Title	Subfunctionalization of sigma factors during the evolution of land plants based on mutant analysis of liverwort (Marchantia polymorpha L.) MpSIG1.
Author(s)	Ueda, Minoru; Takami, Tsuneaki; Peng, Lianwei; Ishizaki, Kimitsune; Kohchi, Takayuki; Shikanai, Toshiharu; Nishimura, Yoshiki
Citation	Genome biology and evolution (2013), 5(10): 1836-1848
Issue Date	2013-09-11
URL	http://hdl.handle.net/2433/182037
Right	© The Author(s) 2013. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution.; This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited.
Type	Journal Article
Textversion	publisher

GBE

Subfunctionalization of Sigma Factors during the Evolution of Land Plants Based on Mutant Analysis of Liverwort (*Marchantia polymorpha* L.) *MpSIG1*

Minoru Ueda[1], Tsuneaki Takami[1,2], Lianwei Peng[1,4], Kimitsune Ishizaki[3,5], Takayuki Kohchi[3], Toshiharu Shikanai[1], and Yoshiki Nishimura[1,*]

[1]Department of Botany, Graduate School of Science, Kyoto University, Japan

[2]Graduate School of Agriculture, Kyushu University, Fukuoka, Japan

[3]Graduate School of Biostudies, Kyoto University, Japan

[4]Present address: Key Laboratory of Photobiology, Institute of Botany, Chinese Academy of Sciences, Beijing, China

[5]Present address: Graduate School of Science, Kobe University, 1-1 Rokkodai, Nada-ku, Kobe, 657-8501 Japan

*Corresponding author: E-mail: yoshiki@pmg.bot.kyoto-u.ac.jp.

Accepted: September 3, 2013

Data deposition: The nucleotide sequence data reported in this paper have been deposited at the DDBJ/EMBL/GenBank database under the accession numbers AB621976 (*Marchantia polymorpha SIG1* [*MpSIG1*]), AB621977 (*Marchantia polymorpha SIG2* [*MpSIG2*]), and AB621978 (*Marchantia polymorpha SIG5* [*MpSIG5*]).

Abstract

Sigma factor is a subunit of plastid-encoded RNA polymerase that regulates the transcription of plastid-encoded genes by recognizing a set of promoters. Sigma factors have increased in copy number and have diversified during the evolution of land plants, but details of this process remain unknown. Liverworts represent the basal group of embryophytes and are expected to retain the ancestral features of land plants. In liverwort (*Marchantia polymorpha* L.), we isolated and characterized a T-DNA-tagged mutant (*Mpsig1*) of sigma factor 1 (*MpSIG1*). The mutant did not show any visible phenotypes, implying that MpSIG1 function is redundant with that of other sigma factors. However, quantitative reverse-transcription polymerase chain reaction and RNA gel blot analysis revealed that genes related to photosynthesis were downregulated, resulting in the minor reduction of some protein complexes. The transcript levels of genes clustered in the *petL*, *psaA*, *psbB*, *psbK*, and *psbE* operons of liverwort were lower than those in the wild type, a result similar to that in the SIG1 defective mutant in rice (*Oryza sativa*). Overexpression analysis revealed primitive functional divergence between the SIG1 and SIG2 proteins in bryophytes, whereas these proteins still retain functional redundancy. We also discovered that the predominant sigma factor for *ndhF* mRNA expression has been diversified in liverwort, *Arabidopsis* (*Arabidopsis thaliana*), and rice. Our study shows the ancestral function of *SIG1* and the process of functional partitioning (subfunctionalization) of sigma factors during the evolution of land plants.

Key words: chloroplast transcription, plastid-encoded RNA polymerase, *Marchantia*, gene duplication.

Introduction

Chloroplasts are considered to be the descendants of free-living cyanobacteria that were engulfed by ancestral eukaryotic cells. After the initial endosymbiotic event, most of the genes encoded by the ancestral endosymbiont genome have either been translocated to the nuclear genome of the host cell or have been lost during evolution (Gray 1992; Timmis et al. 2004). However, approximately 120 genes are retained in the plastid genome of land plants, and plastids have semi-autonomous machinery for the expression of their own genomes (Sugiura 1992; Barkan and Goldschmidt-Clermont 2000).

