Thus, a role for PEP, and for the PEP associated proteins, was demonstrated for the expression of plastid transfer RNA (Williams-Carrier et al., 2014). These results emphasize the complex division of labor between NEP and PEP during the initiation of chloroplast development and that PEP, and its associated proteins, are also essential for the translation of the photosynthetic components.

## AUTHOR CONTRIBUTIONS

Dmitry Kremnev carried out the experiments and analyzed the data. Dmitry Kremnev and Åsa Strand planned the study and wrote the manuscript. Both authors read and approved the final manuscript.

## ACKNOWLEDGMENT

This work was supported by grants from the Swedish research foundation, VR (ÅS).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://www.frontiersin.org/journal/10.3389/fpls.2014.00385/abstract>

## REFERENCES

- Allison, L. A., Simon, L. D., and Maliga, P. (1996). Deletion of *rpoB* reveals a second distinct transcription system in plastids of higher plants. *EMBO J.* 15, 2802–2809.
- Allorent, G., Florence, C., Chevalier, F., and Lerbs-Mache, S. (2013). Plastid gene expression during chloroplast differentiation and dedifferentiation in non-photosynthetic plastids during seed formation. *Plant Mol. Biol.* 82, 59–70. doi: 10.1007/s11103-0037-0
- Aronsson, H., and Jarvis, P. (2002). A simple method for isolating import-competent *Arabidopsis* chloroplasts. *FEBS Lett.* 529, 215–220. doi: 10.1016/S0014-5793(02)03342-2
- Barajas-Lopez Jde, D., Kremnev, D., Shaikhali, J., Pinas-Fernandez, A., and Strand, Å. (2013). PAPP5 is involved in the tetrapyrrole mediated plastid signalling during chloroplast development. *PLoS ONE* 8:e60305. doi: 10.1371/journal.pone.0060305
- Belmonte, M. E., Kirkbride, R. C., Stone, S. L., Pelletier, J. M., Bui, A. Q., Yeung, E. C., et al. (2013). Comprehensive developmental profiles of gene activity in regions and subregions of the *Arabidopsis* seed. *Proc. Natl. Acad. Sci. U.S.A.* 110, E435–E444. doi: 10.1073/pnas.1222061110
- Bollenbach, T. J., Sharwood, R. E., Gutierrez, R., Lerbs-Mache, S., and Stern, D. B. (2009). The RNA-binding proteins CSP41a and CSP41b may regulate transcription and translation of chloroplast-encoded RNAs in *Arabidopsis*. *Plant Mol. Biol.* 69, 541–552. doi: 10.1007/s11103-008-9436-z
- Bryant, N., Lloyd, J., Sweeney, C., Myouga, F., and Meinke, D. (2011). Identification of nuclear genes encoding chloroplast-localized proteins required for embryo development in *Arabidopsis*. *Plant Physiol.* 155, 1678–1689. doi: 10.1104/pp.110.168120
- Cassiday, L. A., and Maher, L. J. III. (2002). Having it both ways: transcription factors that bind DNA and RNA. *Nucleic Acids Res.* 30, 4118–4126. doi: 10.1093/nar/gkf512

- Clemens, K. R., Wolf, V., Mcbryant, S. J., Zhang, P., Liao, X., Wright, P. E., et al. (1993). Molecular basis for specific recognition of both RNA and DNA by a zinc finger protein. *Science* 260, 530–533. doi: 10.1126/science.8475383
- Clough, S. J., and Bent, A. F. (1998). Floral dip: a simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. *Plant J.* 16, 735–743. doi: 10.1046/j.1365-313x.1998.00343.x
- De Santis-Maclossek, G., Kofer, W., Bock, A., Schoch, S., Maier, R. M., Wanner, G., et al. (1999). Targeted disruption of the plastid RNA polymerase genes rpoA, B and C1: molecular biology, biochemistry and ultrastructure. *Plant J.* 18, 477–489. doi: 10.1046/j.1365-313X.1999.00473.x
- Doelling, J. H., and Pikaard, C. S. (1993). Transient expression in *Arabidopsis thaliana* protoplasts derived from rapidly established cell suspension cultures. *Plant Cell Rep.* 12, 241–244. doi: 10.1007/BF00237127
- Emmanuel, C., Von Groll, U., Muller, M., Borner, T., and Weihe, A. (2006). Development- and tissue-specific expression of the RpoT gene family of *Arabidopsis* encoding mitochondrial and plastid RNA polymerases. *Planta* 223, 998–1009. doi: 10.1007/s00425-005-0159-y
- Goffman, F. D., Alonso, A. P., Schwender, J., Shachar-Hill, Y., and Ohlrogge, J. B. (2005). Light enables a very high efficiency of carbon storage in developing embryos of rapeseed. *Plant Physiol.* 138, 2269–2279. doi: 10.1104/pp.105.063628
- Hajdukiewicz, P. T., Allison, L. A., and Maliga, P. (1997). The two RNA polymerases encoded by the nuclear and the plastid compartments transcribe distinct groups of genes in tobacco plastids. *EMBO J.* 16, 4041–4048. doi: 10.1093/emboj/16.13.4041
- Hall, M., Mishra, Y., and Schroder, W. P. (2011). Preparation of stroma, thylakoid membrane, and lumen fractions from *Arabidopsis thaliana* chloroplasts for proteomic analysis. *Methods Mol. Biol.* 775, 207–222. doi: 10.1007/978-1-61779-237-3\_11
- Hanaoka, M., Kanamaru, K., Fujiwara, M., Takahashi, H., and Tanaka, K. (2005). Glutamyl-tRNA mediates a switch in RNA polymerase use during chloroplast biogenesis. *EMBO Rep.* 6, 545–550. doi: 10.1038/sj.embor.7400411
- Hassidim, M., Yakir, E., Fradkin, D., Hilman, D., Kron, I., Keren, N., et al. (2007). Mutations in CHLOROPLAST RNA BINDING provide evidence for the involvement of the chloroplast in the regulation of the circadian clock in *Arabidopsis*. *Plant J.* 51, 551–562. doi: 10.1111/j.1365-313X.2007.03160.x
- He, Y. Q., and Wu, Y. (2009). Oil body biogenesis during *Brassica napus* embryogenesis. *J. Integr. Plant Biol.* 51, 792–799. doi: 10.1111/j.1744-7909.2009.00851.x
- Hedtke, B., Borner, T., and Weihe, A. (1997). Mitochondrial and chloroplast phage-type RNA polymerases in *Arabidopsis*. *Science* 277, 809–811. doi: 10.1126/science.277.5327.809
- Hedtke, B., Borner, T., and Weihe, A. (2000). One RNA polymerase serving two genomes. *EMBO Rep.* 1, 435–440. doi: 10.1093/embo-reports/kvd086
- Hsu, S. C., Belmonte, M. F., Harada, J. J., and Inoue, K. (2010). Indispensable roles of plastids in *Arabidopsis thaliana* embryogenesis. *Curr. Genomics* 11, 338–349. doi: 10.2174/138920210791616716
- Järvi, S., Suorsa, M., Paakkarinen, V., and Aro, E.-M. (2011). Optimized native gel systems for separation of thylakoid protein complexes: novel super- and mega-complexes. *Biochem. J.* 439, 207–214. doi: 10.1042/BJ20102155
- Khrouchtchova, A., Monde, R. A., and Barkan, A. (2012). A short PPR protein required for the splicing of specific group II introns in angiosperm chloroplasts. *RNA* 18, 1197–1209. doi: 10.1261/rna.032623.112
- Kindgren, P., Kremnev, D., Blanco, N. E., De Dios Barajas Lopez, J., Fernandez, A. P., Tellgren-Roth, C., et al. (2012). The plastid redox insensitive 2 mutant of *Arabidopsis* is impaired in PEP activity and high light-dependent plastid redox signalling to the nucleus. *Plant J.* 70, 279–291. doi: 10.1111/j.1365-313X.2011.04865.x
- Koussevitzky, S., Nott, A., Mockler, T. C., Hong, F., Sachetto-Martins, G., Surpin, M., et al. (2007). Signals from chloroplasts converge to regulate nuclear gene expression. *Science* 316, 715–719. doi: 10.1126/science.1140516
- Krause, K., Maier, R. M., Kofer, W., Krupinska, K., and Herrmann, R. G. (2000). Disruption of plastid-encoded RNA polymerase genes in tobacco: expression of only a distinct set of genes is not based on selective transcription of the plastid chromosome. *Mol. Gen. Genet.* 263, 1022–1030. doi: 10.1007/PL00008690
- Le, B. H., Cheng, C., Bui, A. Q., Wagstaiger, J. A., Henry, K. F., Pelletier, J., et al. (2010). Global analysis of gene activity during *Arabidopsis* seed development and identification of seed-specific transcription factors. *Proc. Natl. Acad. Sci. U.S.A.* 107, 8063–8070. doi: 10.1073/pnas.1003530107
- 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. doi: 10.1046/j.1365-313X.2002.01349.x
- Lyubetsky, V. A., Zverkov, O. A., Rubanov, L. I., and Seliverstov, A. V. (2011). Modeling RNA polymerase competition: the effect of sigma-subunit knockout and heat shock on gene transcription level. *Biol. Direct* 6, 3. doi: 10.1186/1745-6150-6-3
- Majeran, W., Friso, G., Asakura, Y., Qu, X., Huang, M., Ponnala, L., et al. (2012). Nucleoid-enriched proteomes in developing plastids and chloroplasts from maize leaves: a new conceptual framework for nucleoid functions. *Plant Physiol.* 158, 156–189. doi: 10.1104/pp.111.188474
- Martin, W., Rujan, T., Richly, E., Hansen, A., Cornelsen, S., Lins, T., et al. (2002). Evolutionary analysis of *Arabidopsis*, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. *Proc. Natl. Acad. Sci. U.S.A.* 99, 12246–12251. doi: 10.1073/pnas.182432999
- McElver, J., Tzafrir, I., Aux, G., Rogers, R., Ashby, C., Smith, K., et al. (2001). Insertional mutagenesis of genes required for seed development in *Arabidopsis thaliana*. *Genetics* 159, 1751–1763.
- Perry, S. E., and Wang, H. (2003). Rapid isolation of *Arabidopsis thaliana* developing embryos. *BioTechniques* 35, 278–280, 282.
- Pfalz, J., Liere, K., Kandlbinder, A., Dietz, K. J., and Oelmuller, 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. doi: 10.1105/tpc.105.036392
- Pfalz, J., and Pfannschmidt, T. (2013). Essential nucleoid proteins in early chloroplast development. *Trends Plant Sci.* 18, 186–194. doi: 10.1016/j.tplants.2012.11.003
- Pfannschmidt, T., and Liere, K. (2005). Redox regulation and modification of proteins controlling chloroplast gene expression. *Antioxid. Redox Signal.* 7, 607–618. doi: 10.1089/ars.2005.7.607
- Pfannschmidt, T., and Link, G. (1994). Separation of two classes of plastid DNA-dependent RNA polymerases that are differentially expressed in mustard (*Sinapis alba* L.) seedlings. *Plant Mol. Biol.* 25, 69–81. doi: 10.1007/BF00024199
- Pfannschmidt, T., Ogrzewalla, K., Baginsky, S., Sickmann, A., Meyer, H. E., and Link, G. (2000). The multisubunit chloroplast RNA polymerase A from mustard (*Sinapis alba* L.). Integration of a prokaryotic core into a larger complex with organelle-specific functions. *Eur. J. Biochem.* 267, 253–261. doi: 10.1046/j.1432-1327.2000.00991.x
- Puthiayavetil, S., Ibrahim, I. M., Jelicic, B., Tomasic, A., Fulgosi, H., and Allen, J. F. (2010). Transcriptional control of photosynthesis genes: the evolutionarily conserved regulatory mechanism in plastid genome function. *Genome Biol. Evol.* 2, 888–896. doi: 10.1093/gbe/evq073
- Qi, Y., Armbruster, U., Schmitz-Linneweber, C., Delannoy, E., De Longevialle, A. F., Ruhle, T., et al. (2012). *Arabidopsis* CSP41 proteins form multimeric complexes that bind and stabilize distinct plastid transcripts. *J. Exp. Bot.* 63, 1251–1270. doi: 10.1093/jxb/err347
- Reiss, T., and Link, G. (1985). Characterization of transcriptionally active DNA-protein complexes from chloroplasts and etioplasts of mustard (*Sinapis alba* L.). *Eur. J. Biochem.* 148, 207–212. doi: 10.1111/j.1432-1033.1985.tb08826.x
- Rolletschek, H., Weber, H., and Borisjuk, L. (2003). Energy status and its control on embryogenesis of legumes. Embryo photosynthesis contributes to oxygen supply and is coupled to biosynthetic fluxes. *Plant Physiol.* 132, 1196–1206. doi: 10.1104/pp.102.017376
- Ruppel, N. J., and Hangarter, R. P. (2007). Mutations in a plastid-localized elongation factor G alter early stages of plastid development in *Arabidopsis thaliana*. *BMC Plant Biol.* 7:37. doi: 10.1186/1471-2229-7-37
- Sakonju, S., Bogenhagen, D. F., and Brown, D. D. (1980). A control region in the center of the 5S RNA gene directs specific initiation of transcription: I. The 5' border of the region. *Cell* 19, 13–25. doi: 10.1016/0092-8674(80)90384-0
- Shaihkhali, J., Barajas-Lopéz, J., Östvös, K., Kremnev, D., Sánchez García, A., Srivastava, V., et al. (2012a). The CRYPTOCHROME1-dependent response

- to excess light is mediated through the transcriptional activators ZINC FIN-GER PROTEIN EXPRESSED IN INFLORESCENCE MERISTEM LIKE1 and 2 in *Arabidopsis thaliana*. *Plant Cell* 24, 3009–3025. doi: 10.1105/tpc.112.100099
- Saikhali, J., Norén, L., Barajas-López, J., Srivastava, V., König, J., Sauer, U. H., et al. (2012b). Redox-mediated mechanisms regulate DNA-binding activity of the G-group of bZIP transcription factors in *Arabidopsis*. *J. Biol. Chem.* 287, 27510–27525. doi: 10.1074/jbc.M112.361394
- Sosso, D., Canut, M., Gendrot, G., Dedieu, A., Chambrier, P., Barkan, A., et al. (2012). PPR8522 encodes a chloroplast-targeted pentatricopeptide repeat protein necessary for maize embryogenesis and vegetative development. *J. Exp. Bot.* 63, 5843–5857. doi: 10.1093/jxb/ers232
- Spencer, M. W., Casson, S. A., and Lindsey, K. (2007). Transcriptional profiling of the *Arabidopsis* embryo. *Plant Physiol.* 143, 924–940. doi: 10.1104/pp.106.087668
- Stroher, E., and Dietz, K. J. (2008). The dynamic thiol-disulphide redox proteome of the *Arabidopsis thaliana* chloroplast as revealed by differential electrophoretic mobility. *Physiol. Plant.* 133, 566–583. doi: 10.1111/j.1399-3054.2008.01103.x
- Suck, R., Zeltz, P., Falk, J., Acker, A., Kossel, H., and Krupinska, K. (1996). Transcriptionally active chromosomes (TACs) of barley chloroplasts contain the alpha-subunit of plastome-encoded RNA polymerase. *Curr. Genet.* 30, 515–521. doi: 10.1007/s002940050164
- Suswam, E. A., Li, Y. Y., Mahtani, H., and King, P. H. (2005). Novel DNA-binding properties of the RNA-binding protein TIAR. *Nucleic Acids Res.* 33, 4507–4518. doi: 10.1093/nar/gki763
- Suzuki, J. Y., Ytterberg, A. J., Beardslee, T. A., Allison, L. A., Wijk, K. J., and Maliga, P. (2004). Affinity purification of the tobacco plastid RNA polymerase and in vitro reconstitution of the holoenzyme. *Plant J.* 40, 164–172. doi: 10.1111/j.1365-313X.2004.02195.x
- Tejos, R. I., Mercado, A. V., and Meisel, L. A. (2010). Analysis of chlorophyll fluorescence reveals stage specific patterns of chloroplast-containing cells during *Arabidopsis* embryogenesis. *Biol. Res.* 43, 99–111. doi: 10.4067/S0716-97602010000100012
- Tzafrir, I., Dickerman, A., Brazhnik, O., Nguyen, Q., McElver, J., Frye, C., et al. (2003). The *Arabidopsis* SeedGenes Project. *Nucleic Acids Res.* 31, 90–93. doi: 10.1093/nar/gkg028
- Wakasugi, T., Tsudzuki, T., and Sugiura, M. (2001). The genomics of land plant chloroplasts: gene content and alteration of genomic information by RNA editing. *Photosynth. Res.* 70, 107–118. doi: 10.1023/A:1013892009589
- Williams-Carrier, R., Zoschke, R., Belcher, S., Pfalz, J., and Barkan, A. (2014). A major role for the plastid encoded RNA polymerase complex in the expression of plastid transfer RNAs. *Plant Physiol.* 164, 239–248. doi: 10.1104/pp.113.228726
- Yamaguchi, K., Beligni, M. V., Prieto, S., Haynes, P. A., McDonald, W. H., Yates, J. R. III, et al. (2003). Proteomic characterization of the *Chlamydomonas reinhardtii* chloroplast ribosome. Identification of proteins unique to the 70 S ribosome. *J. Biol. Chem.* 278, 33774–33785. doi: 10.1074/jbc.M301934200
- Zhelyazkova, P., Sharma, C. M., Forstner, K. U., Liere, K., Vogel, J., and Borner, T. (2012). The primary transcriptome of barley chloroplasts: numerous noncoding RNAs and the dominating role of the plastid-encoded RNA polymerase. *Plant Cell* 24, 123–136. doi: 10.1105/tpc.111.089441
- Zoschke, R., Liere, K., and Borner, 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. doi: 10.1111/j.1365-313X.2007.03084.x

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 14 March 2014; accepted: 19 July 2014; published online: 12 August 2014.

Citation: Kremnev D and Strand Å (2014) Plastid encoded RNA polymerase activity and expression of photosynthesis genes required for embryo and seed development in *Arabidopsis*. *Front. Plant Sci.* 5:385. doi: 10.3389/fpls.2014.00385

This article was submitted to Plant Physiology, a section of the journal *Frontiers in Plant Science*.

Copyright © 2014 Kremnev and Strand. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Establishment of the chloroplast genetic system in rice during early leaf development and at low temperatures

Kensuke Kusumi\* and Koh Iba

Department of Biology, Faculty of Sciences, Kyushu University, Fukuoka, Japan

**Edited by:**

Thomas Pfannschmidt, University  
Joseph Fourier Grenoble, France

**Reviewed by:**

Wataru Sakamoto, Okayama  
University, Japan

Ben Field, Centre National de la  
Recherche Scientifique, France

**\*Correspondence:**

Kensuke Kusumi, Department of  
Biology, Faculty of Sciences, Kyushu  
University, Fukuoka 812-8581, Japan  
e-mail: kusumi.k.239@m.kyushu-u.ac.jp

Chloroplasts are the central nodes of the metabolic network in leaf cells of higher plants, and the conversion of proplastids into chloroplasts is tightly coupled to leaf development. During early leaf development, the structure and function of the chloroplasts differ greatly from those in a mature leaf, suggesting the existence of a stage-specific mechanism regulating chloroplast development during this period. Here, we discuss the identification of the genes affected in low temperature-conditional mutants of rice (*Oryza sativa*). These genes encode factors involved in chloroplast rRNA regulation (*NUS1*), and nucleotide metabolism in mitochondria, chloroplasts, and cytosol (*V<sub>2</sub>*, *V<sub>3</sub>*, *ST1*). These genes are all preferentially expressed in the early leaf developmental stage P4, and depleting them causes altered chloroplast transcription and translation, and ultimately leaf chlorosis. Therefore, it is suggested that regulation of cellular nucleotide pools and nucleotide metabolism is indispensable for chloroplast development under low temperatures at this stage. This review summarizes the current understanding of these factors and discusses their roles in chloroplast biogenesis.

**Keywords:** chloroplast transcription, translation, *Oryza sativa*, low temperature, nucleotide metabolism

## INTRODUCTION

Low temperature is a major abiotic constraint to plant growth. In rice, two stages of development are known to be the most sensitive to low temperatures the young seedling stage and the booting stage (Kaneda and Beachell, 1974; Cruz et al., 2013). At the booting stage, pollen sterility caused by low temperatures decreases the final grain yield. At the seedling stage, low temperatures reduce germination and delay leaf emergence and greening. Leaf chlorosis and yellowing are common symptoms when a low temperature prevails during this stage (Cruz et al., 2013), which suggests the low temperature arrests chloroplast development and functioning.

The effect of a low-temperature environment on chloroplast functions has been studied extensively (Berry and Björkman, 1980; Hasanuzzaman et al., 2013). A low temperature causes swelling of the thylakoid lamellae, vesiculation of the thylakoid, and ultimately breakdown of the entire chloroplast. A low temperature also inhibits electron transport and the carbon assimilation apparatus such as the Calvin cycle, ATP synthase, and ribulose 1,5-bisphosphate carboxylase/oxygenase (RuBisCO; Demmig-Adams and Adams, 1992; Asada, 1999; Yamori et al., 2011; Hasanuzzaman et al., 2013). However, these physiological symptoms have been mainly investigated in mature leaves containing functionally established chloroplasts. The molecular mechanisms underlying early chloroplast development under low temperatures have not yet been extensively studied.

*virescent* is a chlorotic mutant of higher plants causing young leaves to have a reduced chlorophyll content, but the chlorophyll levels recover as they grow (Archer and Bonnett, 1987). In contrast with other chlorotic mutants showing lethality such as *albino*, *chlorina*, and *xantha*, the *virescent* mutants are not

terminal, and can reach maturity and produce seeds. Certain classes of the *virescent* mutations that have been reported are low-temperature conditional. They develop chlorotic leaves under low temperatures, but not under higher temperatures, suggesting a temporal aberration in a factor governing chloroplast development under low-temperature conditions. During the past decade, numerous genes responsible for *virescent* mutations have been identified in rice, and they have been shown to be involved in the chloroplast genetic system, including transcription, translation, and nucleotide metabolism. Because many of these genes are expressed temporally during early leaf development, they are probably involved in the establishment of the plastid genetic system at this phase under low temperatures. Here, we introduce four factors involved in chloroplast biogenesis under low-temperature conditions (*NUS1*, *GKpm*, *RNRS1*, and *RNRL1*) that have been identified through genetic and functional analysis of *virescent* mutants of rice (*v<sub>1</sub>*, *v<sub>2</sub>*, *v<sub>3</sub>*, and *st1*).

## TEMPERATURE-SENSITIVE PHASE OF *virescent* MUTANTS DURING EARLY LEAF DEVELOPMENT

*virescent-1*, -2, and -3 (*v<sub>1</sub>*, *v<sub>2</sub>*, *v<sub>3</sub>*) were the first *virescent* mutants reported in rice and the mutations have been used as classical genetic markers (Omura et al., 1977). They develop chlorotic leaves at a restrictive low temperature (20°C) but nearly normal green leaves at a permissive higher temperature (30°C; **Figure 1A**). They are often hard to distinguish from each other, showing similar phenotypes. An important characteristic of these *virescent* mutants is that the leaf phenotype is not influenced by growth temperature after its emergence (Iba et al., 1991). This indicates that the leaf phenotype is irreversibly determined by

the environmental temperature at a certain developmental stage before emergence. Furthermore, this phenotype can be useful for determining the temperature-sensitive period (TSP), by shifting the temperature from restrictive to permissive during leaf development, or vice versa. This technique was originally performed with *Drosophila* (Suzuki, 1970), in which the TSP for conditional mutants was limited to particular stages of development. Temperature-shift experiments showed that the TSPs of all  $v_1$ ,  $v_2$ , and  $v_3$  mutants was at stage P4 of leaf development (Iba et al., 1991).

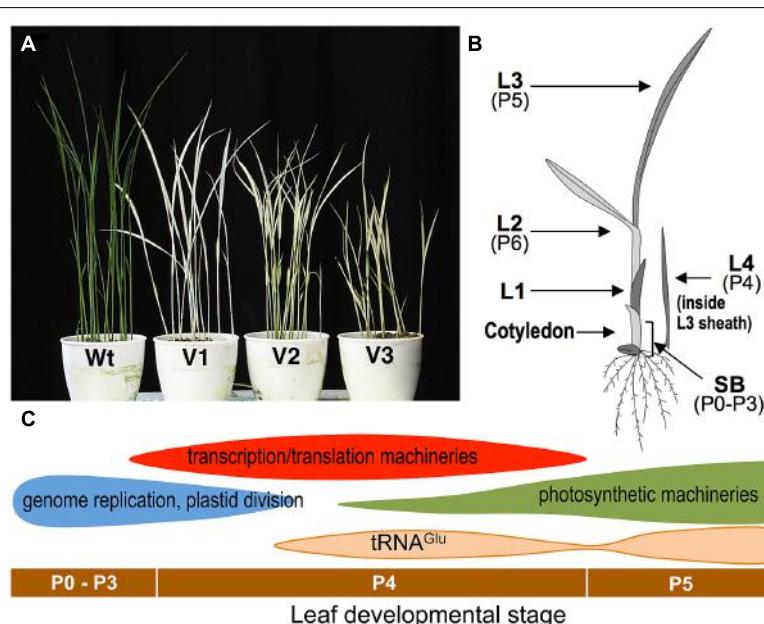
Rice has the striking feature of leaf primordia production (plastochron) that is synchronized with leaf emergence (phyllochron) in shoot development (Nemoto and Yamazaki, 1993; Itoh et al., 2005). This regularity of leaf development enables a series of successive stages to be defined, starting with P0 (leaf founder cells), through P1 (youngest primordium), P2, P3, P4, and P5, to P6 (a fully expanded leaf; **Figure 1B**). Anatomical studies have shown that in rice the P4 stage is characterized by rapid leaf blade elongation (Itoh et al., 2005). Leaves at the P4 stage have an initial length of 3–5 mm and reach a final size of about 8–10 cm. Chlorophyll concentration per unit of fresh weight is negligible in the early P4 stage, and increases to about 40% in a mature leaf (Kusumi et al., 2010a). Electron microscopic observations have indicated that chloroplasts in the leaves at the early P4 stage have a small spherical shape (below 1  $\mu\text{m}$ ) and poor internal thylakoid structures. After the mid-P4 stage, thylakoid extension and grana formation in chloroplasts has been observed within the mid-portion of the leaf (Kusumi et al., 2010b).

The process of chloroplast development is divided roughly into three steps: (i) plastid division and DNA replication; (ii)

establishment of the plastid genetic system; and (iii) activation of the photosynthetic apparatus (Jarvis and Lopez-Juez, 2013; Yagi and Shiina, 2014). These stepwise processes are partially achieved by two plastidial RNA polymerases; a nucleus-encoded phage-type RNA polymerase (NEP), and a plastid-encoded bacterial-type RNA polymerase (PEP; Hajdukiewicz et al., 1997; Yagi and Shiina, 2014). Plastid genes involved in the second step are known to be mainly transcribed by NEP, and those involved in the third step are transcribed by PEP (Yagi and Shiina, 2014). Analyses for chloroplast transcript accumulation revealed that the first step of chloroplast differentiation is likely to start in the leaves at the P0–P3 stages, and will largely finish during the early P4 stage (**Figure 1C**; Kusumi et al., 2010b). The second step occurs significantly in the leaves at mid-P4, and the decline of the second step and onset of the third step take place during late P4 (Kusumi et al., 2010b). The accumulation of tRNA<sup>Glu</sup>, a bifunctional molecule mediating the early steps of chlorophyll synthesis, and the switching of transcription from NEP to PEP (Hanaoka et al., 2005), showed two peaks (late-P4 and P5). The first activation of tRNA<sup>Glu</sup> can be related to the NEP–PEP transition. Therefore, the TSP of  $v_1$ ,  $v_2$ , and  $v_3$  mutants at the P4 stage suggests that they may be related to the establishment of the chloroplast genetic system, which is the major process occurring at this stage.

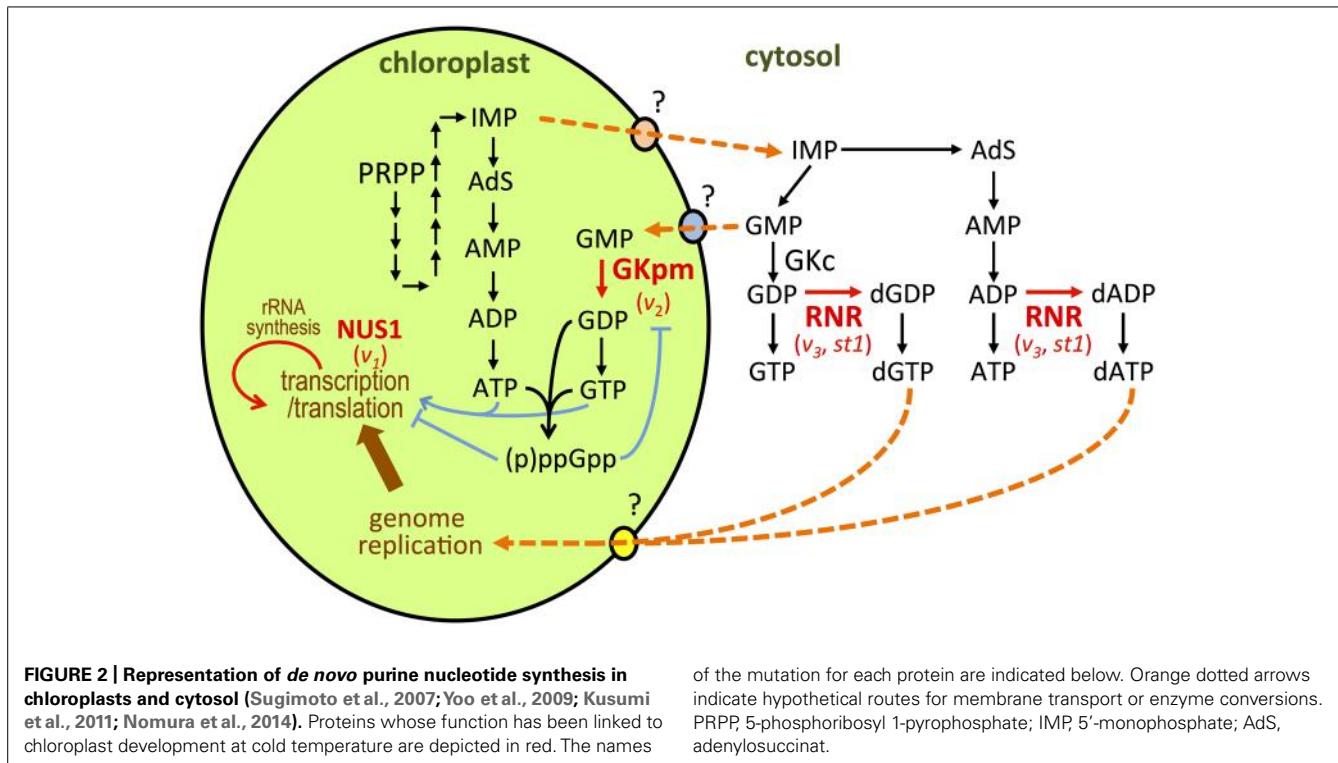
## REGULATION OF CELLULAR NUCLEOTIDE POOLS INVOLVED IN THE CHLOROPLAST DEVELOPMENT

*Virescent-2* ( $V_2$ ) was the first gene isolated from *virescent* mutants of rice (Sugimoto et al., 2007). Functional analyses showed that  $V_2$  encoded guanylate kinase (GK), a key enzyme in guanine nucleotide biosynthesis that catalyzes the conversion of GMP to



**FIGURE 1 | (A)** Phenotypes of the wild-type (Wt) and *virescent* mutants ( $v_1$ ,  $v_2$ , and  $v_3$ ) grown at a restrictive temperature (20°C). **(B)** Schematic of a rice seedling with a fully expanded third leaf. L1, L2, L3, and L4 indicate the first, second, third, and fourth leaf, respectively. Developmental stages (P0–P6) are also indicated. shoot base, SB corresponds to a 5 mm section from the bottom of the shoot and contains pre-emerged leaves at P0–P3 stages. **(C)** Plastid gene expression patterns during leaf development (Kusumi et al., 2010a). Horizontal bars indicate the leaf developmental stages.

also indicated. shoot base, SB corresponds to a 5 mm section from the bottom of the shoot and contains pre-emerged leaves at P0–P3 stages. **(C)** Plastid gene expression patterns during leaf development (Kusumi et al., 2010a). Horizontal bars indicate the leaf developmental stages.



GDP (Figure 2). In bacterial and animal species, GK is localized in the cytoplasm and participates in maintenance of the guanine nucleotide pools. Plants possess two types of GK; cytosolic GK (GKc) and plastid/mitochondrial GK (GKpm; Sugimoto et al., 2007). Analysis of RNAi knockdown plants showed that GKc is essential for the growth and development of plants, but not for chloroplast development (Sugimoto et al., 2007). V<sub>2</sub> is a single-copy gene encoding the GKpm protein. V<sub>2</sub>-encoded GKpm predominantly accumulates in developing leaves at the P0–P4 stages (Sugimoto et al., 2007), which is consistent with a temperature-shift experiment in which the V<sub>2</sub> gene product was shown to be necessary at the P4 stage. A chloroplast possesses its own nucleoside diphosphate kinase that catalyzes subsequent GDP to GTP conversion (Figure 2; Sugimoto et al., 2007; Kihara et al., 2011; Nomura et al., 2014). Therefore, GKpm can limit the GDP/GTP pool in the chloroplast. Reduction of GKpm activity will cause a shortage of the GTP necessary for the assembly and function of the plastid translation machinery. In the v<sub>2</sub> mutant, Val162 has been substituted with Ile, which caused a 20-fold reduction in specific GMP activity (Sugimoto et al., 2007), and severely suppresses chloroplast translation (Sugimoto et al., 2004). Similarly, bacterial GTPases have important roles in ribosome biogenesis and protein translation (Verstraeten et al., 2011). In *Arabidopsis* and tobacco, plastidial GTPases have been reported to be involved in chloroplast rRNA processing and ribosome biogenesis in higher plants (Jeon et al., 2014). It has also been reported that an *Arabidopsis* mutant deficient in GTP-dependent chloroplast elongation factor G developed pale cotyledons and greenish true leaves, as observed in the GKpm-deficient *Arabidopsis* (Albrecht et al., 2006; Sugimoto et al., 2007).

This phenotypic similarity suggests the involvement of GKpm in the regulation of plastid translation, via limitation of the GTP pool.

Additionally, it was recently reported that GKpm is a target of regulation by guanosine 3',5'-bis(diphosphate) (ppGpp) in chloroplasts of rice, as well as those of peas and *Arabidopsis* (Nomura et al., 2014). In bacteria, ppGpp is a key regulatory molecule that controls the stringency of responses through direct interaction with protein factors involved in gene expression such as RNA polymerase, translation factors, and DNA primase (Potrykus and Cashel, 2008; Tozawa and Nomura, 2011). In higher plants, ppGpp is synthesized in chloroplasts from GTP (GDP) and ATP (Figure 2). Major ppGpp synthase/hydrolase enzymes, named RSH (*RelA/SpoT homolog*), are localized to chloroplasts (Mizusawa et al., 2008). It has also been reported that ppGpp can negatively regulate chloroplast RNA polymerase (Sato et al., 2009) and the elongation cycle of translation (Nomura et al., 2012). This suggests that ppGpp functions as a regulatory molecule in chloroplasts, and interaction between GKpm and ppGpp may limit the GTP (and ATP) pool, which will subsequently retard chloroplast transcription and translation.

Nucleotide biosynthesis in the cytosol is also involved in the regulation of chloroplast differentiation at early leaf development under cold stress. Yoo et al. (2009) showed that the genes responsible for the v<sub>3</sub> and st1 mutants of rice encoded the large and small subunits of ribonucleotide reductase (RNR), RNRL1 and RNRS1, respectively (Yoo et al., 2009). RNR is constructed from large RNR (α) and small RNR (β) subunits, which associate to form an active heterodimer complex (α<sub>2</sub>β<sub>2</sub>) and catalyze conversion of nucleotide diphosphates (NDPs) to deoxyribonucleotide

diphosphates (dNDPs; **Figure 2**). Synthesized dNDPs are rapidly converted into dNTPs for DNA replication and repair. Therefore, the RNR activity affects the entire *de novo* nucleotide synthesis pathway *in vivo* (Elledge et al., 1992). As observed in the *v<sub>3</sub>* mutant, *st1* also caused low-temperature-dependent leaf chlorosis. RNRL1 and RNRS1 abundantly accumulated in the leaves at the P0–P4 stages, and this was enhanced by low temperature (Yoo et al., 2009). Both the *v<sub>2</sub>* and *st1* mutations caused missense mutations resulting in reduction of the first ab dimerization, which correlated with the degree of chloroplast disruption (Yoo et al., 2009). This suggests that a threshold level of RNR activity plays an important role in regulating nucleotide flow from the cytosol to chloroplasts. The involvement of cytosolic RNR in plastid nucleotide metabolism is further supported by the report that RNR deficiency causes plastid DNA degradation in pollen in *Arabidopsis* (Tang et al., 2012). Balancing chloroplast biogenesis and cell division during early leaf development would be achieved through optimization of the nucleotide pool in the cellular compartments.

### NUS1 REQUIRED FOR rRNA MATURATION AT LOW TEMPERATURES

In bacteria, synthesis of ribosomes requires a Rho-dependent anti-termination system for the efficient transcription of 16S, 23S, and 5S rRNA from *rrn* operons (Santangelo and Artsimovitch, 2011). All *rrn* operons have anti-terminator sequences in their leader and spacer regions, referred to as BoxB, BoxA, and BoxC, that allow RNA polymerase, modified with protein factors, to transcribe rRNA operons. Previously known protein factors that interact with the anti-terminator include NusA, NusB, NusE, and NusG (Santangelo and Artsimovitch, 2011). Recently, *Virescent-1* (*V<sub>1</sub>*) was identified from a *v<sub>1</sub>* mutant of rice and shown to encode a novel chloroplast RNA binding protein, named NUS1 (Kusumi et al., 2011). The C-terminal region of NUS1 has a structural similarity to the RNA-binding domain of the bacterial NusB, which is classified as alpha helical with seven helices. Accumulation of NUS1 specifically occurs in the developing leaves at the P4 stage, and is enhanced by low-temperature treatment (Kusumi et al., 2011). Although there are no regions identical to bacterial Box regions within the chloroplast *rrn* operon in rice, the gene order of 16S–23S–4.5S–5S and their coding sequences are highly conserved with those of the bacterial *rrn* operon (Bollenbach et al., 2007). RNA-immunoprecipitation and gel mobility shift assays indicated that NUS1 binds to the upstream leader region of the 16S rRNA precursor (Kusumi et al., 2011). The *v<sub>1</sub>* mutant had a nonsense mutation in the helical domain and failed to accumulate the NUS1 protein, and therefore probably represents the null phenotype. In the *v<sub>1</sub>* seedlings grown at low temperatures, the processing and accumulation of chloroplast rRNA and chloroplast translation/transcription was severely suppressed (Kusumi et al., 1997, 2011). Additionally, *Arabidopsis* seedlings deficient in a NUS1 ortholog also exhibited a similar phenotype (Kusumi et al., 2011). Therefore, NUS1 is likely to be involved in the regulation of rRNA maturation, which occurs at the P4 stage. Bacterial NusB is involved in the protein complex that interacts with RNA polymerase, nascent mRNA, and ribosomes (Santangelo and Artsimovitch, 2011; Bubunenko et al.,

2013). Recent proteomics-based techniques have allowed the identification of previously uncharacterized proteins that contribute to the chloroplast genetic system (Majeran et al., 2012; Pfalz and Pfannschmidt, 2013). Majeran et al. (2012) showed that in maize, NUS1 and other factors structurally similar to the bacterial Nus-related factors, such as NusG and Rho, were included in nucleoid-enriched fractions. Examination of the physical interactions among these proteins, and identification of other factors interacting with the NUS1 protein, will be vital for elucidating the role of NUS1 in the regulation of the chloroplast genetic system.

### CONCLUDING REMARKS

Compared with other cereals such as wheat and barley, rice is susceptible to cold stress, probably because of its tropical origin (Cruz et al., 2013). The degrees of low-temperature sensitivity and damage vary according to the growth stage. Yoshida (1981) showed that temperature sensitivity varies between stages and that rice plants have a lower threshold temperature for cold damage during the early young seedling stage (10–13°C) than during the reproductive stage (18–20°C), making them less sensitive to low temperature as young seedlings. In field conditions, sudden low-temperature phases often occur during the early seedling development in spring. Therefore, it is reasonable to infer that rice developed this mechanism to protect leaf and internal chloroplast development against low temperature-induced retardation.

It has been known that processes of chloroplast translation are sensitive to cold stress. Environmental low temperature arrests protein synthesis by causing ribosomal pausing (Grennan and Ort, 2007) or retardation of ribosomal biogenesis and RNA processing (Millerd et al., 1969; Hopkins and Elfman, 1984; Barkan, 1993). Furthermore, loss of translational factors such as ribosomal protein (Rogalski et al., 2008; Fleischmann et al., 2011; Ehrnthal et al., 2014; Song et al., 2014), translation elongation factor (Liu et al., 2010), rRNA methylase (Tokuhisa et al., 1998) and the RNA binding protein required for RNA processing (Kupsch et al., 2012) leads to sensitivity to low temperatures. These reports suggest the existence of a particular mechanism that protects chloroplast translation against cold stress, which can be expected to be associated with *NUS1*, *V<sub>2</sub>*, *V<sub>3</sub>*, and *ST1*.

The observed involvement of control of translation and nucleotide metabolism in low-temperature tolerance/adaptation has also been reported in bacteria. For example, in *Escherichia coli*, a temperature downshift hampers ribosome function, and ribosomes change their composition to function properly (Akanuma et al., 2012). It has also been reported that cold-shock proteins are often induced not only by low temperatures but also by translational inhibitors, such as chloramphenicol and tetracycline (Weber and Marahiel, 2003). Therefore, a reduction in translational capacity may be interpreted as a cellular signal triggering the cold adaptation response. Furthermore, analyses of bacterial mutants deficient in ppGpp synthesis showed that artificially induced high levels of ppGpp diminish the expression of cold-shock proteins, while low levels increase their production (Potrykus and Cashel, 2008). ppGpp synthesis is triggered by occupation of the ribosomal A-site by an uncharged tRNA (Potrykus and Cashel, 2008). Considering the hampered ribosomal function at low temperature,

it is possible that a decrease in cellular ppGpp levels following a temperature downshift plays a physiological role in the regulation of gene expression and adaptation to growth at low temperature. The bacterial NusB protein has also been reported to be involved in cold tolerance. Cells containing a disrupted *nusB* gene are viable under standard growth conditions, but are cold sensitive (Quan et al., 2005). They are defective in rRNA synthesis and have a decreased peptide elongation rate at low temperatures. NusA, another host factor of the Nus complex, is also induced by cold treatment (Phadtare and Severinov, 2010), suggesting the importance of Nus and the anti-termination system in the cold response in bacteria. These similar properties between the chloroplast and the bacterial low-temperature response imply that higher plants have taken over the bacterial protective system in response to low temperature.

Recently, several other genes have been isolated from low temperature-conditional, chloroplast-deficient mutants of rice, such as *OsV4* (*virescent 4*), *wlp1* (*white leaf and panicles 1*), and *tcd9* (*thermo-sensitive chloroplast development 9*; Gong et al., 2014; Jiang et al., 2014; Song et al., 2014). The corresponding genes in *OsV4*, *wlp1*, and *tcd9* mutants encode plastidial pentatricopeptide repeat (PPR) protein, plastid ribosomal protein L13, and a sub-unit of chaperonin 60 (CP60 $\alpha$ ) required for chloroplast division, respectively. Similarly to *NUS1*, *V<sub>2</sub>*, *V<sub>3</sub>*, and *ST1*, their functions are speculated to be involved in early chloroplast development at low temperatures (Gong et al., 2014; Jiang et al., 2014; Song et al., 2014). It is possible that these factors are involved in a closely related mechanism to chloroplast protein expression and assembly, which is required at low temperatures, but not essential for chloroplast development during early leaf development at higher temperatures. Therefore, the maintenance of the developing plastid genetic system will be crucial for tolerance of cold at the seedling stage in rice.

## ACKNOWLEDGMENTS

We are grateful to Dr. Yuzuru Tozawa (Ehime University) for helpful discussion. This work was supported by the Ministry of Education, Culture, Sports, Science, and Technology of Japan (Nos. 21114002, 22570045).

## REFERENCES

- Akanuma, G., Nanamiya, H., Natori, Y., Yano, K., Suzuki, S., Omata, S., et al. (2012). Inactivation of ribosomal protein genes in *Bacillus subtilis* reveals importance of each ribosomal protein for cell proliferation and cell differentiation. *J. Bacteriol.* 194, 6282–6291. doi: 10.1128/JB.01544-12
- Albrecht, V., Ingenfeld, A., and Apel, K. (2006). Characterization of the *snowly cotyledon 1* mutant of *Arabidopsis thaliana*: the impact of chloroplast elongation factor G on chloroplast development and plant vitality. *Plant Mol. Biol.* 60, 507–518. doi: 10.1007/s11103-005-4921-0
- Archer, E. K., and Bonnett, H. T. (1987). Characterization of a virescent chloroplast mutant of tobacco. *Plant Physiol.* 83, 920–925. doi: 10.1104/pp.83.4.920
- Asada, K. (1999). The water-water cycle in chloroplasts: scavenging of active oxygens and dissipation of excess photons. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 50, 601–639. doi: 10.1146/annurev.arplant.50.1.601
- Barkan, A. (1993). Nuclear mutants of maize with defects in chloroplast polysome assembly have altered chloroplast RNA metabolism. *Plant Cell* 5, 389–402. doi: 10.1105/tpc.5.4.389
- Berry, J., and Björkman, O. (1980). Photosynthetic response and adaptation to temperature in higher plants. *Annu. Rev. Plant Physiol.* 31, 491–543. doi: 10.1146/annurev.pp.31.060180.002423
- Bollenbach, T. J., Schuster, G., Portnoy, V., and Stern, D. B. (2007). Processing, degradation and polyadenylation of chloroplast transcripts. *Topics Curr. Genet.* 19, 175–211. doi: 10.1007/4735\_2007\_0235
- Bubunenko, M., Court, D. L., Al Refaii, A., Saxena, S., Korepanov, A., Friedman, D. I., et al. (2013). Nus transcription elongation factors and RNase III modulate small ribosome subunit biogenesis in *Escherichia coli*. *Mol. Microbiol.* 87, 382–393. doi: 10.1111/mmi.12105
- Cruz, R. P. D., Sperotto, R. A., Cargnelutti, D., Adamski, J. M., De Freitas, T., and Fett, J. P. (2013). Avoiding damage and achieving cold tolerance in rice plants. *Food Energy Secur.* 2, 96–119. doi: 10.1002/fes.325
- Demmig-Adams, B., and Adams, W. W. (1992). Photoprotection and other responses of plants to high light stress. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 43, 599–626. doi: 10.1146/annurev.pp.43.060192.003123
- Ehrnthal, M., Scharff, L. B., Fleischmann, T. T., Hasse, C., Ruf, S., and Bock, R. (2014). Synthetic lethality in the tobacco plastid ribosome and its rescue at elevated growth temperatures. *Plant Cell* 26, 765–776. doi: 10.1105/tpc.114.123240
- Elledge, S. J., Zhou, Z., and Allen, J. B. (1992). Ribonucleotide reductase: regulation, regulation, regulation. *Trends Biochem. Sci.* 17, 119–123. doi: 10.1016/0968-0004(92)90249-9
- Fleischmann, T. T., Scharff, L. B., Alkatib, S., Hasdorff, S., Schottler, M. A., and Bock, R. (2011). Nonessential plastid-encoded ribosomal proteins in tobacco: a developmental role for plastid translation and implications for reductive genome evolution. *Plant Cell* 23, 3137–3155. doi: 10.1105/tpc.111.088906
- Gong, X., Su, Q., Lin, D., Jiang, Q., Xu, J., Zhang, J., et al. (2014). The rice *OsV4* encoding a novel pentatricopeptide repeat protein is required for chloroplast development during the early leaf stage under cold stress. *J. Integr. Plant Biol.* 56, 400–410. doi: 10.1111/jipb.12138
- Grennan, A. K., and Ort, D. R. (2007). Cool temperatures interfere with D1 synthesis in tomato by causing ribosomal pausing. *Photosynth. Res.* 94, 375–385. doi: 10.1007/s11120-007-9169-x
- Hajdukiewicz, P. T. J., Allison, L. A., and Maliga, P. (1997). The two RNA polymerases encoded by the nuclear and the plastid compartments transcribe distinct groups of genes in tobacco plastids. *EMBO J.* 16, 4041–4048. doi: 10.1093/emboj/16.13.4041
- Hanaoka, M., Kanamaru, K., Fujiwara, M., Takahashi, H., and Tanaka, K. (2005). Glutamyl-tRNA mediates a switch in RNA polymerase use during chloroplast biogenesis. *EMBO Rep.* 6, 545–550. doi: 10.1038/sj.embo.7400411
- Hasanuzzaman, M., Nahar, K., and Fujita, M. (2013). “Extreme temperature responses, oxidative stress and antioxidant defense in plants,” in *Abiotic Stress – Plant Responses and Applications in Agriculture*, eds K. Vahdati and C. Leslie. (Croatia: InTech Open Access Publisher), 169–205. doi: 10.5772/54833
- Hopkins, W. G., and Elfrman, B. (1984). Temperature-induced chloroplast ribosome deficiency in virescent maize. *J. Hered.* 75, 207–211.
- Iba, K., Takamya, K., Toh, Y., Satoh, H., and Nishimura, M. (1991). Formation of functionally active chloroplast is determined at a limited stage of leaf development in *virescent* mutants of rice. *Dev. Genet.* 12, 342–348. doi: 10.1002/dvg.1020120503
- Itoh, J., Nonomura, K., Ikeda, K., Yamaki, S., Inukai, Y., Yamagishi, H., et al. (2005). Rice plant development: from zygote to spikelet. *Plant Cell Physiol.* 46, 23–47. doi: 10.1093/pcp/pci501
- Jarvis, P., and Lopez-Juez, E. (2013). Biogenesis and homeostasis of chloroplasts and other plastids. *Nat. Rev. Mol. Cell Biol.* 14, 787–802. doi: 10.1038/nrm3702
- Jeon, Y., Ahn, C. S., Jung, H. J., Kang, H., Park, G. T., Choi, Y., et al. (2014). DER containing two consecutive GTP-binding domains plays an essential role in chloroplast ribosomal RNA processing and ribosome biogenesis in higher plants. *J. Exp. Bot.* 65, 117–130. doi: 10.1093/jxb/ert360
- Jiang, Q., Mei, J., Gong, X. D., Xu, J. L., Zhang, J. H., Teng, S., et al. (2014). Importance of the rice *TCD9* encoding alpha subunit of chaperonin protein 60 (Cpn60 $\alpha$ ) for the chloroplast development during the early leaf stage. *Plant Sci.* 215–216, 172–179. doi: 10.1016/j.plantsci.2013.11.003
- Kaneda, C., and Beachell, H. M. (1974). Response of indica-japonica rice hybrids to low temperatures. *SABRAO J.* 6, 17–32.
- Kihara, A., Saburi, W., Wakuta, S., Kim, M. H., Hamada, S., Ito, H., et al. (2011). Physiological and biochemical characterization of three nucleoside diphosphate kinase isozymes from rice (*Oryza sativa* L.). *Biosci. Biotechnol. Biochem.* 75, 1740–1745. doi: 10.1271/bbb.110285