In land plants, including the moss *Physcomitrella patens*, monocots, and dicots, two types of RNA polymerases function in plastid transcription: The multi-subunit, plastid-encoded RNA polymerase (PEP) mainly transcribes photosynthesis-related genes, and the single-subunit, nuclear-encoded RNA polymerase (NEP) mainly mediates the transcription of

house-keeping genes (Liere et al. 2011). Although the core subunits of PEP are encoded by the plastid genome, with some exceptions such as *rpoA* in *P. patens* (Sugiura et al. 2003), the sigma factors are encoded by a small gene family present in the nuclear genomes of land plants. Each member of the family recognizes a different set of gene promoters and thereby facilitates the regulation of transcription in response to changes in the environment and plastid differentiation (Shiina et al. 2009). Six sigma factor genes are present in *Arabidopsis* (*Arabidopsis thaliana*: *AtSIG1–AtSIG6*) (Isono et al. 1997; Tanaka et al. 1997; Fujiwara et al. 2000) and rice (*Oryza sativa*: *OsSIG1*, *OsSIG2A*, *OsSIG2B*, *OsSIG3*, *OsSIG5*, and *OsSIG6*) (Tozawa et al. 1998; Kasai et al. 2004; Kubota et al. 2007). In contrast, sigma factor in the green alga *Chlamydomonas reinhardtii* is a single-copy gene (Carter et al. 2004; Bohne et al. 2006). In *P. patens*, three sigma factors (*PpSIG1*, *PpSIG2*, and *PpSIG5*) are encoded in the nucleus (Hara et al. 2001; Hara, Sugita, et al. 2001; Ichikawa et al. 2004), suggesting that bryophytes (mosses, liverworts, and hornworts) are in an intermediate stage in the evolutionary expansion of the sigma factor family.

To date, *Ossig1*, *Atsig2*, *Atsig3*, *Atsig4*, *Atsig5*, and *Atsig6* mutants have been comprehensively characterized in angiosperms. Expression analyses of these mutants revealed that sigma factors contribute to the transcriptional control of a set of plastid genes, although partial overlaps were observed between the gene sets in the plastid genome. *OsSIG1*, *AtSIG2*, *AtSIG3*, *AtSIG4*, *AtSIG5*, and *AtSIG6* predominantly regulate the expression of the *psaA* operon (Tozawa et al. 2007), tRNAs (Kanamaru et al. 2001; Hanaoka et al. 2003), *psbN* (Zghidi et al. 2007), *ndhF* (Favory et al. 2005), blue-light-responsive *psbD* promoter (Nagashima et al. 2004; Tsunoyama et al. 2004), and photosynthetic genes (Ishizaki et al. 2005) (reviewed in Shiina et al. [2009] and Lerbs-Mache [2011]), respectively. These observations are consistent with the idea that plastid gene expression is regulated by a family of sigma factors, as well as by posttranscriptional steps (Lerbs-Mache 2011; Liere et al. 2011). An open question is to what extent this regulation is established in bryophytes that have a smaller family of sigma factors.

Land plants emerged over 470 Ma (Rubinstein et al. 2010), and liverworts are considered to represent the basal group of terrestrial embryophytes (Steemans et al. 2009; Rubinstein et al. 2010). Liverwort (*Marchantia polymorpha* L.) is an emerging model plant with which various molecular techniques have been established (Ishizaki et al. 2008). Functional analyses of the sigma factor genes in liverwort are expected to reveal the process through which the sigma factor family has expanded in size.

In this study, we have isolated a T-DNA-tagged mutant for *MpSIG1* in liverwort. The lack of a visible phenotype in the *Mpsig1* mutant implied that the function of MpSIG1 is redundant with that of other sigma factors. However, the transcript levels of certain genes were reduced, including those in the *psaA*, *psbB*, and *psbE* operons that are also recognized by OsSIG1 in rice. Moreover, similar reductions in transcript abundance were observed for several genes, including *ndhF* that is recognized by AtSIG4 in *Arabidopsis*. Considering these results and our analyses of *MpSIG1*- and *MpSIG2*-overexpressing plants, we discuss the ancestral function of *SIG1* in liverworts and the functional partitioning (subfunctionalization) of sigma factors coordinated by gene loss from chloroplast genomes during the evolution of land plants.