- Kupsch, C., Ruwe, H., Gusewski, S., Tillich, M., Small, I., and Schmitz-Linneweber, C. (2012). *Arabidopsis* chloroplast RNA binding proteins CP31A and CP29A associate with large transcript pools and confer cold stress tolerance by influencing multiple chloroplast RNA processing steps. *Plant Cell* 24, 4266–4280. doi: 10.1105/tpc.112.103002
- Kusumi, K., Chono, Y., Gotoh, E., Tsuyama, M., and Iba, K. (2010a). Chloroplast biogenesis during the early stage of leaf development in rice. *Plant Biotechnol.* 27, 85–90. doi: 10.5511/plantbiotechnology.27.85
- Kusumi, K., Hirotsuka, S., Shimada, H., Chono, Y., Matsuda, O., and Iba, K. (2010b). Contribution of chloroplast biogenesis to carbon-nitrogen balance during early leaf development in rice. *J. Plant Res.* 123, 617–622. doi: 10.1007/s10265-009-0277-x
- Kusumi, K., Mizutani, A., Nishimura, M., and Iba, K. (1997). A *virescent* gene *V1* determines the expression timing of plastid genes for transcription/translation apparatus during early leaf development in rice. *Plant J.* 12, 1241–1250. doi: 10.1046/j.1365-313x.1997.12061241.x
- Kusumi, K., Sakata, C., Nakamura, T., Kawasaki, S., Yoshimura, A., and Iba, K. (2011). A plastid protein NUS1 is essential for build-up of the genetic system for early chloroplast development under cold stress conditions. *Plant J.* 68, 1039–1050. doi: 10.1111/j.1365-313X.2011.04755.x
- Liu, X., Rodermeier, S. R., and Yu, F. (2010). A *var2* leaf variegation suppressor locus, *SUPPRESSOR OF VARIEGATION3*, encodes a putative chloroplast translation elongation factor that is important for chloroplast development in the cold. *BMC Plant Biol.* 10:287. doi: 10.1186/1471-2229-10-287
- Majeran, W., Friso, G., Asakura, Y., Qu, X., Huang, M., Ponnala, L., et al. (2012). Nucleoid-enriched proteomes in developing plastids and chloroplasts from maize leaves: a new conceptual framework for nucleoid functions. *Plant Physiol.* 158, 156–189. doi: 10.1104/pp.111.188474
- Millerd, A., Goodchild, D. J., and Spencer, D. (1969). Studies on a maize mutant sensitive to low temperature II. Chloroplast structure, development, and physiology. *Plant Physiol.* 44, 567–583. doi: 10.1104/pp.44.4.567
- Mizusawa, K., Masuda, S., and Ohta, H. (2008). Expression profiling of four RelA/SpoT-like proteins, homologues of bacterial stringent factors, in *Arabidopsis thaliana*. *Planta* 228, 553–562. doi: 10.1007/s00425-008-0758-5
- Nemoto, K., and Yamazaki, K. (1993). "Correlative development of vegetative organs," in *Science of the Rice Plant*, eds T. Matsuo and K. Hoshikawa (Tokyo: Food and Agriculture Policy Research Center), 625–627.
- Nomura, Y., Izumi, A., Fukunaga, Y., Kusumi, K., Iba, K., Watanebe, S., et al. (2014). Diversity in guanosine 3',5'-bis(diphosphate) (ppGpp) sensitivity among guanylate kinases of bacteria and plants. *J. Biol. Chem.* 289, 15631–15641. doi: 10.1074/jbc.M113.534768
- Nomura, Y., Takabayashi, T., Kuroda, H., Yukawa, Y., Sattasuk, K., Akita, M., et al. (2012). ppGpp inhibits peptide elongation cycle of chloroplast translation system *in vitro*. *Plant Mol. Biol.* 78, 185–196. doi: 10.1007/s11103-011-9858-x
- Omura, T., Satoh, H., Aiga, I., and Nagao, N. (1977). Studies on the character manifestation in chlorophyll mutants of rice: I. *Virescent* mutants sensitive to low temperature. *J. Fac. Agr. Kyushu Univ.* 21, 129–140.
- Pfälz, J., and Pfannschmidt, T. (2013). Essential nucleoid proteins in early chloroplast development. *Trends Plant Sci.* 18, 186–194. doi: 10.1016/j.tplants.2012.11.003
- Phadare, S., and Severinov, K. (2010). RNA remodeling and gene regulation by cold shock proteins. *RNA Biol.* 7, 788–795. doi: 10.4161/rna.7.6.13482
- Potrykus, K., and Cashel, M. (2008). (p)ppGpp: still magical? *Annu. Rev. Microbiol.* 62, 35–51. doi: 10.1146/annurev.micro.62.081307.162903
- Quan, S., Zhang, N., French, S., and Squires, C. L. (2005). Transcriptional polarity in rRNA operons of *Escherichia coli* *nusA* and *nusB* mutant strains. *J. Bacteriol.* 187, 1632–1638. doi: 10.1128/JB.187.5.1632-1638.2005
- Rogalski, M., Schottler, M. A., Thiele, W., Schulze, W. X., and Bock, R. (2008). Rpl33, a nonessential plastid-encoded ribosomal protein in tobacco, is required under cold stress conditions. *Plant Cell* 20, 2221–2237. doi: 10.1105/tpc.108.060392
- Santangelo, T. J., and Artsimovitch, I. (2011). Termination and antitermination: RNA polymerase runs a stop sign. *Nat. Rev. Microbiol.* 9, 319–329. doi: 10.1038/nrmicro2560
- Sato, M., Takahashi, K., Ochiai, Y., Hosaka, T., Ochi, K., and Nabeta, K. (2009). Bacterial alarmone, guanosine 5'-diphosphate 3'-diphosphate (ppGpp), predominantly binds the  $\beta'$  subunit of plastid-encoded plastid RNA polymerase in chloroplasts. *ChemBioChem* 10, 1227–1233. doi: 10.1002/cbic.200800737
- Song, J., Wei, X., Shao, G., Sheng, Z., Chen, D., Liu, C., et al. (2014). The rice nuclear gene *WLP1* encoding a chloroplast ribosome L13 protein is needed for chloroplast development in rice grown under low temperature conditions. *Plant Mol. Biol.* 84, 301–314. doi: 10.1007/s11103-013-0134-0
- Sugimoto, H., Kusumi, K., Noguchi, K., Yano, M., Yoshimura, A., and Iba, K. (2007). The rice nuclear gene, *VIRESCENT2*, is essential for chloroplast development and encodes a novel type of guanylate kinase targeted to plastids and mitochondria. *Plant J.* 52, 512–527. doi: 10.1111/j.1365-313X.2007.03251.x
- Sugimoto, H., Kusumi, K., Tozawa, Y., Yazaki, J., Kishimoto, N., Kikuchi, S., et al. (2004). The *virescent-2* mutation inhibits translation of plastid transcripts for the plastid genetic system at an early stage of chloroplast differentiation. *Plant Cell Physiol.* 45, 985–996. doi: 10.1093/pcp/pch111
- Suzuki, D. T. (1970). Temperature-sensitive mutations in *Drosophila melanogaster*. *Science* 170, 695–706. doi: 10.1126/science.170.3959.695
- Tang, L. Y., Matsushima, R., and Sakamoto, W. (2012). Mutations defective in ribonucleotide reductase activity interfere with pollen plastid DNA degradation mediated by DPD1 exonuclease. *Plant J.* 70, 637–649. doi: 10.1111/j.1365-313X.2012.04904.x
- Tokuhisa, J. G., Vijayan, P., Feldmann, K. A., and Browse, J. A. (1998). Chloroplast development at low temperatures requires a homolog of DIM1, a yeast gene encoding the 18S rRNA dimethylase. *Plant Cell* 10, 699–711. doi: 10.1105/tpc.10.5.699
- Tozawa, Y., and Nomura, Y. (2011). Signalling by the global regulatory molecule ppGpp in bacteria and chloroplasts of land plants. *Plant Biol.* 13, 699–709. doi: 10.1111/j.1438-8677.2011.00484.x
- Verstraeten, N., Fauvert, M., Versee, W., and Michiels, J. (2011). The universally conserved prokaryotic GTPases. *Microbiol. Mol. Biol. Rev.* 75, 507–542. doi: 10.1128/MMBR.00009-11
- Weber, M. H., and Marahiel, M. A. (2003). Bacterial cold shock responses. *Sci. Prog.* 86, 9–75. doi: 10.3184/003685003783238707
- Yagi, Y., and Shiina, T. (2014). Recent advances in the study of chloroplast gene expression and its evolution. *Front. Plant Sci.* 5:61. doi: 10.3389/fpls.2014.00061
- Yamori, W., Sakata, N., Suzuki, Y., Shikanai, T., and Makino, A. (2011). Cyclic electron flow around photosystem I via chloroplast NAD(P)H dehydrogenase (NDH) complex performs a significant physiological role during photosynthesis and plant growth at low temperature in rice. *Plant J.* 68, 966–976. doi: 10.1111/j.1365-313X.2011.04747.x
- Yoo, S.-C., Cho, S.-H., Sugimoto, H., Li, J., Kusumi, K., Koh, H.-J., et al. (2009). Rice *Virescent-3* and *Stripe-1* encoding the large and small subunits of ribonucleotide reductase are required for chloroplast biogenesis during early leaf development. *Plant Physiol.* 150, 388–401. doi: 10.1104/pp.109.136648
- Yoshida, S. (1981). "Growth and development of the rice plant," in *Fundamentals of Rice Crop Science*. (Los Baños: Philippines: International Rice Research Institute), 1–63.
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Received:** 17 May 2014; **accepted:** 20 July 2014; **published online:** 11 August 2014.
- Citation:** Kusumi K and Iba K (2014) Establishment of the chloroplast genetic system in rice during early leaf development and at low temperatures. *Front. Plant Sci.* 5:386. doi: 10.3389/fpls.2014.00386
- This article was submitted to Plant Physiology, a section of the Journal Frontiers in Plant Science.
- Copyright © 2014 Kusumi and Iba. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# The nucleoid as a site of rRNA processing and ribosome assembly

Alexandra-Viola Bohne \*

Department of Molecular Plant Sciences, Ludwig-Maximilians-University Munich, Planegg-Martinsried, Germany

\*Correspondence: alexandra.bohne@lmu.de

**Edited by:**

Thomas Pfannschmidt, University Joseph Fourier Grenoble, France

**Reviewed by:**

Alice Barkan, University of Oregon, USA

Katharine A. Howell, The University of Western Australia, Australia

**Keywords:** rRNA processing, ribosome assembly, nucleoid, plastid, mitochondria

Protein biosynthesis is one of the key elements of gene expression in living cells. All proteins are synthesized on ribonucleoprotein complexes, the ribosomes. In bacteria and their derivatives—eukaryotic mitochondria and plastids—ribosomes consist of a small and a large subunit which together comprise more than 50 ribosomal proteins and usually two to four ribosomal RNAs (rRNAs). Synthesis of the required rRNAs and proteins, and their correct folding, maturation/modification and assembly into functional particles are highly coordinated. However, whereas ribosomal composition and many mechanistic aspects of their biogenesis are well understood, little is known about the spatial organization of the procedure in bacteria and organelles. In eukaryotes, the individual processes involved occur in defined regions of the cell: ribosomal proteins are synthesized in the cytosol, but most rRNAs are transcribed, processed and modified in the nucleolus, a distinct subnuclear compartment (Lafontaine and Tollervey, 2001; Boisvert et al., 2007). Only the transcription of the small 5S rRNA occurs in the nucleoplasm. After their synthesis, ribosomal proteins and assembly factors are imported into the nucleus, where they are combined with the appropriate rRNAs. Subsequently, small and large ribosomal subunits are exported into the cytosol, where they pair up to form functional ribosomes.

This Opinion Article focuses on recent findings which support the idea that, as in eukaryotes, ribosomal biogenesis in bacteria, mitochondria, and plastids is spatially organized. In these systems, there is

growing evidence that rRNA processing and ribosome assembly most likely take place in association with the nucleoid.

## RNA PROCESSING, MATURATION AND RIBOSOME ASSEMBLY IN BACTERIA

The processes of rRNA maturation and ribosome assembly are probably best understood in bacteria (reviewed in Kaczanowska and Rydén-Aulin, 2007; Shajani et al., 2011). In *E. coli*, the small 30S ribosomal subunit contains 21 ribosomal proteins and a 16S rRNA, while the large 50S subunit consists of 33 proteins and two rRNAs, the 23S and 5S rRNAs (reviewed in Melnikov et al., 2012). All rRNAs are encoded in a polycistronic gene cluster and transcribed as a single precursor, which undergoes extensive processing and maturation to generate the mature rRNAs (reviewed in Deutscher, 2009; Shajani et al., 2011). The processing reactions are carried out by the exo- or endonucleolytic activities of at least five ribonucleases (RNases), including RNases III, E, G, T, and YbeY (reviewed in Deutscher, 2009; Davies et al., 2010). Furthermore, the 23S and 16S rRNAs undergo methylations and pseudouridylation at several positions, which are assumed to influence ribosomal structure and function (reviewed in Shajani et al., 2011).

Many of these maturation steps are coupled and take place on nascent rRNAs. The final rRNA folding pattern is probably determined by an ordered sequence of interactions with ribosomal proteins and assembly factors, which induce conformational changes and stabilize the proper

rRNA structures (reviewed in Shajani et al., 2011).

It was long believed that DNA-associated bacterial ribosomes translate nascent mRNAs while these are being synthesized by the RNA polymerase (co-transcriptional translation). Striking evidence for this hypothesis came from early electron microscopy studies showing ribosome arrays (polysomes) attached to mRNA strands that were being transcribed from DNA by multiple RNA polymerase (RNAP) molecules (Miller et al., 1970). However, recent research suggests that, at least in *E. coli* and *B. subtilis*, most translation is not coupled to transcription, and that cotranscriptional translation may be limited to mRNAs encoding membrane proteins. By tracking the distribution of RNA polymerases (as markers for the nucleoid DNA) and ribosomes by means of fluorescence microscopy, a clear segregation of the nucleoid from ribosome-rich regions of the cytoplasm was observed (Lewis et al., 2000; Bakshi et al., 2012; Chai et al., 2014). Bakshi et al. (2012) found approximately 85% of the ribosomes in the ribosome-rich regions, while only 10 to 15% were detected in close proximity to the nucleoid. The majority fraction probably comprises actively translating “protein factories.” The nucleoid-associated particles are thought to be in various stages of assembly, as several rRNA maturation steps occur in a cotranscriptional and assembly-assisted manner (reviewed in Shajani et al., 2011).

Further data support the identification of the nucleoid as the site of early rRNA processing events. RNase III, which cotranscriptionally cleaves the primary

rRNA transcript to yield 16S, 23S, and 5S precursors, has been shown to be required for localization of the *pre-16S* rRNA 5' leader region to the nucleoid (Malagon, 2013). In RNA-FISH assays for single-cell visualization, Malagon (2013) found that the nucleoid localization of the 16S 5' leader was dependent on the catalytic activity of RNase III, which provides indirect evidence for the presence of the enzyme in this region. Based on evidence for interplay between the Nus transcription elongation factors and RNase III in the modulation of pre-rRNA biogenesis (Bubunenko et al., 2013), Malagon further speculated that Nus proteins might serve to localize pre-rRNAs to the nucleoid.

### LOCALIZED rRNA PROCESSING AND RIBOSOME ASSEMBLY IN ORGANELLES

Eukaryotic plastids and mitochondria are descended from endosymbiotically acquired bacteria, and have retained their own genomes and machineries for gene expression during evolution. It is therefore not surprising that many aspects of these genetic systems, including genome organization and ribosome structure, resemble those of bacterial rather than eukaryotic systems. Similarly, evidence for sublocalization of rRNA processing and ribosome assembly events to the organellar nucleoids is now emerging.

### MITOCHONDRIA

Evidence for rRNA maturation and ribosome assembly in association with the mitochondrial nucleoid mainly derives from studies on mammalian mitochondria. Affinity purification of established protein components of the mitochondrial nucleoid results in copurification of ribosomal proteins and the ribosome assembly factor ERAL1 (He et al., 2012b). ERAL1, a GTP-binding protein with RNA-binding activity, which had previously been linked with a number of nucleoid proteins, interacts both with proteins of the small mitoribosomal subunit and its rRNA component (Dennerlein et al., 2010; Uchiumi et al., 2010). ERAL1 was proposed to act as a chaperone for the small ribosomal RNA, as its depletion leads to destabilization of the small subunit (Dennerlein et al., 2010).

Similarly, another mitoribosome assembly factor, the human GTPase NOA1

(C4orf14), a homolog of the plant Rif1 protein, has been found together with the small ribosomal subunit and translation factors in affinity-purified nucleoids (Flores-Pérez et al., 2008; He et al., 2012a). This finding prompted the suggestion that assembly of the small subunit in the mitochondrial nucleoid enables the direct transfer of newly transcribed mRNAs to the ribosome.

Like bacterial rRNAs, mammalian mitochondrial rRNAs undergo site-specific methylations, although to a lesser extent (reviewed in Rorbach and Minczuk, 2012). The enzymes responsible belong to a family of rRNA methyltransferases, some of whose members have been reported to modify nascent rRNAs in association with the nucleoid (Lee et al., 2013). A more detailed analysis of one of these methyltransferases (named RNMTL1) revealed an additional interaction between this protein and the large ribosomal subunit, suggesting that assembly of mitochondrial ribosomes begins before rRNA transcription is complete (Lee et al., 2013).

Moreover, a very recent study by Bogenhagen et al. (2014) provides evidence that mammalian mitochondrial RNA processing enzymes, like RNase P and ELAC2 (tRNaseZ<sub>L</sub>), as well as a number of nascent mitochondrial ribosomal proteins associate with nucleoids to initiate RNA processing and ribosome assembly. The authors therefore propose the mtDNA nucleoid as a critical control center for mitochondrial biogenesis.

### PLASTIDS

An extensive body of evidence for localized ribosome biogenesis comes from plastids. Besides numerous ribosomal proteins, the majority of proteins with known roles in ribosome biogenesis have been identified in a comprehensive proteomic analysis of the maize chloroplast nucleoid (Majeran et al., 2012; reviewed in Germain et al., 2013). Many of these nucleoid-enriched ribosome biogenesis factors function in rRNA processing, maturation and modification, as well as in ribosome assembly (summarized in Table 1).

Among the processing and splicing factors identified are many plant homologs of bacterial RNases that have been reported to be responsible for exo- and endonucleolytic cleavage of the large rRNA

precursor and maturation of cotranscribed tRNAs (reviewed in Stoppel and Meurer, 2012; Germain et al., 2013). Most of the nucleoid-enriched proteins involved in ribosome assembly and rRNA modification are classified as either GTPases, RNA helicases or rRNA methylases (Table 1).

Two of the nucleoid-enriched plastid proteins listed in Table 1 are homologous to the mitochondrial enzymes ERAL1 and NOA1, for which a nucleoid localization has also been reported (see above). For several others, links with the nucleoid are further supported by alternative approaches as exemplified in the following.

An immunological analysis of maize chloroplast subfractions revealed that the DEAD-box RNA helicase RH3, thought to function in assembly of the 50S subunit, localizes to the chloroplast stroma and thylakoids, as well as to nucleoids (Asakura et al., 2012). For two recently characterized plant proteins required for ribosome biogenesis, DER and RAP, cytological evidence for nucleoid localization comes from GFP fusion experiments (Jeon et al., 2014; Kleinknecht et al., 2014). DER is a Double Era-like GTPase, whose bacterial homolog (also known as EngA) acts as a ribosome assembly factor in *E. coli*, was found to bind to the 50S ribosomal subunit and to play a role in pre-rRNA processing in tobacco (Hwang and Inouye, 2006; Jeon et al., 2014). The *Arabidopsis* protein RAP is a member of the Octotricopeptide Repeat (OPR) protein family, and binds to the 5' leader sequence of the 16S rRNA precursor (Kleinknecht et al., 2014). Depletion of RAP specifically affected the trimming/processing of the chloroplast 16S rRNA precursor, which supports the identification of the nucleoid as the site of 16S rRNA processing in chloroplasts. Preliminary data imply that RAP has no intrinsic RNase activity, and might influence 16S rRNA maturation by conferring sequence specificity on an RNase or by modulating RNA secondary structures to control the accessibility of an RNase recognition site within the 16S precursor (Kleinknecht et al., unpublished results).

However, in the case of many chloroplast proteins reported to be involved in various steps of ribosome biogenesis, GFP-based fusion studies provide no

**Table 1 | Nucleoid-localized proteins with proposed functions in rRNA processing, maturation and ribosome assembly in plant chloroplasts.**

| Protein  | Protein family/domain                                  | Proposed function in ribosome biogenesis                              | Maize identifier                       | Arabidopsis identifier | References <sup>a</sup>  |
|--|--|---|--|------------------------|--|
| <b>PROCESSING/MATURATION OF rRNA PRECURSOR TRANSCRIPTS</b> |  |   |  |                        |  |
| BPG2   | YqeH -GTPase   | Pre-rRNA processing   | GRMZM2G035042_P01                      | AT3G57180              | Komatsu et al., 2010; Kim et al., 2012                               |
| CSP41B   | Endonuclease; Rossmann fold<br>Epimerase/dehydrogenase | Processing of 23S 1st hidden break                                    | GRMZM2G080546_P02                      | AT1G09340              | Beligni and Mayfield, 2008; Bollenbach et al., 2009; Qi et al., 2012 |
| PNPase   | 3' to 5' exonuclease                                   | 23S rRNA 3' end maturation  | GRMZM2G377761_P01                      | AT3G03710              | Walter et al., 2002  |
| PRBP   | Similarity to DEAD box helicases                       | Processing of 4.5S precursors   | GRMZM2G436328_P01                      | AT2G37920              | Park et al., 2011  |
| RAP*   | Octotricopeptide repeat (OPR) protein, RNA binding     | Processing of 16S rRNA precursors                                     | GRMZM2G050845_P01                      | AT2G31890              | Kleinknecht et al., 2014   |
| RBF1   | Ribosome-binding factor A family protein               | Processing of 16S rRNA precursors                                     | GRMZM2G115156_P01                      | AT4G34730              | Fristedt et al., 2014  |
| RH39   | DEAD box RNA helicase                                  | Processing of 23S 2nd hidden break                                    | GRMZM2G175867_P02                      | AT4G09730              | Nishimura et al., 2010   |
| RHON1  | Rho transcription termination domain                   | Processing of 23S-4.5S precursors                                     | GRMZM2G168335_P01                      | AT1G06190              | Stoppel et al., 2012   |
| RimM   | RimM family  | 16S processing, ribosome assembly (in prokaryotes)                    | GRMZM2G165694_P01                      | AT5G46420              | Bylund et al., 1998; Guo et al., 2013                                |
| RNase E  | Endonuclease   | Processing of 23S-4.5S precursor                                      | GRMZM2G328309_P01                      | AT2G04270              | Stoppel et al., 2012   |
| RNase J  | 5' to 3' exonuclease and endonuclease                  | 5' end maturation of 16S and 23S rRNA precursors                      | GRMZM2G103315_P01                      | AT5G63420              | Bollenbach et al., 2005; Sharwood et al., 2011                       |
| RNase P* (PRORP1/2)  | Endonuclease   | 5' end maturation of tRNA precursors                                  | GRMZM2G002642_P01<br>GRMZM2G418206_P01 | AT2G32230              | Gutmann et al., 2012; Bogenhagen et al., 2014                        |
| RNase Z*   | Endonuclease   | 3' end maturation of tRNA precursors                                  | GRMZM2G125844_P01                      | AT2G04530              | Canino et al., 2009; Bogenhagen et al., 2014                         |
| SVR1   | Pseudouridine synthase                                 | rRNA processing   | GRMZM2G315806_P01                      | AT2G39140              | Yu et al., 2008  |
| <b>rRNA MODIFICATION/RIBOSOME ASSEMBLY</b>                 |  |   |  |                        |  |
| DER* (EngA)  | GTPase   | rRNA processing/ ribosome biogenesis                                  | GRMZM2G302233_P01                      | AT3G12080              | Jeon et al., 2014  |
| ERA  | GTPase   | Assembly of small ribosomal subunit (in prokaryotes and mitochondria) | GRMZM2G158024_P02                      | AT5G66470              | Dennerlein et al., 2010; Uchiumi et al., 2010; He et al., 2012b      |
| ObgC   | ObgC - GTPase  | Ribosome assembly   | GRMZM2G077632_P01                      | AT5G18570              | Bang et al., 2012  |
| PFC1   | rRNA methylase   | 16S rRNA methylation  | GRMZM2G174669_P01                      | AT1G01860              | Tokuhisa et al., 1998  |
| RH22   | DEAD RNA box helicase                                  | Assembly of 50S ribosomal subunit                                     | GRMZM2G100043_P01                      | AT1G59990              | Chi et al., 2012   |
| RH3*   | DEAD box RNA helicase                                  | Assembly of 50S ribosomal subunit, group II intron splicing           | GRMZM2G163072_P01<br>GRMZM2G415491_P01 | AT5G26742              | Asakura et al., 2012   |
| RIF1* (NOS1, NOA1)   | GTPase (YqeH )   | Ribosome assembly   | GRMZM2G384293_P01                      | AT3G47450              | Flores-Pérez et al., 2008; Liu et al., 2010                          |
| RsmD   | rRNA methylase   | 16S rRNA methylation (in prokaryotes)                                 | GRMZM2G010801_P01                      | AT3G28460              | Lesnyak et al., 2007   |

All proteins listed were identified as nucleoid proteins by Majeran et al. (2012). Proteins known to be involved in ribosome biogenesis but not identified in their nucleoid fraction, as well as factors that probably have only indirect effects on rRNA maturation and ribosome assembly, are not included in this list.

\*Nucleoid localization demonstrated by groups other than (Majeran et al., 2012) (see text).

<sup>a</sup>Only most relevant publications referring to a function in ribosome biogenesis or nucleoid localization are listed.

support for specific localization to the plastid nucleoid. One possible reason for this may be the use of transit peptide-GFP instead of full-length protein fusions, as

determinants of nucleoid localization are unlikely to be encoded in the transit peptide. Accordingly, Park et al. (2011) also reported the localization of the

protein PRBP, which likely functions in 4.5S rRNA processing, to distinct spots within chloroplasts, which probably represent nucleoids. This sublocalization

was only observed when the full-length protein was fused to GFP and not with a transit peptide-GFP fusion. Nevertheless, other proteins for which full length-GFP fusions were used could not be clearly assigned to the nucleoid (e.g., Flores-Pérez et al., 2008; Yu et al., 2008; Chi et al., 2012). This might be due to weak or transient association of the respective protein with the nucleoid. Alternatively, some of these proteins may perform further functions in other organellar subcompartments leading to ambiguous localization signals.

## CONCLUSION AND FUTURE PERSPECTIVES

Localization of rRNA processing and ribosome assembly to organellar nucleoids seems to be a general phenomenon derived from a bacterial ancestor. Unlike the case of the eukaryotic nucleus, no physical barrier intervenes between bacterial/organellar RNA synthesis and translation. The nucleoid might therefore provide a scaffold for an intra-organellar microenvironment which enables coupling of rRNA transcription to ribosome assembly. This might not only enhance the efficiency of ribosome assembly by substrate channeling, but also largely prevent the precocious association of mRNAs with immature 30S ribosomal subunits. However, it remains to be shown whether the colocalization of rRNA processing and ribosome assembly with nucleoids has functional significance or simply reflects the rapid kinetics with which ribosome biogenesis factors bind to nascent rRNA targets.

Nonetheless, the growing awareness of the requirement for sublocalized cellular processes in apparently less organized bacteria and organelles, and the availability of more sensitive detection techniques including super-resolution imaging, will undoubtedly lead to new insights into the spatiotemporal organization of ribosome biogenesis in the future.

## ACKNOWLEDGEMENT

I thank Joerg Nickelsen for critical reading of this article.

## REFERENCES

- Asakura, Y., Galarneau, E., Watkins, K. P., Barkan, A., and van Wijk, K. J. (2012). Chloroplast RH3 DEAD Box RNA helicases in maize and *Arabidopsis* function in splicing of specific group II introns and affect chloroplast ribosome biogenesis. *Plant Physiol.* 159, 961–974. doi: 10.1104/pp.112.197525
- Bakshi, S., Siryaporn, A., Goulian, M., and Weisshaar, J. C. (2012). Superresolution imaging of ribosomes and RNA polymerase in live *Escherichia coli* cells. *Mol. Microbiol.* 85, 21–38. doi: 10.1111/j.1365-2958.2012.08081.x
- Bang, W. Y., Chen, J., Jeong, I. S., Kim, S. W., Kim, C. W., Jung, H. S., et al. (2012). Functional characterization of OgbC in ribosome biogenesis during chloroplast development. *Plant J.* 71, 122–134. doi: 10.1111/j.1365-313X.2012.04976.x
- Beligni, M., and Mayfield, S. (2008). *Arabidopsis thaliana* mutants reveal a role for CSP41a and CSP41b, two ribosome-associated endonucleases, in chloroplast ribosomal RNA metabolism. *Plant Mol. Biol.* 67, 389–401. doi: 10.1007/s11103-008-9328-2
- Bogenhagen, D. F., Martin, D. W., and Koller, A. (2014). Initial steps in RNA processing and ribosome assembly occur at mitochondrial DNA nucleoids. *Cell Metab.* 19, 618–629. doi: 10.1016/j.cmet.2014.03.013
- Boisvert, F.-M., van Koningsbruggen, S., Navascues, J., and Lamond, A. I. (2007). The multifunctional nucleolus. *Nat. Rev. Mol. Cell Biol.* 8, 574–585. doi: 10.1038/nrm2184
- Bollenbach, T. J., Lange, H., Gutierrez, R., Erhardt, M., Stern, D. B., and Gagliardi, D. (2005). RNRI, a 3'-5' exoribonuclease belonging to the RNR superfamily, catalyzes 3' maturation of chloroplast ribosomal RNAs in *Arabidopsis thaliana*. *Nucleic Acids Res.* 33, 2751–2763. doi: 10.1093/nar/gki576
- Bollenbach, T. J., Sharwood, R. E., Gutierrez, R., Lerbs-Mache, S., and Stern, D. B. (2009). The RNA-binding proteins CSP41a and CSP41b may regulate transcription and translation of chloroplast-encoded RNAs in *Arabidopsis*. *Plant Mol. Biol.* 69, 541–552. doi: 10.1007/s11103-008-9436-z
- Bubunenko, M., Court, D. L., Al Refaii, A., Saxena, S., Korepanov, A., Friedman, D. I., et al. (2013). Nus transcription elongation factors and RNase III modulate small ribosome subunit biogenesis in *Escherichia coli*. *Mol. Microbiol.* 87, 382–393. doi: 10.1111/mmi.12105
- Bylund, G. O., Wipemo, L. C., Lundberg, L. A. C., and Wikström, P. M. (1998). RimM and RbfA are essential for efficient processing of 16S rRNA in *Escherichia coli*. *J. Bacteriol.* 180, 73–82.
- Canino, G., Bocian, E., Barbezier, N., Echeverría, M., Forner, J., Binder, S., et al. (2009). Arabidopsis encodes four tRNase Z enzymes. *Plant Physiol.* 150, 1494–1502. doi: 10.1104/pp.109.137950
- Chai, Q., Singh, B., Peisker, K., Metzendorf, N., Ge, X., Dasgupta, S., et al. (2014). Organization of ribosomes and nucleoids in *Escherichia coli* cells during growth and in quiescence. *J. Biol. Chem.* 289, 11342–11352. doi: 10.1074/jbc.M114.557348
- Chi, W., He, B., Mao, J., Li, Q., Ma, J., Ji, D., et al. (2012). The function of RH22, a DEAD RNA helicase, in the biogenesis of the 50S ribosomal subunits of *Arabidopsis* chloroplasts. *Plant Physiol.* 158, 693–707. doi: 10.1104/pp.111.186775
- Davies, B. W., Köhrer, C., Jacob, A. I., Simmons, L. A., Zhu, J., Aleman, L. M., et al. (2010). Role of *Escherichia coli* YbeY, a highly conserved protein, in rRNA processing. *Mol. Microbiol.* 78, 506–518. doi: 10.1111/j.1365-2958.2010.07351.x
- Dennerlein, S., Rozanska, A., Wydro, M., Chrzanowska-Lightowler, Z. M. A., and Lightowler, R. N. (2010). Human ERAL1 is a mitochondrial RNA chaperone involved in the assembly of the 28S small mitochondrial ribosomal subunit. *Biochem. J.* 430, 551–558. doi: 10.1042/BJ20100757
- Deutscher, M. P. (2009). Maturation and degradation of ribosomal RNA in bacteria. *Prog. Mol. Biol. Transl. Sci.* 85, 369–391. doi: 10.1016/S0079-6603(08)00809-X
- Flores-Pérez, Ú., Sauret-Güeto, S., Gas, E., Jarvis, P., and Rodríguez-Concepción, M. (2008). A mutant impaired in the production of plastome-encoded proteins uncovers a mechanism for the homeostasis of isoprenoid biosynthetic enzymes in *Arabidopsis* plastids. *Plant Cell Online* 20, 1303–1315. doi: 10.1105/tpc.108.058768
- Fristedt, R., Scharff, L. B., Clarke, C. A., Wang, Q., Lin, C., Merchant, S. S., et al. (2014). RBF1, a plant homolog of the bacterial ribosome-binding factor RbfA, acts in processing of the chloroplast 16S ribosomal RNA. *Plant Physiol.* 164, 201–215. doi: 10.1104/pp.113.228338
- Germain, A., Hotto, A. M., Barkan, A., and Stern, D. B. (2013). RNA processing and decay in plastids. *Wiley Interdiscip. Rev. RNA* 4, 295–316. doi: 10.1002/wrna.1161
- Guo, Q., Goto, S., Chen, Y., Feng, B., Xu, Y., Muto, A., et al. (2013). Dissecting the *in vivo* assembly of the 30S ribosomal subunit reveals the role of RimM and general features of the assembly process. *Nucleic Acids Res.* 41, 2609–2620. doi: 10.1093/nar/gks1256
- Gutmann, B., Goertz, A., and Giege, P. (2012). PRORP proteins support RNase P activity in both organelles and the nucleus in *Arabidopsis*. *Genes Dev.* 26, 1022–1027. doi: 10.1101/gad.189514.112
- He, J., Cooper, H. M., Reyes, A., Di Re, M., Kazak, L., Wood, S. R., et al. (2012a). Human C4orf14 interacts with the mitochondrial nucleoid and is involved in the biogenesis of the small mitochondrial ribosomal subunit. *Nucleic Acids Res.* 40, 6097–6108. doi: 10.1093/nar/gks257
- He, J., Cooper, H. M., Reyes, A., Di Re, M., Sembongi, H., Litwin, T. R., et al. (2012b). Mitochondrial nucleoid interacting proteins support mitochondrial protein synthesis. *Nucleic Acids Res.* 40, 6109–6121. doi: 10.1093/nar/gks266
- Hwang, J., and Inouye, M. (2006). The tandem GTPase, Der, is essential for the biogenesis of 50S ribosomal subunits in *Escherichia coli*. *Mol. Microbiol.* 61, 1660–1672. doi: 10.1111/j.1365-2958.2006.05348.x
- Jeon, Y., Ahn, C. S., Jung, H. J., Kang, H., Park, G. T., Choi, Y., et al. (2014). DER containing two consecutive GTP-binding domains plays an essential role in chloroplast ribosomal RNA processing and ribosome biogenesis in higher plants. *J. Exp. Bot.* 65, 117–130. doi: 10.1093/jxb/ert360
- Kaczanowska, M., and Rydén-Aulin, M. (2007). Ribosome biogenesis and the translation process in *Escherichia coli*. *Microbiol. Mol. Biol. Rev.* 71, 477–494. doi: 10.1128/MMBR.00013-07
- Kim, B.-H., Malec, P., Waloszek, A., and Arnim, A. (2012). *Arabidopsis* BPG2: a

- phytochrome-regulated gene whose protein product binds to plastid ribosomal RNAs. *Planta* 236, 677–690. doi: 10.1007/s00425-012-1638-6
- Kleinsteinknecht, L., Wang, F., Stübe, R., Philipp, K., Nickelsen, J., and Bohne, A.-V. (2014) RAP, the sole octotriopeptide repeat protein in *Arabidopsis*, is required for chloroplast 16S rRNA maturation. *Plant Cell* 26, 777–787. doi: 10.1105/tpc.114.122853
- Komatsu, T., Kawaide, H., Saito, C., Yamagami, A., Shimada, S., Nakazawa, M., et al. (2010). The chloroplast protein BPG2 functions in brassinosteroid-mediated post-transcriptional accumulation of chloroplast rRNA. *Plant J.* 61, 409–422. doi: 10.1111/j.1365-313X.2009.04077.x
- Lafontaine, D. L. J., and Tollervey, D. (2001). The function and synthesis of ribosomes. *Nat. Rev. Mol. Cell Biol.* 2, 514–520. doi: 10.1038/35080045
- Lee, K.-W., Okot-Kotber, C., LaComb, J. F., and Bogenhagen, D. F. (2013). Mitochondrial rRNA methyltransferase family members are positioned to modify nascent rRNA in foci near the mtDNA nucleoid. *J. Biol. Chem.* 288, 31386–31399. doi: 10.1074/jbc.M113.515692
- Lesnyak, D. V., Osipiuk, J., Skarina, T., Sergiev, P. V., Bogdanov, A. A., Edwards, A., et al. (2007). Methyltransferase that modifies guanine 966 of the 16S rRNA: functional identification and tertiary structure. *J. Biol. Chem.* 282, 5880–5887. doi: 10.1074/jbc.M608214200
- Lewis, P. J., Thaker, S. D., and Errington, J. (2000). Compartmentalization of transcription and translation in *Bacillus subtilis*. *EMBO J.* 19, 710–718. doi: 10.1093/emboj/19.4.710
- Liu, H., Lau, E., Lam, M. P., Chu, H., Li, S., Huang, G., et al. (2010). OsNOA1/RIF1 is a functional homolog of AtNOA1/RIF1: implication for a highly conserved plant cGTPase essential for chloroplast function. *New Phytol.* 187, 83–105. doi: 10.1111/j.1469-8137.2010.03264.x
- Majeran, W., Friso, G., Asakura, Y., Qu, X., Huang, M., Ponnala, L., et al. (2012). Nucleoid-enriched proteomes in developing plastids and chloroplasts from maize leaves: a new conceptual framework for nucleoid functions. *Plant Physiol.* 158, 156–189. doi: 10.1104/pp.111.188474
- Malagon, F. (2013). RNase III is required for localization to the nucleoid of the 5' pre-rRNA leader and for optimal induction of rRNA synthesis in *E. coli*. *RNA* 19, 1200–1207. doi: 10.1261/rna.038588.113
- Melnikov, S., Ben-Shem, A., Garreau de Loubresse, N., Jenner, L., Yusupova, G., and Yusupov, M. (2012). One core, two shells: bacterial and eukaryotic ribosomes. *Nat. Struct. Mol. Biol.* 19, 560–567. doi: 10.1038/nsmb.2313
- Miller, O. L. Jr., Hamkalo, B. A., and Thomas, C. A. Jr. (1970). Visualization of bacterial genes in action. *Science* 169, 392–395. doi: 10.1126/science.169.3943.392
- Nishimura, K., Ashida, H., Ogawa, T., and Yokota, A. (2010). A DEAD box protein is required for formation of a hidden break in *Arabidopsis* chloroplast 23S rRNA. *Plant J.* 63, 766–777. doi: 10.1111/j.1365-313X.2010.04276.x
- Park, Y.-J., Cho, H.-K., Jung, H., Ahn, C., Kang, H., and Pai, H.-S. (2011). PRBP plays a role in plastid ribosomal RNA maturation and chloroplast biogenesis in *Nicotiana benthamiana*. *Planta* 233, 1073–1085. doi: 10.1007/s00425-011-1362-7
- Qi, Y., Armbruster, U., Schmitz-Linneweber, C., Delannoy, E., de Longevialle, A. F., Röhle, T., et al. (2012). *Arabidopsis* CSP41 proteins form multi-meric complexes that bind and stabilize distinct plastid transcripts. *J. Exp. Bot.* 63, 1251–1270. doi: 10.1093/jxb/err347
- Rorbach, J., and Minczuk, M. (2012). The post-transcriptional life of mammalian mitochondrial RNA. *Biochem. J.* 444, 357–373. doi: 10.1042/BJ20112208
- Shajani, Z., Sykes, M. T., and Williamson, J. R. (2011). Assembly of bacterial ribosomes. *Annu. Rev. Biochem.* 80, 501–526. doi: 10.1146/annurev-biochem-062608-160432
- Sharwood, R., Halpert, M., Luro, S., Schuster, G., and Stern, D. (2011). Chloroplast RNase J compensates for inefficient transcription termination by removal of antisense RNA. *RNA* 17, 2165–2176. doi: 10.1261/rna.028043.111
- Stoppel, R., Manavski, N., Schein, A., Schuster, G., Teubner, M., Schmitz-Linneweber, C., et al. (2012). RHON1 is a novel ribonucleic acid-binding protein that supports RNase E function in the *Arabidopsis* chloroplast. *Nucleic Acids Res.* 40, 8593–8606. doi: 10.1093/nar/gks613
- Stoppel, R., and Meurer, J. (2012). The cutting crew – ribonucleases are key players in the control of plastid gene expression. *J. Exp. Bot.* 63, 1663–1673. doi: 10.1093/jxb/err401
- Tokuhisa, J. G., Vijayan, P., Feldmann, K. A., and Browse, J. A. (1998). Chloroplast development at low temperatures requires a homolog of DIM1, a yeast gene encoding the 18S rRNA dimethylase. *Plant Cell* 10, 699–712. doi: 10.1105/tpc.105.699
- Uchiumi, T., Ohgaki, K., Yagi, M., Aoki, Y., Sakai, A., Matsumoto, S., et al. (2010). ERAL1 is associated with mitochondrial ribosome and elimination of ERAL1 leads to mitochondrial dysfunction and growth retardation. *Nucleic Acids Res.* 38, 5554–5568. doi: 10.1093/nar/gkq305
- Walter, M., Kilian, J., and Kudla, J. (2002). PNase activity determines the efficiency of mRNA 3'-end processing, the degradation of tRNA and the extent of polyadenylation in chloroplasts. *EMBO J.* 21, 6905–6914. doi: 10.1093/emboj/cdf686
- Yu, F., Liu, X., Alsheikh, M., Park, S., and Rodermel, S. (2008). Mutations in SUPPRESSOR OF VARIEGATION1, a factor required for normal chloroplast translation, suppress var2-mediated leaf variegation in *Arabidopsis*. *Plant Cell Online* 20, 1786–1804. doi: 10.1105/tpc.107.054965
- Conflict of Interest Statement:** The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Received:** 27 March 2014; **accepted:** 19 May 2014; **published online:** 05 June 2014.
- Citation:** Bohne A-V (2014) The nucleoid as a site of rRNA processing and ribosome assembly. *Front. Plant Sci.* 5:257. doi: 10.3389/fpls.2014.00257
- This article was submitted to Plant Physiology, a section of the journal *Frontiers in Plant Science*.
- Copyright © 2014 Bohne. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Complex(iti)es of the ubiquitous RNA-binding CSP41 proteins

Dario Leister<sup>1,2\*</sup>

<sup>1</sup> Department Biology I, Plant Molecular Biology (Botany), Ludwig-Maximilians-University Munich, Martinsried, Germany

<sup>2</sup> Department of Plant and Environmental Sciences, Copenhagen Plant Science Centre, University of Copenhagen, Copenhagen, Denmark

\*Correspondence: leister@lmu.de

**Edited by:**

Thomas Pfannschmidt, University Joseph Fourier Grenoble, France

**Reviewed by:**

Zhong-Nan Yang, Shanghai Normal University, China

**Keywords:** *Arabidopsis*, chloroplast, gene expression, RNA, RNA-binding protein, transcription, translation

Photosynthetic eukaryotes encode two copies of the CSP41 (Chloroplast Stemloop binding Protein of 41 kDa) protein that are of cyanobacterial origin. In *Arabidopsis thaliana*, the two CSP41 proteins belong to the group of most-abundant chloroplast proteins. Multiple functions have been described for CSP41 proteins, including roles in chloroplast rRNA metabolism and transcription. CSP41a and CSP41b interact physically. Recent data show that CSP41b is an essential and major component of high-molecular weight complexes that form in the dark, disassemble in the light, and bind chloroplast mRNAs coding for photosynthetic proteins and some ribosomal RNAs, but not the plastid-encoded RNA polymerase (PEP). This, together with the effects seen in leaves of plants lacking CSP41b, implies that complexes containing CSP41 proteins stabilize untranslated mRNAs and precursor rRNAs. This occurs in a redox-dependent manner and seems to be important in the absence light when the translation is less active. In this scenario, translation and transcription is secondarily affected by the decreased transcript stability.

## CSP41 PROTEINS ARE ABUNDANT AND CONSTITUENTS OF COMPLEXES

CSP41 proteins are highly abundant chloroplast proteins. Zybailov et al. (2008) grouped chloroplast stromal proteins into seven abundance classes and CSP41b is found in the group of highest abundance together with the Calvin cycle enzymes for instance. CSP41b is more abundant than CSP41a, which is found in the group of second highest abundance together with most chloroplast ribosomal proteins.

Therefore, it does not come as surprise that CSP41 proteins have been detected in several stromal complexes, but not all of these associations are necessarily of physiological significance in the light of the ubiquity of these proteins.

The first report on CSP41a function described its binding *in vitro* to the 3' end of the *petD* mRNA (Yang et al., 1995, 1996). Subsequently, CSP41 proteins were also found in preparations that were enriched for the plastid-encoded RNA PEP (Pfannschmidt et al., 2000) or plastid ribosomes (Yamaguchi et al., 2003). However, later studies could not confirm that CSP41 proteins are part of the PEP complex (Suzuki et al., 2004; Pfalz et al., 2006). A ribosome association of CSP41a and b was also observed by Peltier et al. (2006) in their analysis of the stromal proteome in its oligomeric state extracted from highly purified chloroplasts of *Arabidopsis thaliana*. The two CSP41 proteins were each found at three different locations of the stromal colorless native (CN)-PAGE native gels: (i) in a complex larger than 950 kDa most likely associated with 70 S ribosomes, (ii) at 224 kDa, and (iii) at 106–126 kDa. At 224 kDa, the only obvious potential partners are the ribosomal proteins L5 and L31. At 106–126 kDa, CSP41a and b possibly form a heterotrimer. A further proteomics analysis found CSP41b mostly in 0.8–2 MDa fractions of stromal high-molecular-weight (HMW) complexes, together with other proteins like subunits of the 30S part of the plastid ribosome and subunit E2 of the plastid pyruvate decarboxylase (LTA2) (Olinares et al., 2010).

Recently, Qi et al. (2012) found by co-immunoprecipitation experiments that

the major interactor of CSP41b is the CSP41a protein. The majority of both CSP41 proteins comigrates in several distinct spots during 2D BN/SDS PAGE, implying that they are present in multimeric protein complexes mainly comprised of these two subunits. Because the HMW CSP41 complexes are disrupted by treatment with RNase, they should be associated with RNAs (Qi et al., 2012). RIP-chip analysis points to chloroplast mRNAs coding for photosynthetic proteins and some ribosomal RNAs, but no tRNAs or mRNAs for ribosomal proteins, as putative ligands of CSP41 complexes (see below). The only other protein, besides CSP41a, found in immunoprecipitates of tagged CSP41b was LTA2 (Qi et al., 2012), corroborating the results of the proteomic study of Olinares et al. (2010). Interestingly, LTA2 is not a highly abundant stromal protein (Zybailov et al., 2008); therefore, this interaction might be specific and not due to contamination.

Taken together, because of their high abundance, CSP41 proteins can be found as contaminants in preparations of several stromal complexes. However, the results obtained by Qi et al. (2012) imply that CSP41 does not functionally interact with PEP or the plastid ribosome, as proposed before. Indeed, CSP41 complexes appear to contain chloroplast mRNAs coding for photosynthetic proteins and some ribosomal RNAs.