Materials and Methods

Plant Materials and Growth Conditions

Male (accession Takaragaike-1 [Tak1]) and female (accession Takaragaike-2 [Tak2]) liverworts (*M. polymorpha* L.) were asexually maintained and propagated through the growth of gemmae as previously described (Ishizaki et al. 2008). The plants were grown on 1/2 Gamborg's B5 media containing 1% sucrose in a growth chamber at 20 °C under continuous light. Fifteen-day-old plants grown under continuous white light (40 μmol photons $m^{-2}s^{-1}$) at 20 °C were used for blue native-polyacrylamide gel electrophoresis (BN-PAGE), circularized-RNA reverse-transcription polymerase chain reaction (CR-RT-PCR), immunoblotting, northern blotting, quantitative RT-PCR (qRT-PCR), and rapid amplification of 5′ cDNA ends (5′-RACE). Rice *Tos17* mutants for OsSIG1 (a mutant with an insertion of the retrotransposon *Tos17* at the *OsSIG1* locus: NE8184) were grown on Murashige–Skoog media containing 3% sucrose in a growth chamber at 25 °C under a 16-h-light, 8-h-dark cycle. Total RNA was isolated from leaves of 9-day-old plants.

Database Analysis

All similarity searches were conducted using the default parameters of the Internet-based Basic Local Alignment Search Tool (Blast) available through the National Center for Biotechnology Information (NCBI). ClustalX (ftp://ftp-igbmc.u-strasbg.fr/pub/ClustalX/, last accessed September 25, 2013) was used for the multiple alignments of nucleotide sequences.

Phylogenetic Analysis of the Genes for Plastid Sigma Factors

The C-terminal conserved regions (region 2.1–4.2) of the sigma factors were used to construct a phylogenetic tree. The evolutionary history was inferred using the neighbor-joining method (Saitou and Nei 1987). The bootstrap consensus tree inferred from 1,000 replicates was taken to represent the evolutionary history of the analyzed taxa (Felsenstein 1985). The percentages of replicate trees in which the associated taxa clustered together in the bootstrap test (1,000 replicates) are shown next to the branches (Felsenstein 1985). The analysis involved 27 amino acid sequences. All of the positions

containing gaps and missing data were eliminated. There were a total of 210 positions in the final data set. The evolutionary analyses were conducted using MEGA4 (Tamura et al. 2007). The accession numbers for the 27 deduced amino acid sequences used in the phylogenetic analysis are shown in supplementary table S1, Supplementary Material online.

Transformation

Agrobacterium tumefaciens strain C58C1, transformed with the pCAMBIA1300 binary vector, was used to prepare the T-DNA-tagged lines. Transformations were conducted according to the methods of Ishizaki et al. (2008) and Chiyoda et al. (2008). The resulting spores were grown on 1/2 Gamborg's B5 media containing 10 μg/ml hygromycin B (Wako Pure Chemical, Osaka, Japan) for selection. For the complementation test, the hygromycin-resistant gene (*hpt*) encoded in pCAMBIA1300 was substituted with a mutated acetolactate synthase (mALS) gene conferring resistance to chlorosulfuron (Ishizaki et al., in preparation. Selections were conducted with 0.5 μM chlorosulfuron (DuPont, Wilmington, DE).

Genotyping

Total DNA was isolated from thalli using conventional methods and diluted in 200 μl of Tris–EDTA (ethylenediaminetetraacetic acid) buffer (10 mM Tris–HCl [pH 8.0], 1 mM EDTA [pH 8.0]). To confirm the position of the T-DNA insertion in the *Mpsig1* mutant and the success of the complementation, 1 μl from the extract was used in genomic PCR with primers P1, P2, and P3. All of the primers used in this article are shown in supplementary tables S2–S5, Supplementary Material online. Each primer position and the primer combinations for the genotyping of both types are shown in figure 1*A*. To confirm the position of the *tos17* insertion in the *Ossig1* mutant (NE8184), 1 μl from the extract was used for genomic PCR with primers P4, P5, and P6. Each primer position and the primer combinations for the genotyping of both types are

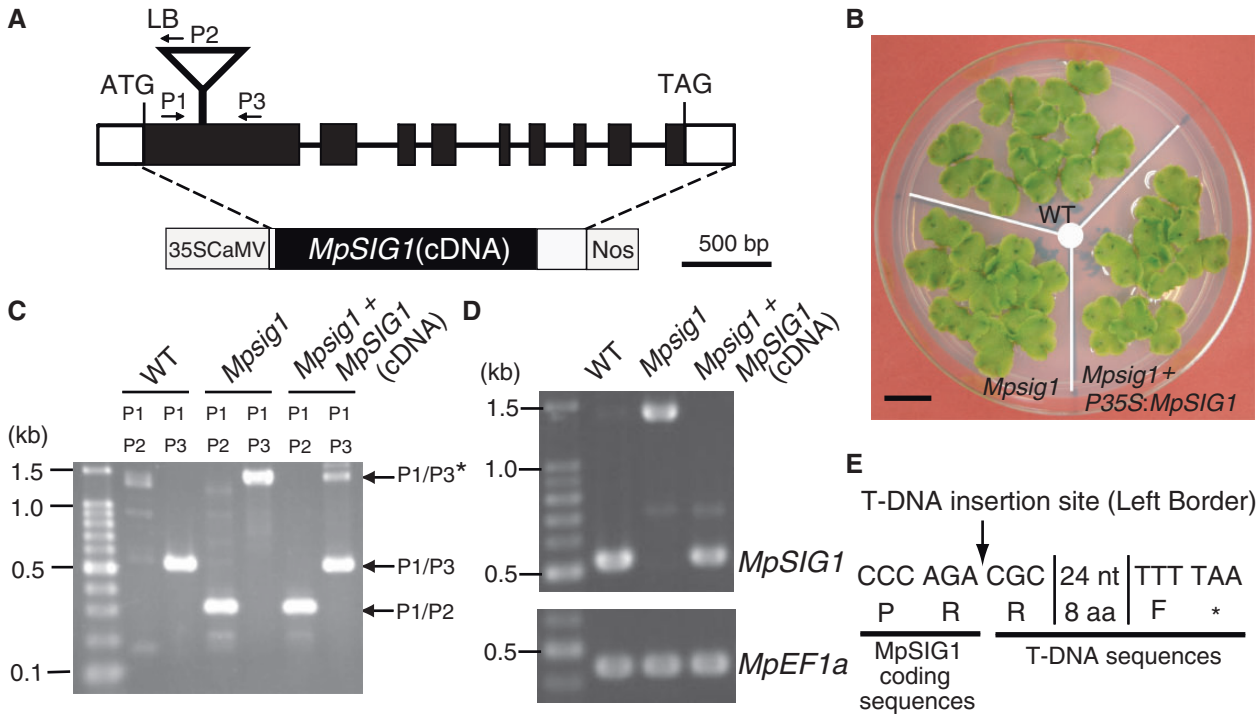

Fig. 1.—Tagging the gene for sigma factor 1 (*Mpsig1*) in liverwort (*Marchantia polymorpha*). (*A*) The exon–intron structure of *MpSIG1*. The black boxes, white boxes, and horizontal lines indicate exons, untranslated regions, and introns, respectively. The black arrows show the positions of the primers designed for genotyping. The position of a T-DNA insertion is shown with a triangle. The black arrows and the triangle are not to scale. The dotted lines show the region used for the complementation test. The sequence used for complementation, which is under the control of the cauliflower mosaic virus (CaMV) 35S promoter and the NOS transcription terminator, is shown in the lower boxes. (*B*) Morphological phenotype comparison between WT plants and the *Mpsig1* mutant. Fifteen-day-old plants were grown under continuous white light (40 μmol photons m^{-2} s^{-1}) at 20 °C. Scale bar: 1 cm. (*C*) Genotyping to confirm complementation of the *Mpsig1* mutant. The P1 and P3 (P1/P3 in the panel, 533 bp; P1/P3 marked by an asterisk in the panel, 1,493 bp containing a rearranged T-DNA) primer pair detects WT and complemented plants, and the P1 and P2 (P1/P2 in the panel, 321 bp) primer pair detects an *Mpsig1* mutant. The primer sequences are shown in supplementary table S2, Supplementary Material online. (*D*) Expression analysis of *MpSIG1* in plants. (*E*) The nucleotide and deduced amino acid sequences at the junction of the T-DNA insertion site. A nonsense mutation in the MpSIG1 reading frame, shown as an asterisk, was found around a T-DNA insertion site.

shown in figure 6*A*. The PCRs were conducted using KOD-FX DNA polymerase (TOYOBO, Osaka, Japan).

DNA Cloning of Reverse-Transcribed PCR (RT-PCR) and the Genomic PCR Products of *MpSIG1*

Total DNA and RNA were isolated from the thalli of plants using the Isoplant II kit (Nippon Gene, Tokyo, Japan) and the RNeasy Plant Mini Kit (Qiagen, Valencia, CA), respectively. The isolated RNA was further purified by incubation with RNase-free DNase I (Qiagen), according to the manufacturer's instructions. First-strand cDNA synthesis was performed using 1 μg of total RNA, 20 pmol of random hexamer primer, and 0.5 units of SuperScript III reverse transcriptase (Invitrogen, Carlsbad, CA). The resulting cDNAs and total DNA were used as templates to amplify *MpSIG1*. P7 and P8 were the primer pairs for *MpSIG1*. Genomic PCR and RT-PCR were conducted using KOD-plus-Neo DNA polymerase (TOYOBO). The resulting products, as well as products from experiments described later, were cloned into the pGEM-T Easy vector (Promega, Madison, WI) after the addition of adenosine according to the manufacturer's instructions, and the clones were sequenced using universal primers or randomly designed primers.