## CSP41 FUNCTIONS AT THE BIOCHEMICAL LEVEL

Multiple functions have been assigned to CSP41 proteins. (1) RNase activity, with a preference for 3' stem-loops

(Yang and Stern, 1997; Bollenbach and Stern, 2003a,b). (2) Ribosomal biogenesis, because of the association of CSP41b with pre-ribosomal particles (Beligni and Mayfield, 2008). (3) Plastid transcription (Bollenbach et al., 2009). (4) Cytosolic functions based on its interaction with heteroglycans in the cytosol (Fettke et al., 2011).

Not all of the four tentative functions of CSP41 proteins described above might be physiologically relevant and rather represent artifacts similar to the multiple presences of CP41 proteins in HMW complexes. Moreover, the high abundance of CSP41 proteins argues against a specific catalytic function but points in the direction of a more general function, requiring relative large quantities of the protein. Recently, Qi et al. (2012) demonstrated by RIP-chip analysis that CSP41 can bind various chloroplast RNAs. This includes transcripts for the large Rubisco subunit (*rbcL*), PSI (*psaA*, *psaB*), and PSII (*psbA*, *psbC*, *psbD*) core proteins, and 16S and 23S rRNAs. Therefore, Qi et al. (2012) concluded that the CSP41 proteins might serve to stabilize RNAs. Indeed, in their *in-organello* assay the stability of two tentative target RNAs (one of them 23S rRNA) was found to be decreased in mutants lacking CSP41b. Consequently, such destabilization of the precursors of 23S and 16S rRNA might result in fewer functional ribosomes, and in turn in a decrease of the rate of chloroplast translation. As a further consequence of a reduced translation rate, a decline of the levels of PEP synthesis can be expected and, in turn, of the rate of transcription. Nevertheless, a direct effect on transcription/translation through binding of CSP41 to target transcripts cannot be entirely ruled out yet.

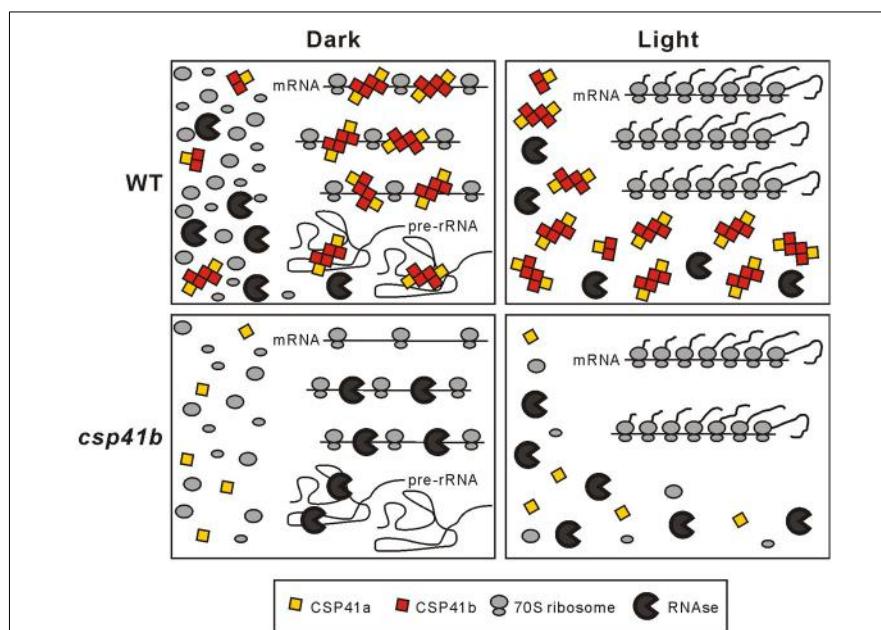
The findings that CSP41 displays endonuclease activity *in vitro* (Yang et al., 1996; Yang and Stern, 1997; Bollenbach and Stern, 2003a,b) and CSP41 proteins stabilize target RNAs *in vivo* are not necessarily mutually exclusive, because the endoribonuclease activity of CSP41 could be highly regulated *in vivo*, e.g., by phosphorylation (Qi et al., 2012). Thus, complexes of CSP41 in its inactive state (without endonuclease activity) might stabilize RNAs by binding to protect them against degradation. Certain conditions could then activate the ribonucleolytic

activity of CSP41, leading to the degradation of the target transcripts; however, it remains to be clarified whether CSP41 actually plays a role as RNase *in vivo*. In this context it is interesting to note that changes in the pIs of CSP41b species between dark and light conditions suggests that redox-dependent post-translational modifications of CSP41 might regulate the capacity of CSP41 complexes to bind RNA (Qi et al., 2012).

Immunoprecipitates of tagged CSP41b contain also LTA2, the E2 subunit of the plastid pyruvate decarboxylase (Qi et al., 2012). Interestingly, the counterpart of LTA2 in the green alga *Chlamydomonas reinhardtii*, DLA2, binds *psbA* mRNA and has been implicated in the reciprocal regulation of protein synthesis and carbon metabolism for thylakoid membrane biogenesis (Bohne et al., 2013). Therefore, their *psbA* transcript binding activity might bring together CSP41 proteins and LTA2 in the same complex.

## MUTANT ANALYSES OF CSP41 FUNCTIONS IN PLANTS

The *csp41b* mutation affects the morphology of chloroplasts, photosynthesis and circadian rhythms (Hassidim et al., 2007). Based on the observation that *A. thaliana* mutants without both CSP41 proteins are not viable, Beligni and Mayfield (2008) proposed that CSP41a and CSP41b have redundant functions. However, recent data by Qi et al. (2012) argue in favor of the notion that CSP41a and CSP41b do not have entirely redundant functions and that CSP41b is functionally more important than CSP41a. (1) While loss of CSP41a does not result in obvious phenotypic effects, chloroplast RNA levels and plant performance are impaired when CSP41b is inactive. Moreover, the *csp41ab* double mutant behaves like *csp41b* mutant plants (Qi et al., 2012). (2) Although CSP41 protein complexes seem to contain both proteins in wild-type plants, only CSP41b is essential for their formation. (3) Phylogenetic analysis of CSP41 sequences



**FIGURE 1 | Model for action of CSP41 protein complexes.** In the dark, CSP41 protein complexes associate with various mRNAs and some pre-rRNAs and protect them from nucleolytic cleavage. Untranslated RNAs not stabilized by CSP41 protein complexes are degraded by ribonucleases. Newly synthesized precursors of rRNAs are rapidly incorporated in the light into functional ribosomes, which in turn stabilize plastid transcripts during translation. In the absence of CSP41b, protection of mRNAs and pre-rRNAs in the dark is impaired and HMW RNA-CSP41 complexes are not formed. In consequence less functional ribosomes and mRNAs are available in the light. Therefore, less photosynthetic subunits are synthesized and the translational capacity is generally decreased. The latter can explain the pleiotropic effects seen in *csp41b* plants. Modified from Qi et al. (2012).

from *A. thaliana*, *C. reinhardtii* and the cyanobacterium *Synechocystis* suggest that CSP41a might be less constrained evolutionarily (Qi et al., 2012).

The major CSP41 protein, CSP41b, accumulates predominantly in mature leaves (Fettke et al., 2011). Accordingly, the function of the CSP41 proteins appears to be particularly required in mature leaves, as determined by measurements of translation rate and photosynthetic activity (Qi et al., 2012). Interestingly, mutants with reduced PEP levels show opposite behavior compared to *csp41b* mutant plants with normal mature leaves but compromised younger leaves (Chi et al., 2008; Chateigner-Boutin et al., 2011). Moreover, the sets of mRNAs that are bound by CSP41 complexes or which are transcribed by the PEP overlap. Therefore, it can be concluded that in young leaves sufficient transcripts are synthesized by PEP such that the function of CSP41 complexes is not required. In older leaves, however, chloroplast gene expression can only be maintained by transcript stabilization through CSP41 complexes. In line with this, it has been described that the stability of chloroplast transcripts increases with the leaf age (Klaff and Gruissem, 1991). In fact, the post-translational modification of CSP41 proteins could represent a development-dependent regulatory mechanism by which the function of CSP41 is controlled.

In the light with highest activity of the chloroplast translational machinery the CSP41 proteins fail to form significant amounts of HMW complexes. On the contrary, darkness induces formation of HMW CSP41 complexes that are sensitive to treatment with RNase (Qi et al., 2012). This, together with the assumption that increased polysomal association of transcripts serves to stabilize them in the light (Qi et al., 2012), is the basis for our working model for the mechanism by which CSP41 complexes stabilize mRNAs in the dark (**Figure 1**). In this model, CSP41 proteins bind in the dark to non-translated mRNAs and rRNA precursors (which are not incorporated in ribosomes) to protect them against degradation. As soon as translation is activated in the light, this would allow the rapid initiation of translation and elongation from these stabilized transcripts, as well

as *de-novo* assembly of ribosomes. In the mutants that lack CSP41 complexes, untranslated target mRNAs and precursors of rRNAs are prone to increased degradation. A role of the chloroplast redox state in the regulation of the association of CSP41 proteins with their RNA targets in the dark (and their light-induced dissociation) became evident when studying a photosynthetic mutant line with a more oxidized redox state of the stroma (Qi et al., 2012): here, HMW CSP41 complexes can persist also in the light.

## CONCLUSIONS

The key to understand the multiple functions of CSP41 proteins probably lies in their abundance. Actually, CSP41 proteins are the most abundant RNA-binding proteins in the chloroplast stroma. Their function becomes critical in mature leaves when transcripts produced by PEP might become limiting and need to be stabilized and protected. Therefore, lack of CSP41 proteins decreases transcripts for photosynthetic proteins and of some ribosomal RNAs, which in turn, appears to result in pleiotropic effects due to a decrease in the translational activity of chloroplasts.

## REFERENCES

- Beligni, M. V., and Mayfield, S. P. (2008). *Arabidopsis thaliana* mutants reveal a role for CSP41a and CSP41b, two ribosome-associated endonucleases, in chloroplast ribosomal RNA metabolism. *Plant Mol. Biol.* 67, 389–401. doi: 10.1007/s11103-008-9328-2
- Bohne, A. V., Schwarz, C., Schottkowski, M., Lidschreiber, M., Piotrowski, M., Zerges, W., et al. (2013). Reciprocal regulation of protein synthesis and carbon metabolism for thylakoid membrane biogenesis. *PLoS Biol.* 11:e1001482. doi: 10.1371/journal.pbio.1001482
- Bollenbach, T. J., Sharwood, R. E., Gutierrez, R., Lerbs-Mache, S., and Stern, D. B. (2009). The RNA-binding proteins CSP41a and CSP41b may regulate transcription and translation of chloroplast-encoded RNAs in Arabidopsis. *Plant Mol. Biol.* 69, 541–552. doi: 10.1007/s11103-008-9436-z
- Bollenbach, T. J., and Stern, D. B. (2003a). Secondary structures common to chloroplast mRNA 3'-untranslated regions direct cleavage by CSP41, an endoribonuclease belonging to the short chain dehydrogenase/reductase superfamily. *J. Biol. Chem.* 278, 25832–25838. doi: 10.1074/jbc.M303559200
- Bollenbach, T. J., and Stern, D. B. (2003b). Divalent metal-dependent catalysis and cleavage specificity of CSP41, a chloroplast endoribonuclease belonging to the short chain dehydrogenase/reductase superfamily. *Nucleic Acids Res.* 31, 4317–4325. doi: 10.1093/nar/gkg640
- Chateigner-Boutin, A. L., des Francs-Small, C. C., Delannoy, E., Kahlau, S., Tanz, S. K., de Longevialle, A. F., et al. (2011). OTP70 is a pentatricopeptide repeat protein of the E sub-group involved in splicing of the plastid transcript *rpoC1*. *Plant J.* 65, 532–542. doi: 10.1111/j.1365-313X.2010.04441.x
- Chi, W., Ma, J., Zhang, D., Guo, J., Chen, F., Lu, C., et al. (2008). The pentatricopeptide repeat protein DELAYED GREENING1 is involved in the regulation of early chloroplast development and chloroplast gene expression in *Arabidopsis*. *Plant Physiol.* 147, 573–584. doi: 10.1104/pp.108.116194
- Fettke, J., Nunes-Nesi, A., Fernie, A. R., and Steup, M. (2011). Identification of a novel heteroglycan-interacting protein, HIP 1.3, from *Arabidopsis thaliana*. *J. Plant Physiol.* 168, 1415–1425. doi: 10.1016/j.jplph.2010.09.008
- Hassidim, M., Yakir, E., Fradkin, D., Hilman, D., Kron, I., Keren, N., et al. (2007). Mutations in chloroplast RNA binding provide evidence for the involvement of the chloroplast in the regulation of the circadian clock in *Arabidopsis*. *Plant J.* 51, 551–562. doi: 10.1111/j.1365-313X.2007.03160.x
- Klaff, P., and Gruissem, W. (1991). Changes in chloroplast mRNA stability during leaf development. *Plant Cell* 3, 517–529. doi: 10.1105/tpc.3.5.517
- Olinares, P. D., Ponnola, L., and van Wijk, K. J. (2010). Megadalton complexes in the chloroplast stroma of *Arabidopsis thaliana* characterized by size exclusion chromatography, mass spectrometry and hierarchical clustering. *Mol. Cell. Proteomics* 9, 1594–1615. doi: 10.1074/mcp.M000038-MCP201
- Peltier, J. B., Cai, Y., Sun, Q., Zabrouskov, V., Giacomelli, L., Rudella, A., et al. (2006). The oligomeric stromal proteome of *Arabidopsis thaliana* chloroplasts. *Mol. Cell. Proteomics* 5, 114–133. doi: 10.1074/mcp.M500180-MCP200
- 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. doi: 10.1105/tpc.105.036392
- Pfannschmidt, T., Ogrzewalla, K., Baginsky, S., Sickmann, A., Meyer, H. E., and Link, G. (2000). The multisubunit chloroplast RNA polymerase A from mustard (*Sinapis alba* L.). Integration of a prokaryotic core into a larger complex with organelle-specific functions. *Eur. J. Biochem.* 267, 253–261. doi: 10.1046/j.1432-1327.2000.00991.x
- Qi, Y., Armbruster, U., Schmitz-Linneweber, C., Delannoy, E., de Longevialle, A. F., Rühle, T., et al. (2012). Arabidopsis CSP41 proteins form multi-meric complexes that bind and stabilize distinct plastid transcripts. *J. Exp. Bot.* 63, 1251–1270. doi: 10.1093/jxb/err347
- Suzuki, J. Y., Ytterberg, A. J., Beardslee, T. A., Allison, L. A., Wijk, K. J., and Maliga, P. (2004). Affinity purification of the tobacco plastid RNA polymerase and *in vitro* reconstitution of the holoenzyme. *Plant J.* 40, 164–172. doi: 10.1111/j.1365-313X.2004.02195.x
- Yamaguchi, K., Beligni, M. V., Prieto, S., Haynes, P. A., McDonald, W. H., Yates, J. R. 3rd, et al. (2003).

- Proteomic characterization of the *Chlamydomonas reinhardtii* chloroplast ribosome. Identification of proteins unique to the 70S ribosome. *J. Biol. Chem.* 278, 33774–33785. doi: 10.1074/jbc.M301934200
- Yang, J., Schuster, G., and Stern, D. B. (1996). CSP41, a sequence-specific chloroplast mRNA binding protein, is an endoribonuclease. *Plant Cell* 8, 1409–1420. doi: 10.1105/tpc.8.8.1409
- 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.* 272, 12874–12880. doi: 10.1074/jbc.272.19.12874
- Yang, J., Usack, L., Monde, R. A., and Stern, D. B. (1995). The 41 kDa protein component of the spinach chloroplast *petD* mRNA 3' stem-loop:protein complex is a nuclear encoded chloroplast RNA-binding protein. *Nucleic Acids Symp. Ser.* 33, 237–239.
- Zybailov, B., Rutschow, H., Friso, G., Rudella, A., Emanuelsson, O., Sun, Q., et al. (2008). Sorting signals, N-terminal modifications and abundance of the chloroplast proteome. *PLoS ONE* 3:e1994. doi: 10.1371/journal.pone.0001994

Received: 24 March 2014; accepted: 19 May 2014;  
published online: 06 June 2014.

Citation: Leister D (2014) Complex(iti)es of the ubiquitous RNA-binding CSP41 proteins. *Front. Plant Sci.* 5:255. doi: 10.3389/fpls.2014.00255

This article was submitted to Plant Physiology, a section of the journal *Frontiers in Plant Science*.

Copyright © 2014 Leister. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# A purification strategy for analysis of the DNA/RNA-associated sub-proteome from chloroplasts of mustard cotyledons

**Yvonne Schröter<sup>1</sup>, Sebastian Steiner<sup>1,2</sup>, Wolfram Weisheit<sup>1,3</sup>, Maria Mittag<sup>1,3</sup> and Thomas Pfannschmidt<sup>1,4,5,6,7\*</sup>**

<sup>1</sup> Lehrstuhl für Pflanzenphysiologie, Institut für Allgemeine Botanik und Pflanzenphysiologie, Friedrich-Schiller-Universität Jena, Jena, Germany

<sup>2</sup> KWS SAAT AG, Einbeck, Germany

<sup>3</sup> Department of General Botany, Institute of General Botany and Plant Physiology, Friedrich Schiller University Jena, Jena, Germany

<sup>4</sup> University of Grenoble-Alpes, Grenoble, France

<sup>5</sup> CNRS, UMR5168, Grenoble, France

<sup>6</sup> Commissariat à L'énergie Atomique (CEA), iRTSV, Laboratoire de Physiologie Cellulaire & Végétale, Grenoble, France

<sup>7</sup> INRA, USC1359, Grenoble, France

**Edited by:**

Steven Carl Huber, United States  
Department of  
Agriculture - Agricultural Research  
Service, USA

**Reviewed by:**

Frederik Börnke, Leibniz-Institute for  
Vegetable and Ornamental Crops  
(IGZ), Germany  
Karin Krupinska, Christian-Albrechts  
University of Kiel, Germany

**\*Correspondence:**

Thomas Pfannschmidt,  
Commissariat à L'énergie Atomique  
(CEA), iRTSV, Laboratoire de  
Physiologie Cellulaire & Végétale,  
17 Rue des Martyrs, 38000  
Grenoble, France  
e-mail: Thomas.Pfannschmidt@  
ujf-grenoble.fr

Plant cotyledons are a tissue that is particularly active in plastid gene expression in order to develop functional chloroplasts from pro-plastids, the plastid precursor stage in plant embryos. Cotyledons, therefore, represent a material being ideal for the study of composition, function and regulation of protein complexes involved in plastid gene expression. Here, we present a pilot study that uses heparin-Sepharose and phospho-cellulose chromatography in combination with isoelectric focussing and denaturing SDS gel electrophoresis (two-dimensional gel electrophoresis) for investigating the nucleic acids binding sub-proteome of mustard chloroplasts purified from cotyledons. We describe the technical requirements for a highly resolved biochemical purification of several hundreds of protein spots obtained from such samples. Subsequent mass spectrometry of peptides isolated out of cut spots that had been treated with trypsin identified 58 different proteins within 180 distinct spots. Our analyses indicate a high enrichment of proteins involved in transcription and translation and, in addition, the presence of massive post-translational modification of this plastid protein sub-fraction. The study provides an extended catalog of plastid proteins from mustard being involved in gene expression and its regulation and describes a suitable purification strategy for further analysis of low abundant gene expression related proteins.

**Keywords:** *Sinapis alba*, cotyledon, chloroplast, nucleic acids binding protein, post-translational modification, mass spectrometry

## INTRODUCTION

Plant chloroplasts are semiautonomous cell organelles of endosymbiotic origin that emerged from a cyanobacteria-like ancestor (Lopez-Juez and Pyke, 2005). One evolutionary remnant of this origin is their own genome (called plastome) comprising 100–120 genes and a pre-dominantly bacteria-like gene-expression machinery being essential for its proper expression. The plastome gene set in vascular plants is highly conserved and encodes mainly proteins with a function in photosynthesis and the gene expression machinery (Sugiura, 1992). However, for full functionality plastids require the import of many proteins that are encoded by the nuclear compartment since during evolution the endosymbiotic ancestor lost most of its genes to the nucleus of the host cell *via* horizontal gene transfer (Martin et al., 2002; Stoebe and Maier, 2002). These nuclear-encoded plastid proteins are translated in the cytoplasm as precursor molecules that are subsequently imported into plastids with the help of N-terminal transit peptides directing them to their correct sub-compartment (Soll and Schleiff, 2004). After removal of the transit peptide the

mature proteins are then assembled into their final configuration together with the plastid-expressed proteins and, therefore, all major multi-subunit complexes (such as photosystems, ribosomes, or metabolic enzyme complexes) represent a patchwork of nuclear as well as plastid expressed proteins (Allen et al., 2011).

Based on the prediction of transit peptides and genome-scale proteomics it was estimated that plastids may contain around 1500–4000 different proteins (Abdallah et al., 2000; Baerenfaller et al., 2008; Ferro et al., 2010; van Wijk and Baginsky, 2011). Reference proteomes generated for maize and *Arabidopsis* cover 1564 and 1559 proteins, respectively, so far (Huang et al., 2013) indicating that a large part of the predicted plastid proteome has yet not been detected. This might be caused by the fact that plastids from different tissues (for instance roots, cotyledons, leaves, flowers, and fruits) likely contain different protein compositions, but also from the fact that especially regulatory proteins are present in only trace amounts that are difficult to detect in a matrix of highly abundant proteins, e.g., from the photosynthetic apparatus (Huang et al., 2013). Further complexity in the plastid

protein complement may derive from the occurrence of multiple post-translational modifications that are essential for regulatory events.

Cotyledons display a high activity in plastid transcription and translation being essential for the light-induced development of chloroplasts out of the embryonic pro-plastids (Baumgartner et al., 1989, 1993). Thus, the proteome of cotyledon plastids comprises a high amount of proteins implicated in gene expression providing a useful source material for the characterization of the nucleic acids binding proteome. The chloroplast proteome of the dicotyledonous model organism *Arabidopsis thaliana* is well studied in adult leaves, however, an analysis of that of cotyledons is lacking mainly because the small size of the cotyledons is not very suitable for the isolation of chloroplasts and subsequent analyses of their proteins *via* chromatography. In recent investigations, the fast growing cruciferous plant mustard (*Sinapis alba*) demonstrated a high suitability for performing biochemical and physiological analyses of plastid gene expression in cotyledons since the seedlings and their cotyledons are much larger than that of *Arabidopsis* (Oelmüller et al., 1986; Tiller and Link, 1993; Pfannschmidt and Link, 1994; Link, 1996; Baginsky et al., 1997). Isolation of cotyledons in the order of kilograms is easily achieved after just 5 days of growth and provides enough material even for the biochemical analysis of low-abundant proteins by chromatography followed by mass spectrometry. Since *Sinapis* is a close relative of *Arabidopsis*, peptide data evaluation for the identification of mustard plastid proteins was found to be applicable for well conserved proteins by using the *A. thaliana* or *Brassicaceae* protein databases (Schröter et al., 2010; Steiner et al., 2011). Thus, the use of mustard as a source for cotyledons combines the advantages of mustard chloroplast preparation with the availability of protein data of well studied organisms like *A. thaliana* or some *Brassica* species.