Thermal Asymmetric Interlaced PCR

Thermal asymmetric interlaced PCR (TAIL-PCR) was performed according to the protocol developed by Liu et al. (1995) with minor modifications. Total DNA purified by Isoplant II (Nippon Gene) was used as the template for PCR. The PCR was conducted using KOD-plus-Neo DNA polymerase (TOYOBO) with three nested primers: P9, P10, and P11. Each of these primers was paired with one of the two degenerate primers, P12 or P13. The secondary or tertiary PCR reactions were loaded on a 1% agarose gel. The specific bands were excised from the gel and eluted using a QIAquick PCR purification kit (Qiagen). The purified PCR products were subsequently cloned into the pGEM-T Easy vector.

5′-RACE

The 5′-RACE experiments were performed using a SMART RACE cDNA Amplification Kit (BD Bioscience Clontech, Palo Alto, CA) according to the manufacturer's instructions and KOD-plus-Neo DNA polymerase (TOYOBO). The gene-specific primers for the 5′-RACE PCR were P14 and P15.

Overexpression of Sigma Factor Genes in the *Mpsig1* Mutant

The coding sequences of *MpSIG1* and *MpSIG2* were amplified using the RT-PCR product or first-strand cDNA as templates with the following primers: *MpSIG1*, P16 and P17; *MpSIG2*, P18 and P19. Furthermore, for cloning into the pGWB2 binary vector, an in vitro reaction was performed according to the manufacturer's protocol, using LR clonase (Invitrogen). The resulting sequences, encoding a sigma factor protein under the control of the cauliflower mosaic virus (CaMV) 35S promoter and 3′ nopaline synthase (NOS) transcription terminator, were subcloned into pCAMBIA1300 (*hpt* was substituted for mALS).

Northern Blotting

Two micrograms of total RNA was fractionated through a 1% agarose/formaldehyde gel and blotted onto a Hybond-N^+ membrane (GE Healthcare, Buckinghamshire, UK). The gene-specific digoxigenin (DIG)-labeled probes were prepared for hybridization using DIG-11-dUTP (Jena BioScience GmbH, Jena, Germany). The primer pairs and sequences for preparing the DIG-labeled probes are listed in supplementary table S3, Supplementary Material online. Hybridization using DIG Easy Hyb buffer (Roche Diagnostics GmbH, Mannheim, Germany) was carried out at 50 °C overnight according to the manufacturer's instructions.

BN-PAGE and Immunoblot Analysis

BN-PAGE was performed as described previously (Peng et al. 2009). For these analyses, thylakoid proteins were loaded onto an SDS–polyacrylamide gel or blue native polyacrylamide gel on an equal chlorophyll basis. Samples of 0.5 μg chlorophyll thylakoid proteins were used for immunoblotting, except for the blot probed with the anti-NdhM antibody (Ueda et al. 2012). Chlorophyll thylakoid proteins of 2 and 6 μg were used for immunoblotting with anti-NdhM antibody and BN-PAGE, respectively. The signals were detected using an ECL Plus Western Blotting Detection Kit (GE Healthcare UK Ltd., Buckinghamshire, UK) and visualized with an LAS3000 chemiluminescence analyzer (Fuji Film, Tokyo, Japan). The results presented are representative of three independent experiments.

CR-RT-PCR

The CR-RT-PCR method was conducted according to the protocols of Kuhn and Binder (2002) and Forner et al. (2007) with minor modifications. After RNase-free DNase I treatment, 10 μg of total RNA was incubated with tobacco acid pyrophosphatase (Nippon Gene) at 37 °C for 1 h to improve the ligation efficiency. The dephosphorylated RNA was incubated with RNA ligase at 16 °C for 16 h (TaKaRa, Ohtsu, Japan). RT-PCR was conducted with the following primer pairs for each gene: liverwort *psaA*, P20 and P21; liverwort *ndhF*, P22 and P23; rice *psaA*, P24 and P25; and rice *ndhF*, P26 or P27 and P28. These primer pairs were designed from liverwort (GenBank accession numbers: X04465) and rice (GenBank accession numbers: AY522330) chloroplast genomes. These reactions were conducted according to the manufacturer's instructions.

qRT-PCR

The transcript levels were assayed using the Mx30000p qPCR instrument (Stratagene, La Jolla, CA) and FastStart Universal SYBR Green Master (Rox) reaction mix (Roche). The error bars show the standard deviations calculated from three repeated experiments. To standardize the data, the ratio of the absolute transcript level of each gene to the absolute transcript level of elongation factor 1-α (*EF-1a*) was calculated for each sample. The transcript abundance of genes is shown as the expression relative to the wild-type (WT) plant. The primer pairs were designed from the liverwort chloroplast genome (Ohyama et al. 1986) and their sequences for qRT-PCR are listed in supplementary tables S4 and S5, Supplementary Material online. The scores of the qRT-PCR are shown in supplementary table S6, Supplementary Material online. The experiments were conducted using male plants and were repeated twice. The genes for which qRT-PCR was conducted are identified by red bold letters in supplementary figure S1, Supplementary Material online, that was prepared using OrganellarGenomeDRAW (Lohse et al. 2007).

Results

Isolation of the Sigma Factor 1 (*Mpsig1*) Mutant in Liverwort

A liverwort mutant defective in the *SIG1* gene was identified serendipitously during the process of screening mutants for a different purpose. The sequence obtained by TAIL-PCR from a T-DNA left border was identical with an expressed sequence tag (EST; Genbank accession number: AB621976) in the population of T-DNA tagged lines. The EST encodes an open reading frame for a 574 amino acid (aa) protein that shows high similarity with plastidic sigma factor 1 proteins in land plants by a BlastX search (supplementary fig. S2, Supplementary Material online). Sequencing the 5′-RACE PCR products identified several transcription initiation sites for this gene, from −432 to −364 upstream of the translation initiation codon. As a result, MpSIG1 was confirmed to be composed of 574 residues because 5′-RACE did not detect additional N-terminal extensions (data not shown). On the basis of the high sequence similarity to land plant SIG1 proteins, we designated this gene *MpSIG1*.

Plant sigma factors typically consist of subdomains putatively involved in binding to the core RNA polymerase subunits (2.1 and 3), melting DNA (2.3), and recognizing promoters (−10 and −35 consensus motif) (2.4 and 4.2) (Vassylyev et al. 2002; Lysenko 2007). These subdomains were conserved in the C-terminal 327 aa of MpSIG1 (supplementary fig. S2, Supplementary Material online). We tested whether MpSIG1 is classified into the same group with sigma factor 1 from land plants. The exon–intron positions are completely conserved in the *Arabidopsis*, liverwort, rice, and *P. patens SIG1* genes (fig. 1*A*; supplementary fig. S2, Supplementary Material online). Phylogenetic analysis using the sigma factor conserved domain (region 2.1–4.2) showed that MpSIG1 was clustered with orthologs of SIG1 from flowering plants (supplementary fig. S3, Supplementary Material online). These results strongly suggest that MpSIG1 is a land plant SIG1 ortholog.

The coding sequence of *MpSIG1* consists of 1,722 nt, and the T-DNA was inserted 337 nt downstream of the translation initiation codon in the *Mpsig1* mutant (fig. 1*A* and *C*). RT-PCR did not detect any transcripts of *MpSIG1* in the mutant except for a 1,493-bp RT-PCR product that was not present in the WT plant (fig. 1*D*). This unusual transcript included the T-DNA with a large deletion and did not code for MpSIG1. Sequencing the RT-PCR product revealed that a nonsense mutation occurred around the T-DNA insertion site (fig. 1*E*), indicating that *MpSIG1* is a knockout allele of the gene.

Identification of Genes Regulated by *MpSIG1* by Genetic Analysis

Despite the complete loss of the functional *MpSIG1* gene, the mutant did not show any visible phenotypes (fig. 1*B*); however, transcription in chloroplasts might be affected to some extent. To assess this possibility, we measured the relative transcript abundance of 33 plastid-encoded genes expressed in the WT plant, the *Mpsig1* mutant and the *Mpsig1* mutant complemented by introduction of *MpSIG1* cDNA. Supplementary figure S1, Supplementary Material online, summarizes the genes used for the mRNA expression analysis. In rice, *OsSIG1* is involved in the transcription of at least three operons, *psaA/psaB/rps14*, *psbB/psbT/psbH/petB/petD*, and *psbE/psbF/psbL/psbJ* (Tozawa et al. 2007). In the *Mpsig1* mutant, the levels of *petB*, *psaA*, *psaB*, *psbB*, *psbE*, *psbF*, and *rps14* were mildly reduced (fig. 2*A*). The accumulation levels of these transcripts were approximately 60–90% of the WT plants (supplementary table S6, Supplementary Material online), suggesting that the function of MpSIG1 overlaps with that of other sigma factors in the WT plant and/or that its function is partially complemented by other sigma factors in the mutant. In addition, more than a 30% reduction was observed for three other genes (*ndhF*, *psbK*, and *rps18*) (fig. 2*B*). The transcript abundances of *ndhF*, *rps18* (*petL* operon: *petL/petG/psaJ/rpl33/rps18*), and *psbK* (*psbK* operon: *psbK/psbI*) were also reduced to 48%, 65%, and 67% compared with those of WT plants, respectively (supplementary table S6, Supplementary Material online).