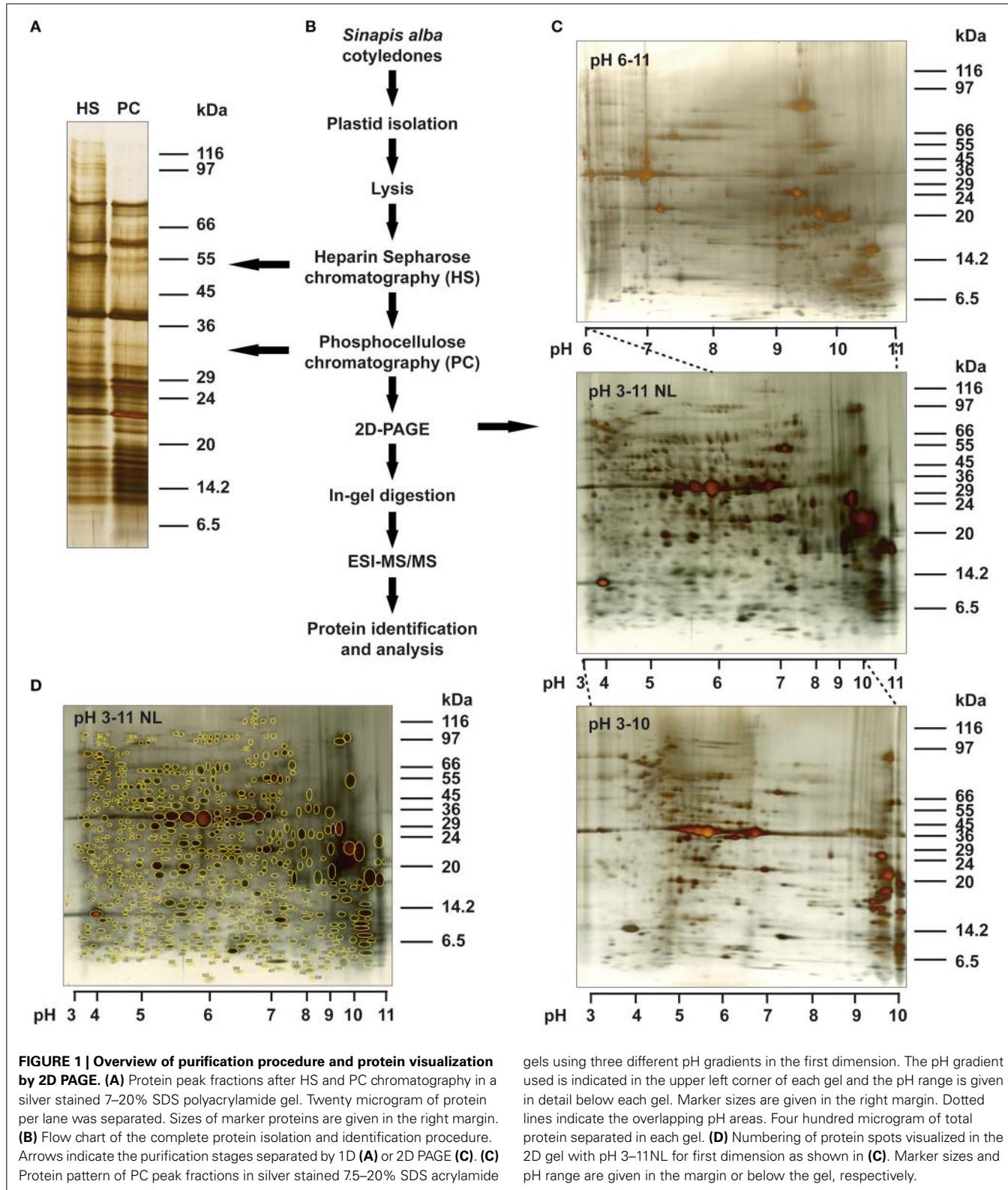
In recent studies, proteins implicated in plastid gene expression in mustard have been isolated by a number of different purification schemes. These include the isolation of the membrane bound insoluble transcriptionally active chromosome (TAC) by ultracentrifugation and gel filtration (Hallick et al., 1976; Bülow et al., 1987; Pfalz et al., 2006) and the isolation of soluble proteins such as RNA polymerases, kinases, RNA binding proteins and sigma factors by various chromatographic steps (Tiller et al., 1991; Nickelsen and Link, 1993; Tiller and Link, 1993; Pfannschmidt and Link, 1994; Liere and Link, 1995; Baginsky et al., 1999). Recently, we applied the purification scheme of plastid isolation followed by protein enrichment *via* heparin-Sepharose (HS) chromatography and visualization by two-dimensional (2D) blue native (BN)-PAGE to isolate protein complexes such as the RNA polymerase complex as well as a number of gene expression related proteins (Schröter et al., 2010; Steiner et al., 2011). However, these HS purified fractions still included a number of metabolic enzymes which exacerbate the analysis of the nucleic acids binding sub-proteome as they tend to cover low abundant proteins or even hinder their visualization and identification. Here, we present a pilot characterization of the nucleic acids binding sub-proteome of chloroplasts from mustard cotyledons. To this end we used HS chromatography followed by a second chromatographic step with phosphocellulose (PC) which

was shown to be very effective for isolating nucleic acids binding enzymes like RNA polymerases (Bottomley et al., 1970; Tiller and Link, 1993). This was followed by isoelectric focussing (IF) and 2D gel electrophoresis that allowed us to estimate the size of the nucleic acids binding sub-proteome and the ideal IF range for its visualization and protein determination using mass spectrometry. The use of 2D gel electrophoresis also revealed massive post-translational modifications of the sub-proteome.

## RESULTS

### ENRICHMENT OF NUCLEIC ACIDS BINDING PROTEINS FROM MUSTARD CHLOROPLASTS

In previous studies we analyzed gene expression related protein complexes from isolated mustard chloroplasts using a combination of HS chromatography followed by a two dimensional BN/SDS polyacrylamid gel-electrophoresis (2D BN-PAGE) and electro-spray ionization-tandem mass spectrometry (ESI-MS/MS). Besides the plastid-encoded RNA polymerase, various CSP41 complexes and translation related proteins, we identified several metabolic enzyme complexes such as GAP-dehydrogenase, ATPases, or RubisCO that co-purify in this affinity chromatography. These abundant proteins exacerbated the identification of further low-abundant proteins (Schröter et al., 2010; Steiner et al., 2011). In addition, these studies were focussed on the analysis of large native protein complexes using a BN-PAGE approach. This limited the characterization of gene expression related proteins that may occur in small complexes or as individual proteins. In this study, we aimed a deeper investigation of the size, composition and complexity of the nucleic acids binding subproteome of mustard chloroplasts. To this end, we performed chloroplast isolation and HS chromatography from mustard cotyledons precisely as described before (Schröter et al., 2010). Bound proteins were eluted with a high-salt step, concentrated by dialysis and, for further enrichment of gene expression related proteins, applied to a cation exchange column with PC as matrix as described earlier (see above). Proteins were eluted by a second high-salt step and dialyzed against a low-salt storage buffer for analysis and further use (see Materials and Methods). A first comparison of peak fractions with equal protein amounts of both purification steps was done by SDS-PAGE and silver staining (Figure 1A). The PC fraction exhibited a selective enrichment of many protein bands between 5 and 75 kDa and a strong exclusion of proteins larger than 75–80 kDa. For a more detailed resolution of this protein fraction, we performed 2D gel electrophoresis with an IF as first dimension followed by a SDS-PAGE (Figures 1B,C) as second dimension. Using IPG stripes with a non-linear (NL) pH range from 3 to 11 for the IF and a gradient polyacrylamide gel, we could obtain an overview of the total protein content leading to the identification of around 600 individual spots. We observed two major areas where multiple proteins accumulated on the gel which were located between approximately pH 4.5–7 and pH 9–11. Because of the non-linearity of the IF gradient, proteins at the outer ranges of the IF stripe were poorly resolved which became mainly evident at the basic pH values. Therefore, linear IPG gels were used in addition, overlapping with the first one between pH 3–10 and pH 6–11. The latter gradient resolved the problem with spot accumulation especially observed



at the cathode. The higher resolution led to the identification of further proteins leading to a total count of 1079 individual protein spots within the PC fraction which could be distinguished between the different gels. We regard this as the nucleic acids

binding sub-proteome of mustard plastids. Our data indicate a significant higher complexity of this specific sub-proteome as it was estimated earlier from the HS fractions (Schröter et al., 2010).

## IDENTIFICATION OF PROTEINS FROM THE PC FRACTION BY LC-ESI-MS/MS

All 1079 spots were cut out and proteins were subjected to an ingel tryptic digest. In 153 cases, selected spots were pooled from duplicate gels in order to increase the protein amount for the subsequent measurements. Since a database from *S. alba* is currently not available, protein identification was performed by comparing the determined mass spectrometry data to the *Brassicales* and *A. thaliana* databases (compare Materials and Methods). By this means 225 proteins were reliably identified with at least two different peptides in 180 spots indicating that several spots contained more than one protein. In addition, 36 particular proteins were identified in more than one spot (up to 40 different ones) suggesting post-translational modification of these proteins (Table 1). In total, 58 different proteins were identified. In further analyses, the identified gene models were checked for presence of a plastid transit peptide using TargetP (Emanuelsson et al., 2000) (Figure 2). Plastid-directing transit peptides could be predicted for 36 of these proteins, ten of them exhibit an additional luminal transit peptide and four plastid-encoded proteins were identified. Considering a detection probability of 73% for a transit peptide, we estimated the percentage of true plastid proteins within the PC fraction to be around 94%. Some of the identified proteins were found before in mustard (Pfannschmidt et al., 2000; Pfalz et al., 2006; Schröter et al., 2010), but 36 were identified here for the first time (Table 1).

Based on functional similarities and structural homologies, a categorization of proteins into protein families or subgroups was conducted (Figure 3). A practical classification mode is given by the modified MapMan bin system (Thimm et al., 2004) of the Plant Proteomics Data Base (PPDB) (Sun et al., 2009). In Table 1, proteins were listed following the PPDB bin grouping as given in column 2. For further comparison, we summarized identified proteins into five major groups. The first group comprises transcription and transcript related proteins, namely sub-units of the plastid encoded RNA polymerase (PEPs) and PEP associated proteins (PAPs) as defined in Steiner et al. (2011), other pTACs (pTAC proteins not belonging to the PAPs) and RNA and DNA related proteins (bin 27 and 28, not belonging to PAPs and pTACs). A second large group comprises translation related proteins (bin 29.2 and 29.5). Three further groups cover proteins involved in protein homeostasis (bin 29 and 21 not belonging to PEPs and PAPs), photosynthesis (bin 1) and a miscellaneous group called “others” including various enzymes catalyzing metabolic reactions or protein modifications.

### PEPs, PAPs, AND OTHER pTACs

We detected most subunits of the soluble PEP complex including PAP3, PAP4, PAP5, PAP6, PAP8, PAP10, PAP11, PAP12 as well as the PEP core subunit RpoA (Pfalz and Pfannschmidt, 2013). Other PEP core subunits (RpoB, RpoC1, RpoB) and PAP1, PAP2, PAP7, and PAP9 were not identifiable in spots of these gels. Most of the identified proteins of this group became visible as single isolated spots in the acidic range (pH 3–6) on the gel (Figure 4) and at their expected molecular weight. An exception was PAP representing the protein fructokinase-like 1 (FLN1) that contains a protein domain of the pfkB-carbohydrate kinase family

(Arsova et al., 2010; Steiner et al., 2011). This protein appeared in a chain of five spots of the same apparent molecular weight but with slightly varying isoelectric points from which the two strongest spots were identified as PAP6 here. This observation suggests post-translational modification of this kinase. In addition, for PAP6 but also for PAP3 and PAP11 one or two spots of lower molecular weight, respectively, were detected suggesting a targeted degradation or proteolytic modification of these two proteins (Figure 4). For PAP4 and PAP12 only a degradation product was detectable, while a spot of the full length protein was not identified.

Besides PEP and PAP proteins, we identified two proteins described as component of the TAC in mustard, PTAC4 and PTAC18 (Pfalz et al., 2006). PTAC4 is the vesicle-inducing protein in plastids 1 (VIPP1) which plays a crucial role in membrane stability (Zhang et al., 2012). The PTAC18 protein belongs to the cupin superfamily that merges proteins with a conserved β-barrel fold, giving this type of protein a strong thermal stability. It represents a family of very diverse members including enzymes and seed storage proteins, but also transcription factors (Dunwell et al., 2001). However, the exact function of pTAC18 is largely unknown. PTAC18 was identified in spot 255 being smaller and more in the acidic range as expected from the predicted protein representing likely a fragment. PTAC4 was identified in spots 243, 278, 280, and 295. Two hundred and seventy-eight and 280 are on the same size but with slightly different IPs suggesting post-translational modification of the protein.

An exceptional constituent of the PC protein fraction represents the protein CSP41 that appears in two forms, CSP41a and CSP41b. Originally described as the chloroplast stem-loop binding protein of 41 kDa (Yang et al., 1996) it has been discussed to be involved in RNA processing and stabilization as well as in RNA protection (Qi et al., 2012). As described for the HS fractions it represents a dominant protein of the nucleic acids binding proteome of plastids being present in multiple multimeric complexes of highly variable sizes (Schröter et al., 2010; Qi et al., 2012). In the PC fractions, the two forms of CSP41 appear to be especially enriched as they can be detected in 10 spots of the same apparent molecular weight of around 34 kDa but with different IPs (three for CSP41a and seven for CSP41b). The main accumulation is visible in the middle of the gels between pH 5.5 and 7. The CSP41a spots are by far the strongest spots observed in the whole gel followed by the spots for CSP41b. Roughly estimated they account for 30–40% of the total protein content in this fraction making a precise estimate difficult. In addition, the proteins are detectable in 34 less stained and smaller spots of different sizes suggesting massive post-translational modifications as well as multiple degradation or targeted proteolytic events acting on both protein forms. These smaller protein spots of CSP41a/b appear to contain not only random fragments of the proteins but could be observed as reproducible spot pattern in all replicates of nucleic acids binding sub-proteome preparations from mustard.

### TRANSLATION ASSOCIATED PROTEINS

Numerous proteins identified in this work are directly or indirectly related to translation. In total 12 ribosomal proteins of the large 50S subunit of plastid ribosomes (PRPL) were identified,

**Table 1 | Functional categorization and characterization of proteins from the phosphocellulose fraction identified by LC-ESI-MS/MS.**

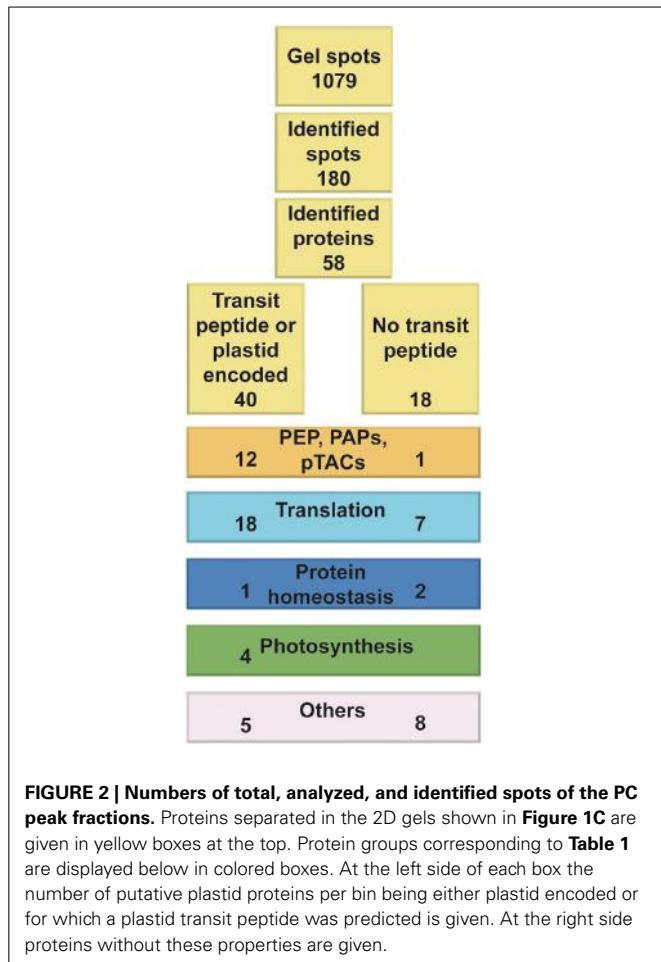
| Protein  | Spots | MapMan bin      | Accession | ATG       | cTP            | References                |
|--|-------|-----------------|-----------|-----------|----------------|---------------------------|
| <b>PEP, PAPs, pTACs</b>  |       |                 |           |           |                |                           |
| Chloroplast RNA stem-loop binding protein 41a (CSP41a)                                 | 8     | 27.3.99         | 15229384  | AT3G63140 | cTP 0.983 RC 1 | Pfannschmidt et al., 2000 |
| Chloroplast RNA stem-loop binding protein 41b (CSP41b)                                 | 40    | 27              | 15217485  | AT1G09340 | –              | Schröter et al., 2010     |
| Fructokinase-like 1 (FLN1)   | 3     | 29.4.1          | 15232415  | AT3G54090 | cTP 0.624 RC 3 | Pfalz et al., 2006        |
| Iron superoxide dismutase 3 (FSD3)   | 1     | 21.6            | 15237281  | AT5G23310 | cTP 0.945 RC 2 | Pfannschmidt et al., 2000 |
| Plastid encoded RNA polymerase alpha (RpoA)  | 2     | 27.2            | 7525065   | AtCg00740 | –              | Pfannschmidt et al., 2000 |
| Plastid transcriptionally active chromosome 4 (pTAC4)                                  | 4     | 29.3.3          | 18408237  | AT1G65260 | cTP 0.965 RC 1 | Pfalz et al., 2006        |
| Plastid transcriptionally active chromosome 6 (pTAC6)                                  | 1     | 28.3            | 79318316  | AT1G21600 | cTP 0.802 RC 2 | Pfalz et al., 2006        |
| Plastid transcriptionally active chromosome 7 (pTAC7)                                  | 1     | 35.2            | 334187898 | AT5G24314 | cTP 0.899 RC 2 | Pfalz et al., 2006        |
| Plastid transcriptionally active chromosome 10 (pTAC10)                                | 3     | 28.3            | 297816052 | –         | cTP 0.791 RC 3 | Pfalz et al., 2006        |
| Plastid transcriptionally active chromosome 12 (pTAC12)                                | 1     | 28.3            | 30686151  | AT2G34640 | cTP 0.563 RC 4 | Pfalz et al., 2006        |
| Plastid transcriptionally active chromosome 18 (pTAC18)                                | 1     | 35.2            | 15225202  | AT2G32180 | cTP 0.712 RC 4 | Pfalz et al., 2006        |
| Thioredoxin z (TRX z)  | 1     | 21.1            | 15230779  | AT3G06730 | cTP 0.858 RC 3 | Pfalz et al., 2006        |
| UDP-N-acetylmuamoylalanyl-d-glutamate-2,6-diaminopimelate ligase (MurE)                | 3     | 28.3            | 240254313 | AT1G63680 | cTP 0.695 RC 3 | Pfalz et al., 2006        |
| <b>TRANSLATION</b>   |       |                 |           |           |                |                           |
| Alpha-nascent polypeptide associated complex like protein 1 (Alpha-NAC-like protein 1) | 1     | 29.2.4          | 15230476  | AT3G12390 | –              | This work                 |
| Alpha-nascent polypeptide associated complex like protein 3 (Alpha-NAC-like protein 3) | 1     | 29.2.4          | 240256288 | AT5G13850 | –              | This work                 |
| Cytosolic ribosomal protein L11 (CRPL11)   | 1     | 29.2.1.2.2.11   | 79595462  | AT2G42740 | –              | This work                 |
| Cytosolic ribosomal protein L22-2 (CRPL22-2)   | 2     | 29.2.1.2.2.22   | 145331980 | AT3G05560 | –              | This work                 |
| Elongation factor 1-alpha4 (EF1-alpha4)  | 1     | 29.2.4          | 186532608 | AT5G60390 | –              | This work                 |
| Elongation factor tu (EFtu)  | 1     | 29.2.4          | 15237059  | AT4G20360 | cTP 0.975 RC 1 | Pfalz et al., 2006        |
| Eukaryotic translation initiation factor 1A (eIF1A)                                    | 1     | 29.2.3          | 334188030 | AT5G35680 | –              | This work                 |
| Eukaryotic translation initiation factor 3 (eIF3)                                      | 2     | 29.5.11.20      | 15225611  | AT2G39990 | cTP 0.797 RC 2 | This work                 |
| Plastid ribosomal protein L1 (PRPL1)   | 21    | 29.2.1.1.1.2.1  | 15229443  | AT3G63490 | cTP 0.937 RC 1 | This work                 |
| Plastid ribosomal protein L4 (PRPL4)   | 12    | 29.2.1.1.1.2.4  | 79317147  | AT1G07320 | cTP 0.826 RC 2 | This work                 |
| Plastid ribosomal protein L5 (PRPL5)   | 1     | 29.2.1.1.1.2.5  | 15234136  | AT4G01310 | cTP 0.828 RC 3 | This work                 |
| Plastid ribosomal protein L6 (PRPL6)   | 17    | 29.2.1.1.1.2.6  | 15220443  | AT1G05190 | cTP 0.495 RC 5 | Schröter et al., 2010     |
| Plastid ribosomal protein L10 (PRPL10)   | 6     | 29.2.1.1.1.2.10 | 15240644  | AT5G13510 | cTP 0.887 RC 2 | This work                 |
| Plastid ribosomal protein L12-1 (PRPL12-1)   | 2     | 29.2.1.1.1.2.12 | 15232274  | AT3G27830 | cTP 0.941 RC 1 | Pfalz et al., 2006        |
| Plastid ribosomal protein L12-3 (PRPL12-3)   | 2     | 29.2.1.1.1.2.12 | 15232276  | AT3G27850 | cTP 0.955 RC 1 | This work                 |

(Continued)

**Table 1 | Continued**

| <b>Protein</b>  | <b>Spots</b> | <b>MapMan bin</b> | <b>Accession</b> | <b>ATG</b> | <b>cTP</b>     | <b>References</b>     |
|---|--------------|-------------------|------------------|------------|----------------|-----------------------|
| Plastid ribosomal protein L14 (PRPL14)                                  | 1            | 29.2.1.1.1.2      | 297848252        | –          | –              | This work             |
| Plastid ribosomal protein L15 (PRPL15)                                  | 2            | 29.2.1.1.1.2.15   | 15230931         | AT3G25920  | cTP 0.866 RC 2 | This work             |
| Plastid ribosomal protein L18 (PRPL18)                                  | 1            | 29.2.1.1.1.2.18   | 15221153         | AT1G48350  | cTP 0.863 RC 2 | This work             |
| Plastid ribosomal protein L21 (PRPL21)                                  | 2            | 29.2.1.1.1.2.21   | 15219695         | AT1G35680  | cTP 0.981 RC 1 | This work             |
| Plastid ribosomal protein L24 (PRPL24)                                  | 2            | 29.2.1.1.1.2.24   | 30696487         | AT5G54600  | cTP 0.920 RC 2 | This work             |
| Plastid ribosomal protein L29 (PRPL29)                                  | 1            | 29.2.1.1.1.2.29   | 257717595        | –          | cTP 0.924 RC 2 | Pfalz et al., 2006    |
| Plastid ribosomal protein S5 (PRPS5)                                    | 14           | 29.2.1.1.1.5      | 15226167         | AT2G33800  | cTP 0.929 RC 1 | This work             |
| Translation initiation factor 2 (IF2)                                   | 7            | 29.2.3            | 15220055         | AT1G17220  | cTP 0.537 RC 3 | This work             |
| Translation initiation factor 3 (IF3)                                   | 2            | 29.2.3            | 18417644         | AT4G30690  | cTP 0.782 RC 2 | This work             |
| tRNA/rRNA methyltransferase (SpoU)                                      | 3            | 29.2.7            | 30680811         | AT2G19870  | –              | This work             |
| <b>PROTEIN HOMEOSTASIS</b>  |              |                   |                  |            |                |                       |
| Chloroplast heat shock cognate protein 70-2 (cpHSC70-2)                 | 1            | 29.6              | 15240578         | AT5G49910  | cTP 0.993 RC 1 | This work             |
| Protein disulfide isomerase like 2-1 (PDI-like 2-1)                     | 2            | 21.1              | 145331431        | AT2G47470  | –              | This work             |
| T-complex protein 1/chaperonin60 family protein (TCP1/cpn60)            | 4            | 29.6              | 15242093         | AT5G20890  | –              | This work             |
| <b>PHOTOSYNTHESIS</b>   |              |                   |                  |            |                |                       |
| ATPsynthase alpha   | 1            | 1.1.4             | 7525018          | AtCg00120  | –              | Schröter et al., 2010 |
| ATPsynthase beta  | 3            | 1.1.4             | 7525040          | AtCg00480  | –              | Schröter et al., 2010 |
| Rieske Cluster  | 2            | 1.1.3             | 30679426         | AT4G03280  | cTP 0.652 RC 3 | This work             |
| RubisCO activase  | 2            | 1.3.13            | 30687999         | AT2G39730  | cTP 0.888 RC 1 | This work             |
| <b>OTHERS</b>   |              |                   |                  |            |                |                       |
| Acetyl-coenzyme A carboxylase carboxyl transferase subunit alpha (CAC3) | 1            | 11.1.1            | 30687368         | AT2G38040  | cTP 0.927 RC 2 | This work             |
| Actin   | 1            | 31.1              | 79324605         | AT2G37620  | –              | This work             |
| Cruciferin 3 (CRU3)   | 2            | 33.1              | 15235321         | AT4G28520  | –              | This work             |
| Cystein synthase  | 1            | 13.1.5.3.1        | 334184908        | AT2G43750  | cTP 0.938 RC 1 | This work             |
| Fatty acid biosynthesis z (FabZ)  | 2            | 11.1.5            | 72255615         | –          | cTP 0.872 RC 2 | This work             |
| Malate dehydrogenase (MDH)  | 3            | 8.2.99            | 15232820         | AT3G47520  | cTP 0.911 RC 1 | This work             |
| Malate synthase (MLS)   | 4            | 6.2               | 334187411        | AT5G03860  | –              | This work             |
| Multi-functional protein 2 (MFP2)                                       | 5            | 11.9.4.9          | 15231317         | AT3G06860  | –              | Schröter et al., 2010 |
| Myrosinase  | 1            | 16.5.1            | 127734           | –          | –              | This work             |
| Phosphoserine aminotransferase (PSAT)                                   | 4            | 13.1.5.1.2        | 15237069         | AT4G35630  | cTP 0.938 RC 1 | This work             |
| Pyrroline-5-carboxylate reductase (P5CR)                                | 3            | 13.2.2.3          | 145334418        | AT5G14800  | –              | This work             |
| Serine hydroxymethyltransferase 1 (SHMT 1)                              | 8            | 25.1              | 15235745         | AT4G37930  | –              | Schröter et al., 2010 |
| Serine hydroxymethyltransferase 2 (SHMT 2)                              | 8            | 25.1              | 30690404         | AT5G26780  | –              | Schröter et al., 2010 |

Identified proteins are named in the first column according to the annotation of the respective gene at NCBI. They are grouped into different classes written bold at the beginning of each group as defined in results. Proteins within each group are sorted alphabetically. Spots: number of the spots containing the respective protein. MapMan bin: classification groups for proteins according to the modified MapMan system of the plant proteome database (ppdb) [<http://ppdb.tc.cornell.edu/dbsearch/mapman.aspx>] © Klaas J. van Wijk Lab, Cornell University; Sun et al., 2009] based on the MapManBins of Thimm et al. (2004); Accession: gi identification number and At gene accession number; cTP: possibility of a plastid transit peptide and the respective reliability class (RC); References: first identification of the protein in mustard by mass spectrometry.