To assess whether the *Mpsig1* mutant phenotype observed in chloroplast RNA was due to the lack of *MpSIG1*, the mutant was transformed with a WT copy of *MpSIG1* cDNA controlled by the CaMV 35S promoter and the NOS transcription terminator. Compared with the WT plants, the complemented lines accumulated approximately a 19-fold higher level of the *MpSIG1* transcript (fig. 2*C*). The overexpression of *MpSIG1* did not cause any visible abnormality (fig. 1*B*). The transcript levels of the genes that were downregulated in the *Mpsig1*

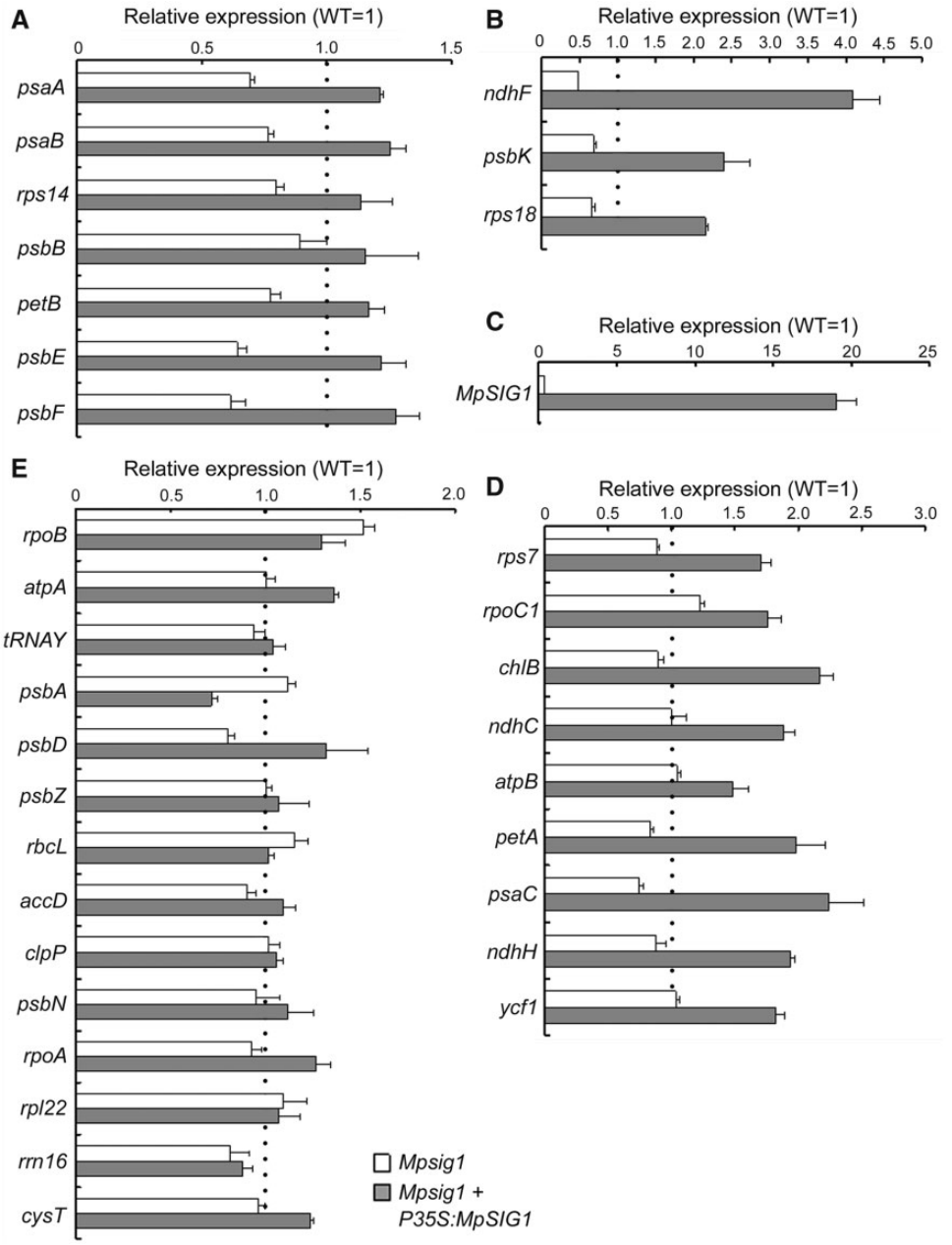

Fig. 2.—The levels of plastid transcripts in the *Mpsig1* mutant and its complemented lines. The abundance of mRNAs, tRNAs, and rRNA transcripts was measured in the *Mpsig1* mutant using qRT-PCR. (*A*, *B*, *D*, and *E*) Expression levels of chloroplast-encoded genes. (*A*) Genes that are downregulated in the *Mpsig1* mutant. (*B*) Genes downregulated by more than 30% in the *Mpsig1* mutant. (*C*) Expression levels of *MpSIG1* in the nucleus. (*D*) Genes upregulated by at least 1.5-fold in the complemented *Mpsig1* mutant. (*E*) Other genes. The standard deviations ($n = 3$, n stands for technical replicates) are indicated by lines extending from the bars. Each mean represents the ratio of the expression level of transcripts in the *Mpsig1* mutant or complemented plants compared with that of WT plants. The means are depicted by white and