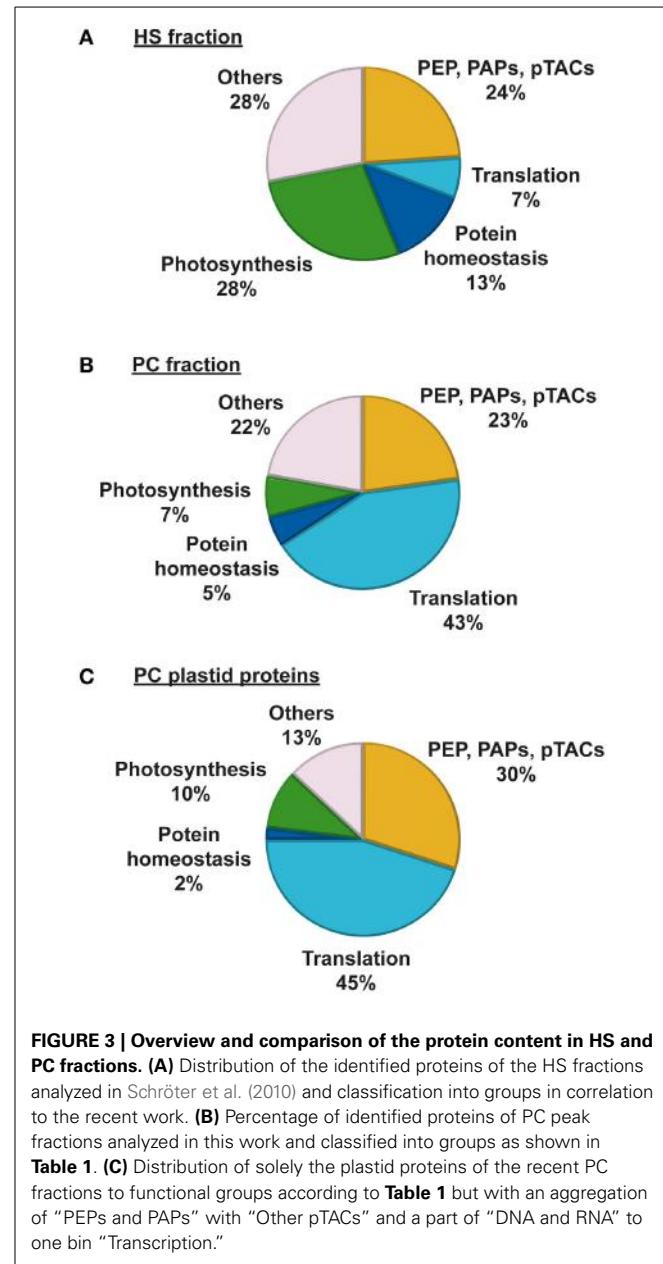


**FIGURE 2 | Numbers of total, analyzed, and identified spots of the PC peak fractions.** Proteins separated in the 2D gels shown in Figure 1C are given in yellow boxes at the top. Protein groups corresponding to Table 1 are displayed below in colored boxes. At the left side of each box the number of putative plastid proteins per bin being either plastid encoded or for which a plastid transit peptide was predicted is given. At the right side proteins without these properties are given.

namely PRPL1, -4, -5, -6, -10, -12, -14, -15, -18, -21, -24, -29. The solely identified protein of the small 30S ribosomal subunit (PRPS) is PRPS5. We also identified two ribosomal subunits that belong to the large subunit of the cytosolic 80S ribosomes (CRPL), CRPL11 and -22-2. *S. alba* proteins of PRPL12-1 and PRPL29 were formerly identified by Pfalz et al. (2006) and PRPL6 by Schröter et al. (2010). The remaining ribosomal proteins listed in Table 1 are identified in mustard plastid protein samples here for the first time.

Beside the ribosomal subunits a number of translation initiation factors (IF) were present in the fractions and were detected here for the first time in *S. alba*. Except of eIF1A (a subunit of the cytosolic translation initiation complex) all of them contain a predicted plastid transit peptide. This accounts also to eIF3 which is known as a subunit of a eukaryotic IF (eIF). IF2 and IF3 represent plastid translation IF while elongation factors (EF) EF-Tu and the eukaryotic EF1alpha4 are involved in translation elongation. eIF1A, EF-Tu, and EF1-alpha4 appear as single spots while the others were found in several spots suggesting post-translational modifications here, too.

Furthermore, we identified a SpoU methylase that belongs to the class of SPOUT enzymes and introduces a methylation of 2'-OH groups of tRNA or rRNA riboses (Cavaillé et al., 1999; Tkaczuk et al., 2007), and two proteins that are subunits of

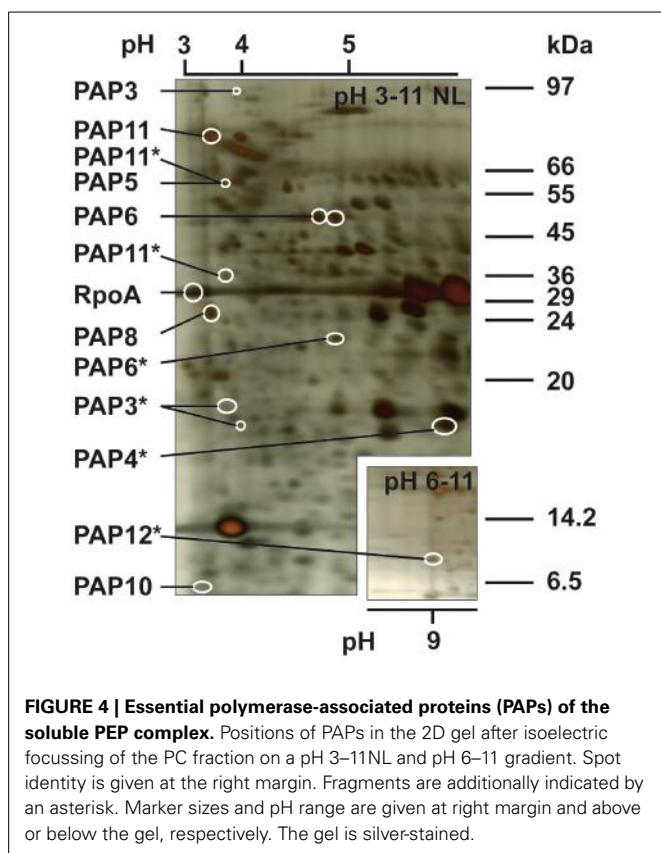


**FIGURE 3 | Overview and comparison of the protein content in HS and PC fractions.** **(A)** Distribution of the identified proteins of the HS fractions analyzed in Schröter et al. (2010) and classification into groups in correlation to the recent work. **(B)** Percentage of identified proteins of PC peak fractions analyzed in this work and classified into groups as shown in Table 1. **(C)** Distribution of solely the plastid proteins of the recent PC fractions to functional groups according to Table 1 but with an aggregation of “PEPs and PAPs” with “Other pTACs” and a part of “DNA and RNA” to one bin “Transcription.”

the nascent polypeptide associated complex (NAC). This dimeric complex is composed of an alpha- and beta-chain and may reversibly bind to ribosomes (Wiedmann et al., 1994). The alpha-NAC-like proteins identified during this work are encoded by different genes in *Arabidopsis* but exhibit a strong similarity within their amino acid sequence. The α-NAC like protein 1 and 3 were determined in the same two spots on the gels representing double spots.

#### PROTEINS INVOLVED IN PROTEIN HOMEOSTASIS, PHOTOSYNTHESIS, AND METABOLISM

We identified the chloroplast heat shock cognate protein 70-2 (cpHsc70-2) which is the analog of one of only two stromal Hsp70s in *A. thaliana* plastids (Su and Li, 2008). In addition, we



**FIGURE 4 | Essential polymerase-associated proteins (PAPs) of the soluble PEP complex.** Positions of PAPs in the 2D gel after isoelectric focussing of the PC fraction on a pH 3–11NL and pH 6–11 gradient. Spot identity is given at the right margin. Fragments are additionally indicated by an asterisk. Marker sizes and pH range are given at right margin and above or below the gel, respectively. The gel is silver-stained.

found a TCP-1/cpn60 family chaperonin and a protein disulfide isomerase like 2-1 (PDIL 2-1) belonging to the thioredoxin superfamily and acting as folding catalyst. All proteins are identified in mustard fractions here for the first time and likely function in protein stability or formation. The correct folding of proteins is the last but essential step of gene expression.

The group of photosynthesis related proteins contains four proteins. The alpha and beta subunits of the plastid ATP synthase were formerly identified in *S. alba* (Schröter et al., 2010). Another ATPase, the RubisCO activase and the Rieske cluster of the cytochrome b6/f complex were detected here first by mass spectrometry in the mustard plastid proteome. These proteins are most likely not involved in gene expression but co-purify in the column chromatography because of their substrate affinities. This is also true for the group of the miscellaneous proteins including the malate dehydrogenases (MDH) and the malate synthase (MLS), both identified in several spots.

Proteins involved in fatty acid metabolism were identified as well. These include acetyl-coenzyme A carboxylase carboxyl transferase subunit alpha (CAC3) and FabZ, a beta-hydroxyacyl acyl carrierprotein (ACP) dehydratase. An earlier study on the purification of the acetyl-CoA carboxylase multienzyme complex also resulted in the enrichment of nucleoid-associated proteins (Phinney and Thelen, 2005) suggesting a potential physical link between these two larger protein associations.

A third protein found (MFP2) is involved in lipid degradation. It was already identified in the HS-fractions in former experiments (Schröter et al., 2010). We found also a cystein

synthase and phosphoserine aminotransferase (PSAT) as well as a pyrroline-5-carboxylate reductase (P5CR) known to be essential for amino acid metabolism and a serine hydroxymethyltransferase (SHMT) being essential for photorespiration. The mustard protein in the PC fractions matches to mitochondrial SHMT1 and 2 peptides of several *Brassicaceae*. The exact affiliation to one of these SHMTs remains unclear since the matching peptides fit to both proteins (Table 1). The PC fractions contain also the myrosinase MB3 (involved in glucosinolate degradation) and a cruciferin fitting best to *A. thaliana* CRU3 (Table 1). Finally, also actin was detected in one spot, although mustard peptides of PC fractions match to different actin types of different *Brassicaceae*.

## DISCUSSION

### THE PLASTID NUCLEIC ACIDS BINDING PROTEOME OF MUSTARD

Goal of our study was the establishment of a purification scheme allowing the estimation of size and composition of the plastid nucleic acids binding sub-proteome from mustard. By using HS and PC chromatography coupled to IF and SDS-PAGE we could reproducibly isolate 1079 protein spots from which we could identify 180 protein spots by mass spectrometry. However, to our surprise these 180 protein spots were found to represent just 58 individual proteins indicating a high degree of post-translational modification of this specific sub-proteome which in part might be caused by differential phosphorylation (Reiland et al., 2009, 2011). Since we used NaF as phosphatase inhibitor in all preparation steps, the differential phosphorylation states of the analyzed proteins should be well conserved. In contrast, different redox states of thiol groups were not maintained during our purification procedure since reducing agents were included in all steps. Detection of a differential redox state in these fractions will require more specific methods such as redox difference gel electrophoresis (redox-DIGE) (Hurd et al., 2007, 2009). We also observed numerous smaller fragments from several proteins indicating degradation events. These, however, were not random as the spot pattern was reproducible between different preparations suggesting that it is not caused by action of proteases during purification, but by targeted events in the chloroplast. Whether these products represent intermediate steps of protein degradation or whether these fragments perform distinct functions remains to be determined. In summary, this high degree of post-translational modification indicates that the size of the sub-proteome is certainly smaller than the 1079 spots detected. If we assume a similar percentage of individual proteins as within the identified spots (32.2%) for the complete fraction then we estimate 347 proteins for the total nucleic acids binding sub-proteome. Since we identified a number of co-purifying proteins involved in metabolic processes (29.3%), we had to reduce this number to 236 proteins. However, our mass spectrometry determination has a certain bias since we could detect only the fraction of sufficiently abundant proteins which likely is enriched in metabolic enzymes. In addition, a significant part of post-translational modification detected in our fractions is focussed on only two proteins, CSP41a and b which partly compromise our estimate. Without these two proteins, we estimate 314 proteins for the chloroplast nucleic acids binding sub-proteome. This appears a reasonable number taking into account the proteins

that are already known to be involved in the regulation of plastid gene expression such as NEP, PEP, PAPs, pTACs, PPRs, ribosomal proteins and so on. It, however, leaves still some space for the discovery of as yet unidentified regulators that might appear only in trace amounts such as eukaryotic transcription factors (Wagner and Pfannschmidt, 2006).

#### SPECIFIC FEATURES OF THE PROTEIN FRACTION AFTER PC CHROMATOGRAPHY

PC chromatography is a well established purification step for nucleic acids binding proteins from chloroplasts (Bottomley et al., 1970; Tiller and Link, 1993). Crucial for the quality of these fractions, however, are a thorough chloroplast preparation *via* sucrose gradient centrifugation and a pre-purification step of the chloroplast lysate using HS chromatography. In comparison to results from earlier work using just HS fractions (Schröter et al., 2010) we observed a high enrichment of translation associated proteins and especially of CSP41 proteins. Co-purification of metabolic enzymes as well as components from other cell compartments was clearly reduced. We obtained a good coverage of the subunits for the plastid RNA polymerase complex PEP; however, surprisingly the larger subunits of this complex were not detectable. We observed a significant reduction of proteins above 80 kDa in size within the PC fractions (**Figure 1**), however, this might be not the reason for the failure of detection since all other components of the complex were identified in the fractions and especially RpoC2 and RpoB are known to bind DNA/RNA. Since these large subunits are highly conserved and have been successfully detected earlier in HS fractions (Steiner et al., 2011) it is likely that they are not well separated on the IEF. Further analyses using additional enrichment methodologies before the IEF step such as size-exclusion chromatography might help to target this problem in the future.

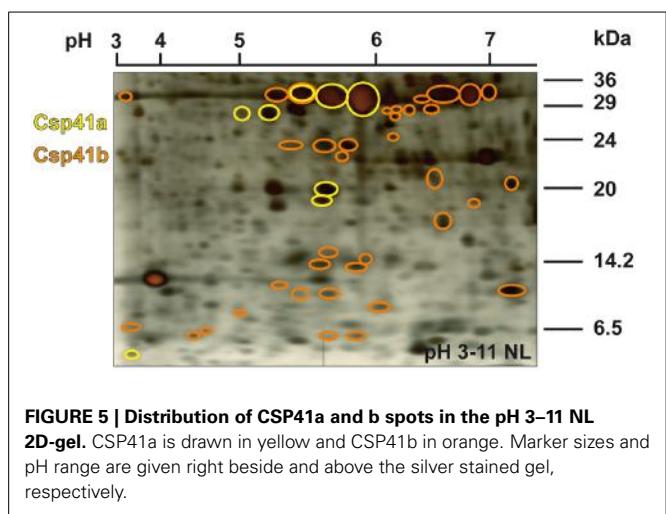
The largest amount of all identified proteins in the PC fractions is dedicated to translational processes with 43% of all proteins (**Figure 3C**). The 50S subunit of plastid ribosomes contains 33 subunits with 31 orthologs to *Escherichia coli* and the two plastid specific subunits PRPL5 and PRPL6 (Yamaguchi and Subramanian, 2000). The 30S subunit is composed of 21 *E. coli* orthologs and four plastid specific proteins with no homologs in other ribosomes (Yamaguchi et al., 2000). Most ribosomal proteins have contact to RNA in various ways, either they are structural components or directly involved in the translational process. Thus, ribosomal proteins contain nucleic acids binding structures which adhere to the used column materials and represent one main component of the nucleic acids binding sub-proteome of plastids. On the 2D-gels most of them accumulate at the higher pH-ranges and the use of the basic IPG-gels of pH 6–11 led to a good resolution of this group of proteins. The identification of 80S ribosomal proteins in plastid fractions is likely caused by the co-purification of particles attached to the outer chloroplast membrane, like known for tonoplast membrane fragments (Schröter et al., 2010). The main regulation of translation occurs at the level of initiation which is performed by initiation factors (IF). In eukaryotes this process is assured *via* 12 eIFs comprised by 23 polypeptides, whereas in prokaryotes three IFs are sufficient (Kapp and Lorsch, 2004). In plastids orthologs

for all bacteria-type translation factors can be found, but the translational complex contains additional proteins not present in bacteria (Beligni et al., 2004). Three of the four IFs identified in this study contain a cTP although only IF2 and IF3 are plastid IFs with a prokaryotic origin. The third one, eIF3f, is a subunit of the eIF3 and is important for the basic cell growth and development and influences the expression of about 3000 genes in *A. thaliana* also in interaction with two other eIF3 subunits (Xia et al., 2010). ChloroP predicts a plastid transit peptide of 40 amino acids for eIF3f of *A. thaliana* and it was previously also identified in fractions enriched in plastid nucleoids (Huang et al., 2013). Thus, it seems to be a true plastid protein and not a co-purification of the cytosolic translational apparatus. However, it might be also possible that this protein possesses a dual localization both in nucleus and plastids contributing to the coordination of gene expression between the two genetic compartments as proposed for other plant cell proteins (Krause and Krupinska, 2009). The elucidation of the precise role of eIF3f in plastids and whether it is involved in the regulation of plastid gene expression will be an interesting field of future research.

The dominant proteins in the PC fractions are the two proteins named CSP41a and CSP41b (Yang et al., 1996; Yang and Stern, 1997). CSP41a and b were also detected in isolates of the PEP-complex as one of the most abundant component (Pfannschmidt et al., 2000; Suzuki et al., 2004; Schröter et al., 2010) but they appear not to belong to the PAPs but co-purify with these fractions because of the enormous size of their largest conglomerates (Peltier et al., 2006; Schröter et al., 2010; Qi et al., 2012). Here, we identified CSP41a in 8 and CSP41b in 40 spots of diverse sizes and isoelectric points. Thereby, both form a defined spot pattern which was congruent in most replicates of the 2D-gels prepared for this work. This suggests that not only a multimerization of CSP41a/b occurs but maybe also an integration of defined fragment species of the proteins that might be important for specific functions. In addition to targeted fragmentation, the spot pattern after 2D SDS-PAGE suggests also a strong post-translational modification of the two proteins. Indeed, phosphorylation and lysine acetylation have been reported for the corresponding *Arabidopsis* proteins (Reiland et al., 2009, 2011; Finkemeier et al., 2011). The spot pattern as well as the positions of the two proteins in the 2D-gels is highly reminiscent to those recently reported for *Arabidopsis* (Qi et al., 2012). The only difference occurs in the number of identified spots which were 6 Csp41a and 5 Csp41b in *Arabidopsis* while in mustard we observed 3 Csp41a and 7 Csp41b variants (besides the fragmented versions) (**Figure 5**). This suggests the action of at least some species-specific modifications of the proteins.

#### CONCLUSION

Here, we describe the technical requirements for a highly resolved biochemical purification of several hundreds of protein spots representing the nucleic acids binding sub-proteome of plastids. Our analyses indicate a high enrichment of proteins involved in transcription and translation and, in addition, the presence of massive post-translational modification of this plastid protein sub-fraction. Furthermore, our study provides an extended catalog of plastid proteins from mustard being involved in gene



expression and its regulation and describes a suitable purification strategy for further analysis of low abundant gene expression related proteins.

## MATERIALS AND METHODS

### PLANT GROWTH AND ISOLATION OF PLASTIDS

Mustard seedlings (*Sinapis alba* L., var. Albatros) were cultivated under permanent white light illumination at 20°C and 60% humidity. Cotyledons were harvested under the respective light and stored on ice before homogenization in ice-cold isolation buffer in a Waring Blender and filtering through muslin and nylon. Chloroplast isolation by differential centrifugation and sucrose gradient centrifugation in a gradient between 30 and 55% sucrose was conducted as described earlier (Schröter et al., 2010).

### ISOLATION OF NUCLEIC ACIDS BINDING PROTEINS BY HS- AND PC-CHROMATOGRAPHY

Lysis of plastids and the chromatography at HS CL-6B was performed according to (Tiller and Link, 1993; Steiner et al., 2009). Proteins were washed, eluted with 1.2 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> and the peak fractions detected via protein quantification assays (RC DC™, Bio-Rad Laboratories, Inc., Hercules, CA, USA) (Schröter et al., 2010). For PC chromatography pooled HS peak fractions were diluted to 10% (v/v) glycerol with dilution buffer [50 mM Tris/HCl, pH 7.6, 0.1 mM EDTA, 0.1% (v/v) TritonX-100, 10 mM sodium fluoride, 62.5 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 6 mM 2-mercaptoethanol]. Activation and equilibration of PC (cellulose phosphate ion-exchanger P11, Whatman™ GE healthcare UK Limited, Little Chalfont, UK) to pH 7.6 occurred following the distributor's instructions. Diluted HS proteins were applied to disposable PD-10 columns (Amersham™ GE healthcare UK Limited, Little Chalfont, UK) filled with activated PC, closed carefully and rotated gently for 60 min at 4°C. After fixing the column on a stand and washing with washing buffer [50 mM Tris/HCl, pH 7.6, 0.1 mM EDTA, 0.1% (v/v) TritonX-100, 10 mM sodium fluoride, 50 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 5 mM 2-mercaptoethanol, 10% (v/v) glycerol] proteins were eluted in 3 ml fractions with elution buffer [50 mM Tris/HCl, pH 7.6, 0.1 mM EDTA, 0.1% (v/v) TritonX-100, 10 mM sodium fluoride, 1.2 M

(NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 5 mM 2-mercaptoethanol, 10% (v/v) glycerol] and dialyzed against storage buffer [50 mM Tris/HCl, pH 7.6, 0.1 mM EDTA, 0.1% (v/v) TritonX-100, 10 mM sodium fluoride, 50 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 5 mM 2-mercaptoethanol, 50% (v/v) glycerol]. Peak fractions were determined by a protein quantification assay (RC DC™, Bio-Rad Laboratories, Inc., Hercules, CA, USA), pooled and stored at -20°C.

### 2D GEL ELECTROPHORESIS

For 2D gel electrophoresis an acetone precipitation of dialyzed proteins from PC chromatography was used to remove the storage buffer following manuals instruction (2-D Electrophoresis principles and methods, 2004, GE healthcare UK Limited, Little Chalfont, UK). For first dimension 18 cm IPG-stripes pH 3–11NL, pH 6–11, and pH3–10 were used (GE Healthcare UK Limited, Little Chalfont, UK). IPG-stripes were rehydrated in rehydration buffer [8 M urea, 0.5% (w/v) chaps, 0.2% (w/v) DTT, 0.5% (v/v) IPG-Buffer, 0.002% (w/v) bromophenol blue] for about 14–16 h. An amount of 400 µg precipitated and dried protein per stripe was resolved in rehydration buffer and applied on the IPG-stripes as cup-loading procedure following manufacturer's instructions (2-D Electrophoresis principles and methods, 2004, GE Healthcare UK Limited, Little Chalfont, UK). For focussing of proteins the following protocol was used on IPGphor (Amersham™ GE Healthcare UK Limited, Buckinghamshire, UK) 6 h step and hold 150 V, 3 h step and hold 300 V, 6 h gradient 1200 V, 3 h gradient 8000 V, 3 h step and hold 8000 V. After IF IPG-stripes were equilibrated twice in equilibration solution [50 mM Tris-HCl, pH 8.8, 6 M urea, 30% (v/v) glycerol, 2% (w/v) SDS, 0.002% (w/v) bromophenol blue] first with addition of 2% (w/v) DTT for 15 min under gentle agitation and after removing the first solution second with addition of 2.5% (w/v) iodacetamide (IAA, for alkylation of thiol groups) again for 15 min and gently agitated as described (2-D Electrophoresis principles and methods, 2004, GE Healthcare UK Limited, Little Chalfont, UK). As second dimension a SDS PAGE in gradient gels of 7.5–20% acrylamide with Rhinohide™ gel strengthener (Molecular Probes, Inc., Eugene, OR, USA) was used following manual instructions. Afterwards gels were stained with silver according to manufacturer's instruction (Amersham™ GE Healthcare UK Limited, Buckinghamshire, UK).

### TRYPTIC DIGEST, LC/ESI-MS/MS AND DATA ANALYSIS

The spot pattern of the different gels was compared. Matching low abundant spots were pooled (as indicated in **Supplemental Table 1**) to increase the detectable protein amount. Tryptic digest of protein spots was conducted after destaining as referred (Mørzt et al., 1994; Stauber et al., 2003). Mass spectrometry was carried out at LCQ™-DecaXP ion trap mass spectrometer (Thermo Finnigan, San Jose, CA, USA) using a data-dependent scan procedure with four cyclic scan events as described in Schröter et al. (2010). The first cycle, a full MS scan of the mass range m/z 450–1200, was followed by three dependent MS/MS scans of the three most abundant ions. Sample run and data acquisition was performed using the Xcalibur™ software (Version1.3 © Thermo Finnigan 1998–2001). Seventy-six of the low abundant spots were measured at a Finnigan LTQ linear

ion trap mass spectrometer (Thermo Finnigan, Thermo Fisher Scientific Inc., Waltham, MA, USA) coupled online after a nano HPLC Ultimate 3000 (Dionex, Thermo Fisher Scientific Inc., Waltham, MA, USA) (Schmidt et al., 2006). After one full MS the instrument was set to measure the collision induced dissociation pattern of the four most abundant ions and exclude the measured once for 10 s from newly measuring.

The resulting spectra were analyzed using the Proteome Discoverer vs. 1.0 (Thermo Fisher Scientific Inc., Waltham, MA, USA) with the implemented Sequest algorithm (Link et al., 1999). Therefore, a database of all RefSeq (reference sequence) sequences of *A. thaliana* and *Arabidopsis lyrata* as well as the complete *Brassica napus* and *Capsella rubella* and the remaining *Brassicaceae* proteins of NCBI was created [NCBI 2012.03.19 109146 sequences: *Arabidopsis* RefSeq 67924 sequences (35375 *A. thaliana*, 32549 *A. lyrata*) + *B. napus* 10622 sequences + *C. rubella* 4246 sequences + other *Brassicaceae* 26354 sequences]. The Proteome Discoverer Software was set to adjust the Xcorr to reach a false discovery rate of  $\leq 1\%$  (Veith et al., 2009). All proteins with at least two unique peptides were taken for further analysis.

For transit peptide prediction the web-based tools TargetP (<http://www.cbs.dtu.dk/services/TargetP/>) (Emanuelsson et al., 2000) was used and for prediction of the transit peptide length the web-tool ChloroP (<http://www.cbs.dtu.dk/services/ChloroP/>) (Emanuelsson et al., 1999). For further analyses identified proteins were grouped into bins according to the modified MapMan system of the plant proteome database (ppdb) (<http://ppdb.tc.cornell.edu/dbsearch/mapman.aspx>) (Sun et al., 2009) based on the MapManBins of Thimm et al. (2004) (Supplemental Table 2).

## ACKNOWLEDGMENT

This work was supported by the Deutsche Forschungsgemeinschaft (Grants Pf323/4, Mi373/11-1, and Mi373/15-1).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://www.frontiersin.org/journal/10.3389/fpls.2014.00557/abstract>

**Supplemental Figure S1 | Silver stained 2D-gels of the PC fractions with isoelectric focusing for the first dimension in pH gradients between 6–11, 3–11 NL, and 3–10 indicated in the upper left corner of each gel.** The second dimension is performed in a 7.5–20% SDS polyacrylamide gel. Spots are marked and numbered in yellow. Marker sizes and pH range are given right beside and below the gel, respectively.

**Supplemental Table 1 | Identified peptides from the PC fraction.** Spots are listed in numerical order. Accession numbers of proteins belonging to the same spot are listed in an order starting with the highest peptide coverage. Spot nr., identification number of the protein containing spot on the 2D-gels (see Supplemental Figure S1). Descriptions of depicted proteins are given as stated in the databases (see Materials and Methods). Coverage, coverage of the depicted proteins by the identified peptides; calc. pl, calculated pl of the depicted proteins based on the protein sequences in the database; MW, calculated molecular weight based on the protein sequences in the databases; z, peptide ion charge;

lower case "m" in the peptide sequence, oxidized form of methionine; lower case "w" oxidized form of tryptophane; lower case "c" cysteine with carbamidomethylation; lower case "k" acetylation of lysine.

**Supplemental Table 2 | Detailed characterization of proteins from the PC fraction identified by LC-ESI-MS/MS.** Identified proteins are given in the first column according to the annotation of the respective gene at NCBI. Proteins were sorted according to the MapMan bin numbering in the second column representing classification groups for proteins according to the modified MapMan system of the plant proteome database (ppdb) [<http://ppdb.tc.cornell.edu/dbsearch/mapman.aspx>] ©Klaas J. van Wijk Lab, Cornell University (Sun et al., 2009) based on the MapManBins of Thimm et al. (2004)]. Spots: number of spots containing the respective protein. Spot nr.: identification number of the protein containing spot on the 2D-gels. NCBI accession: gi identification number at NCBI (The National Center for Biotechnology Information <http://www.ncbi.nlm.nih.gov/>); ATG: gene accession of the first matching *Arabidopsis thaliana* hit or (if no *A. thaliana* protein was matching) the best matching other organism and the respective *A. thaliana* gene accession determined by a protein-protein blast at NCBI in brackets; cTP: probability of a plastid transit peptide; MW [kDa]: calculated theoretical molecular weight in kilodalton; MW(-cTP) [kDa]: calculated theoretical molecular weight without chloroplast transit peptide (cTP) in kilodalton; MW(-ITP) [kDa]: calculated theoretical molecular weight without cTP and luminal transit peptide (ITP) in kilodalton; PI: calculated isoelectric point; PI (-cTP): calculated isoelectric point without cTP; PI (-ITP): calculated isoelectric point without cTP and ITP. The last six parameter were provided at the related accession entry by ppdb (<http://ppdb.tc.cornell.edu/dbsearch/searchacc.aspx>).

## REFERENCES

- Abdallah, F., Salamini, F., and Leister, D. (2000). A prediction of the size and evolutionary origin of the proteome of chloroplasts of *Arabidopsis*. *Trends Plant Sci.* 5, 141–142. doi: 10.1016/S1360-1385(00)01574-0
- Allen, J. F., de Paula, W. B., Puthiyaveetil, S., and Nield, J. (2011). A structural phylogenetic map for chloroplast photosynthesis. *Trends Plant Sci.* 16, 645–655. doi: 10.1016/j.tplants.2011.10.004
- Arsova, B., Hoja, U., Wimmelbacher, M., Greiner, E., Ustun, S., Melzer, M., et al. (2010). Plastidial thioredoxin z interacts with two fructokinase-like proteins in a thiol-dependent manner: evidence for an essential role in chloroplast development in *arabidopsis* and *Nicotiana benthamiana*. *Plant Cell* 22, 1498–1515. doi: 10.1105/tpc.109.071001
- Baerenfaller, K., Grossmann, J., Grobeij, M. A., Hull, R., Hirsch-Hoffmann, M., Yalovsky, S., et al. (2008). Genome-scale proteomics reveals *Arabidopsis thaliana* gene models and proteome dynamics. *Science* 320, 938–941. doi: 10.1126/science.1157956
- Baginsky, S., Tiller, K., and Link, G. (1997). Transcription factor phosphorylation by a protein kinase associated with chloroplast RNA polymerase from mustard (*Sinapis alba*). *Plant Mol. Biol.* 34, 181–189. doi: 10.1023/A:100580290902
- Baginsky, S., Tiller, K., Pfannschmidt, T., and Link, G. (1999). PTK, the chloroplast RNA polymerase-associated protein kinase from mustard (*Sinapis alba*), mediates redox control of plastid *in vitro* transcription. *Plant Mol. Biol.* 39, 1013–1023. doi: 10.1023/A:1006177807844
- Baumgartner, B. J., Rapp, J. C., and Mullet, J. E. (1989). Plastid transcription and DNA copy number increase early in barley chloroplast development. *Plant Physiol.* 89, 1011–1018. doi: 10.1104/pp.89.3.1011
- 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. *Plant Physiol.* 101, 781–791.
- Beligni, M. V., Yamaguchi, K., and Mayfield, S. P. (2004). The translational apparatus of *Chlamydomonas reinhardtii* chloroplast. *Photosynth. Res.* 82, 315–325. doi: 10.1007/s11120-004-2440-5

- Bottomley, W., Smith, H. J., and Bogorad, L. (1970). RNA polymerases of maize: partial purification and properties of the chloroplast enzyme. *Proc. Natl. Acad. Sci. U.S.A.* 71, 2412–2416.
- Bülow, S., Reiss, T., and Link, G. (1987). DNA-binding proteins of the transcriptionally active chromosome from mustard (*Sinapis alba* L.) chloroplasts. *Curr. Genet.* 12, 157–159. doi: 10.1007/BF00434671
- Cavaillé, J., Chetouani, F., and Bachellerie, J. P. (1999). The yeast *Saccharomyces cerevisiae* YDL112w ORF encodes the putative 2'-O-ribose methyltransferase catalyzing the formation of Gm18 in tRNAs. *RNA* 5, 66–81. doi: 10.1017/S1355838299981475
- Dunwell, J. M., Culham, A., Carter, C. E., Sosa-Aguirre, C. R., and Goodenough, P. W. (2001). Evolution of functional diversity in the cupin superfamily. *Trends Biochem. Sci.* 26, 740–746. doi: 10.1016/S0968-0004(01)01981-8
- Emanuelsson, O., Nielsen, H., Brunak, S., and von Heijne, G. (2000). Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J. Mol. Biol.* 300, 1005–1016. doi: 10.1006/jmbi.2000.3903
- Emanuelsson, O., Nielsen, H., and von Heijne, G. (1999). ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. *Protein Sci.* 8, 978–984. doi: 10.1110/ps.8.5.978
- Ferro, M., Brugiére, S., Salvi, D., Seigneurin-Berny, D., Court, M., Moyet, L., et al. (2010). AT\_CHLORO, a comprehensive chloroplast proteome database with subplastidial localization and curated information on envelope proteins. *Mol. Cell. Proteomics* 9, 1063–1084. doi: 10.1074/mcp.M900325-MCP200
- Finkemeier, I., Laxa, M., Miguët, L., Howden, A. J., and Sweetlove, L. J. (2011). Proteins of diverse function and subcellular location are lysine acetylated in *Arabidopsis*. *Plant Physiol.* 155, 1779–1790. doi: 10.1104/pp.110.171595
- Hallick, R. B., Lipper, C., Richards, O. C., and Rutter, W. J. (1976). Isolation of a transcriptionally active chromosome from chloroplasts of *Euglena gracilis*. *Biochemistry* 15, 3039–3043. doi: 10.1021/bi00659a016
- Huang, M., Friso, G., Nishimura, K., Qu, X., Olinares, P. D., Majeran, W., et al. (2013). Construction of plastid reference proteomes for maize and *Arabidopsis* and evaluation of their orthologous relationships; the concept of orthoproteomics. *J. Proteome Res.* 12, 491–504. doi: 10.1021/pr300952g
- Hurd, T. R., James, A. M., Lilley, K. S., and Murphy, M. P. (2009). Measuring redox changes to mitochondrial protein thiols with redox difference gel electrophoresis (redox-DIGE). *Methods Enzymol.* 456, 343–361. doi: 10.1016/S0076-6879(08)04419-4
- Hurd, T. R., Prime, T. A., Harbour, M. E., Lilley, K. S., and Murphy, M. P. (2007). Detection of reactive oxygen species-sensitive thiol proteins by redox difference gel electrophoresis: implications for mitochondrial redox signaling. *J. Biol. Chem.* 272, 22040–22051. doi: 10.1074/jbc.M703591200
- Kapp, L. D., and Lorsch, J. R. (2004). The molecular mechanics of eukaryotic translation. *Annu. Rev. Biochem.* 73, 657–704. doi: 10.1146/annurev.biochem.73.030403.080419
- Krause, K., and Krupinska, K. (2009). Nuclear regulators with a second home in organelles. *Trends Plant Sci.* 14, 194–199. doi: 10.1016/j.tplants.2009.01.005
- Liere, K., and Link, G. (1995). RNA-binding activity of the matK protein encoded by the chloroplast trnK intron from mustard (*Sinapis alba* L.). *Nucleic Acids Res.* 23, 917–921. doi: 10.1093/nar/23.6.917
- Link, A. J., Eng, J., Schieltz, D. M., Carmack, E., Mize, G. J., Morris, D. R., et al. (1999). Direct analysis of protein complexes using mass spectrometry. *Nat. Biotechnol.* 17, 676–682. doi: 10.1038/10890
- Link, G. (1996). Green life: control of chloroplast gene transcription. *Bioessays* 18, 465–471. doi: 10.1002/bies.950180608
- Lopez-Juez, E., and Pyke, K. A. (2005). Plastids unleashed: their development and their integration in plant development. *Int. J. Dev. Biol.* 49, 557–577. doi: 10.1387/ijdb.051997el
- Martin, W., Rujan, T., Richly, E., Hansen, A., Cornelissen, S., Lins, T., et al. (2002). Evolutionary analysis of *Arabidopsis*, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. *Proc. Natl. Acad. Sci. U.S.A.* 99, 12246–12251. doi: 10.1073/pnas.182432999
- Mørtz, E., Vorm, O., Mann, M., and Roepstorff, P. (1994). Identification of proteins in polyacrylamide gels by mass spectrometric peptide mapping combined with database search. *Biol. Mass Spectrom.* 23, 249–261. doi: 10.1002/bms.1200230503
- Nickelsen, J., and Link, G. (1993). The 54 kDa RNA-binding protein from mustard chloroplasts mediates endonucleolytic transcript 3' end formation *in vitro*. *Plant J.* 3, 537–544. doi: 10.1046/j.1365-313X.1993.03040537.x
- Oelmüller, R., Dietrich, G., Link, G., and Mohr, H. (1986). Regulatory factors involved in gene-expression (subunits of ribulose-1,5-bisphosphate carboxylase) in mustard (*Sinapis alba* L.) cotyledons. *Planta* 169, 260–266. doi: 10.1007/BF00392323
- Peltier, J. B., Cai, Y., Sun, Q., Zabrouskov, V., Giacomelli, L., Rudella, A., et al. (2006). The oligomeric stromal proteome of *Arabidopsis thaliana* chloroplasts. *Mol. Cell. Proteomics* 5, 114–133. doi: 10.1074/mcp.M500180-MCP200
- Pfälz, J., Liere, K., Kandlbinder, A., Dietz, K. J., and Oelmüller, R. (2006). PTAC-2, -6, and -12 are components of the transcriptionally active plastid chromosome that are required for plastid gene expression. *Plant Cell* 18, 176–197. doi: 10.1105/tpc.105.036392
- Pfälz, J., and Pfannschmidt, T. (2013). Essential nucleoid proteins in early chloroplast development. *Trends Plant Sci.* 18, 186–194. doi: 10.1016/j.tplants.2012.11.003
- Pfannschmidt, T., and Link, G. (1994). Separation of 2 classes of plastid DNA-dependent RNA-polymerases that are differentially expressed in mustard (*Sinapis alba* L.) Seedlings. *Plant Mol. Biol.* 25, 69–81. doi: 10.1007/BF00024199
- Pfannschmidt, T., Ogrzewalla, K., Baginsky, S., Sickmann, A., Meyer, H. E., and Link, G. (2000). The multisubunit chloroplast RNA polymerase A from mustard (*Sinapis alba* L.) – integration of a prokaryotic core into a larger complex with organelle-specific functions. *Eur. J. Biochem.* 267, 253–261. doi: 10.1046/j.1432-1327.2000.00991.x
- Phinney, B. S., and Thelen, J. J. (2005). Proteomic characterization of a Triton-insoluble fraction from chloroplasts defines a novel group of proteins associated with macromolecular structures. *J. Proteome Res.* 4, 497–506. doi: 10.1021/pr049791k
- Qi, Y., Arnbruster, U., Schmitz-Linneweber, C., Delannoy, E., de Longevialle, A. F., Rühle, T., et al. (2012). *Arabidopsis* CSP41 proteins form multimeric complexes that bind and stabilize distinct plastid transcripts. *J. Exp. Bot.* 63, 1251–1270. doi: 10.1093/jxb/err347
- Reiland, S., Finazzi, G., Endler, A., Willig, A., Baerenfaller, K., Grossmann, J., et al. (2011). Comparative phosphoproteome profiling reveals a function of the STN8 kinase in fine-tuning of cyclic electron flow (CEF). *Proc. Natl. Acad. Sci. U.S.A.* 108, 12955–12960. doi: 10.1073/pnas.1104734108
- Reiland, S., Messerli, G., Baerenfaller, K., Gerrits, B., Endler, A., Grossmann, J., et al. (2009). Large-scale *Arabidopsis* phosphoproteome profiling reveals novel chloroplast kinase substrates and phosphorylation networks. *Plant Physiol.* 150, 889–903. doi: 10.1104/pp.109.138677
- Schmidt, M., Gessner, G., Luff, M., Heiland, I., Wagner, V., Kaminski, M., et al. (2006). Proteomic analysis of the eyespot of *Chlamydomonas reinhardtii* provides novel insights into its components and tactic movements. *Plant Cell* 18, 1908–1930. doi: 10.1105/tpc.106.041749
- Schröter, Y., Steiner, S., Matthai, K., and Pfannschmidt, T. (2010). Analysis of oligomeric protein complexes in the chloroplast sub-proteome of nucleic acid-binding proteins from mustard reveals potential redox regulators of plastid gene expression. *Proteomics* 10, 2191–2204. doi: 10.1002/pmic.200900678
- Soll, J., and Schleiff, E. (2004). Protein import into chloroplasts. *Nat. Rev. Mol. Cell Biol.* 5, 198–208. doi: 10.1038/nrm1333
- Stauber, E. J., Fink, A., Markert, C., Kruse, O., Johanningmeier, U., and Hippler, M. (2003). Proteomics of *Chlamydomonas reinhardtii* light-harvesting proteins. *Eukaryot. Cell* 2, 978–994. doi: 10.1128/EC.2.5.978-994.2003
- Steiner, S., Dietzel, L., Schröter, Y., Fey, V., Wagner, R., and Pfannschmidt, T. (2009). The role of phosphorylation in redox regulation of photosynthesis genes psaA and psbA during photosynthetic acclimation of mustard. *Mol. Plant* 2, 416–429. doi: 10.1093/mp/ssp007
- Steiner, S., Schröter, Y., Pfälz, J., and Pfannschmidt, T. (2011). Identification of essential subunits in the plastid-encoded RNA polymerase complex reveals building blocks for proper plastid development. *Plant Physiol.* 157, 1–13. doi: 10.1104/pp.111.184515
- Stoebe, B., and Maier, U. G. (2002). One, two, three: nature's tool box for building plastids. *Protoplasma* 219, 123–130. doi: 10.1007/s007090200013
- Su, P. H., and Li, H. M. (2008). *Arabidopsis* stromal 70-kD heat shock proteins are essential for plant development and important for thermotolerance of germinating seeds. *Plant Physiol.* 146, 1231–1241. doi: 10.1104/pp.107.114496
- Sugiura, M. (1992). The chloroplast genome. *Plant Mol. Biol.* 19, 149–168. doi: 10.1007/BF00015612
- Sun, Q., Zybalov, B., Majeran, W., Friso, G., Olinares, P. D. B., and Van Wijk, K. J. (2009). PPDB, the plant proteomics database at cornell. *Nucleic Acids Res.* 37, D969–D974. doi: 10.1093/nar/gkn654

- Suzuki, J. Y., Ytterberg, A. J., Beardslee, T. A., Allison, L. A., Wijk, K. J., and Maliga, P. (2004). Affinity purification of the tobacco plastid RNA polymerase and *in vitro* reconstitution of the holoenzyme. *Plant J.* 40, 164–172. doi: 10.1111/j.1365-313X.2004.02195.x
- Thimm, O., Blasing, O., Gibon, Y., Nagel, A., Meyer, S., Kruger, P., et al. (2004). MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. *Plant J.* 37, 914–939. doi: 10.1111/j.1365-313X.2004.02016.x
- Tiller, K., Eisermann, A., and Link, G. (1991). The chloroplast transcription apparatus from mustard (*Sinapis alba L.*). Evidence for three different transcription factors which resemble bacterial sigma factors. *Eur. J. Biochem.* 198, 93–99. doi: 10.1111/j.1432-1033.1991.tb15990.x
- Tiller, K., and Link, G. (1993). Phosphorylation and dephosphorylation affect functional-characteristics of chloroplast and etioplast transcription systems from mustard (*Sinapis alba L.*). *EMBO J.* 12, 1745–1753.
- Tkaczuk, K. L., Dunin-Horkawicz, S., Purta, E., and Bujnicki, J. M. (2007). Structural and evolutionary bioinformatics of the SPOUT superfamily of methyltransferases. *BMC Bioinformatics* 8:73. doi: 10.1186/1471-2105-8-73
- van Wijk, K. J., and Baginsky, S. (2011). Update on plastid proteomics in higher plants; current state and future goals. *Plant Physiol.* 155, 1578–1588. doi: 10.1104/pp.111.172932
- Veith, T., Brauns, J., Weisheit, W., Mittag, M., and Büchel, C. (2009). Identification of a specific fucoxanthin-chlorophyll protein in the light harvesting complex of photosystem I in the diatom *Cyclotella meneghiniana*. *Biochim. Biophys. Acta* 1787, 905–912. doi: 10.1016/j.bbabi.2009.04.006
- Wagner, R., and Pfannschmidt, T. (2006). Eukaryotic transcription factors in plastids – bioinformatic assessment and implications for the evolution of gene expression machineries in plants. *Gene* 381, 62–70. doi: 10.1016/j.gene.2006.06.022
- Wiedmann, B., Sakai, H., Davis, T. A., and Wiedmann, M. (1994). A protein complex required for signal-sequence-specific sorting and translocation. *Nature* 370, 434–440. doi: 10.1038/370434a0
- Xia, C., Wang, Y. J., Li, W. Q., Chen, Y. R., Deng, Y., Zhang, X. Q., et al. (2010). The *Arabidopsis* eukaryotic translation initiation factor 3, subunit F (AtEF3f), is required for pollen germination and embryogenesis. *Plant J.* 63, 189–202. doi: 10.1111/j.1365-313X.2010.04237.x
- Yamaguchi, K., and Subramanian, A. R. (2000). The plastid ribosomal proteins. Identification of all the proteins in the 50S subunit of an organelle ribosome (chloroplast). *J. Biol. Chem.* 275, 28466–28482. doi: 10.1074/jbc.M005012200
- Yamaguchi, K., von Knoblauch, K., and Subramanian, A. R. (2000). The plastid ribosomal proteins. Identification of all the proteins in the 30S subunit of an organelle ribosome (chloroplast). *J. Biol. Chem.* 275, 28455–28465. doi: 10.1074/jbc.M004350200
- Yang, J. J., Schuster, G., and Stern, D. B. (1996). CSP41, a sequence-specific chloroplast mRNA binding protein, is an endoribonuclease. *Plant Cell* 8, 1409–1420. doi: 10.1105/tpc.8.8.1409
- Yang, J. 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.* 272, 12874–12880. doi: 10.1074/jbc.272.19.12874
- Zhang, L., Kato, Y., Otters, S., Vothknecht, U. C., and Sakamoto, W. (2012). Essential role of VIPP1 in chloroplast envelope maintenance in *Arabidopsis*. *Plant Cell* 24, 3695–3707. doi: 10.1105/tpc.112.103606

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Received: 11 July 2014; accepted: 29 September 2014; published online: 29 October 2014.*

*Citation: Schröter Y, Steiner S, Weisheit W, Mittag M and Pfannschmidt T (2014) A purification strategy for analysis of the DNA/RNA-associated sub-proteome from chloroplasts of mustard cotyledons. *Front. Plant Sci.* 5:557. doi: 10.3389/fpls.2014.00557*

*This article was submitted to Plant Physiology, a section of the journal *Frontiers in Plant Science*.*

*Copyright © 2014 Schröter, Steiner, Weisheit, Mittag and Pfannschmidt. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*

# Advantages of publishing in Frontiers



## OPEN ACCESS

Articles are free to read,  
for greatest visibility



## COLLABORATIVE PEER-REVIEW

Designed to be rigorous  
– yet also collaborative,  
fair and constructive



## FAST PUBLICATION

Average 85 days from  
submission to publication  
(across all journals)



## COPYRIGHT TO AUTHORS

No limit to article  
distribution and re-use



## TRANSPARENT

Editors and reviewers  
acknowledged by name  
on published articles



## SUPPORT

By our Swiss-based  
editorial team



## IMPACT METRICS

Advanced metrics  
track your article's impact



## GLOBAL SPREAD

5'100'000+ monthly  
article views  
and downloads



## LOOP RESEARCH NETWORK

Our network  
increases readership  
for your article

### Frontiers

EPFL Innovation Park, Building I • 1015 Lausanne • Switzerland  
Tel +41 21 510 17 00 • Fax +41 21 510 17 01 • info@frontiersin.org  
[www.frontiersin.org](http://www.frontiersin.org)

### Find us on