gray bars for the *Mpsig1* mutant (*Mpsig1*) and complemented plant (*Mpsig1* + *P35S:MpSIG1*), respectively. Fifteen-day-old plants grown under continuous white light (40 μmol photons $m^{-2}s^{-1}$) at 20 °C were used.

mutant became even higher in the complemented lines than in the WT plants (fig. 2*A* and *B*). In particular, *ndhF*, *rps18*, and *psbK* transcript levels were more than twice as high as those in the WT plants, probably because of the overexpression of MpSIG1 (fig. 2*B*). Furthermore, more than 1.5-fold upregulation in overexpressors was observed for nine transcripts (*atpB*, *chlB*, *ndhC*, *ndhH*, *petA*, *psaC*, *rpoC1*, *rps7*, and *ycf1*) that were not significantly affected in the *Mpsig1* mutant (fig. 2*D*). The validity of the qRT-PCR was confirmed by performing Northern hybridization analysis for six selected transcripts (*ndhF*, *psaA*, *psbB*, *psbD*, *psbE/psbF*, and *rpl33/rps18*) (fig. 3). On the basis of the phenotypes of the mutant and

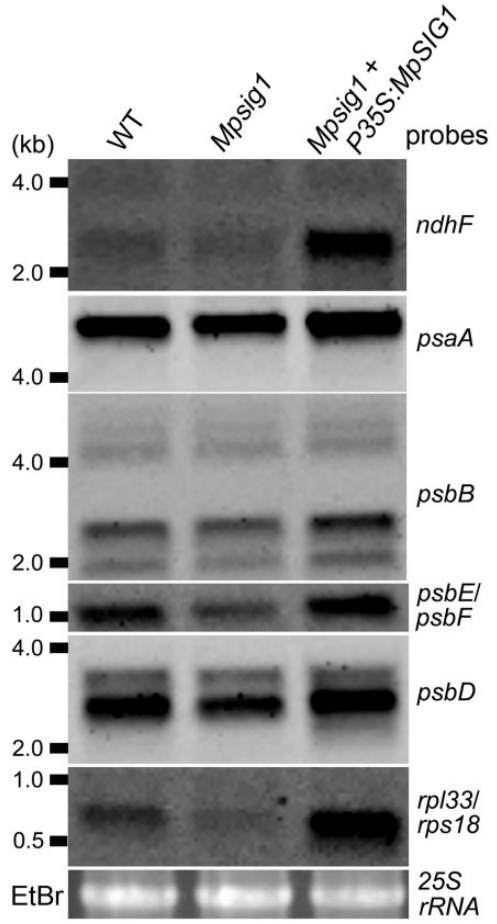

Fig. 3.—RNA gel blot hybridizations for chloroplast-encoded genes in the *Mpsig1* mutant. Specific probes for each gene were prepared using primer pairs listed in supplementary table S3, Supplementary Material online. Fifteen-day-old plants grown under continuous white light (40 μmol photons $m^{-2}s^{-1}$) at 20 °C were used. Equal loading was confirmed by comparing the ethidium bromide staining of rRNA.

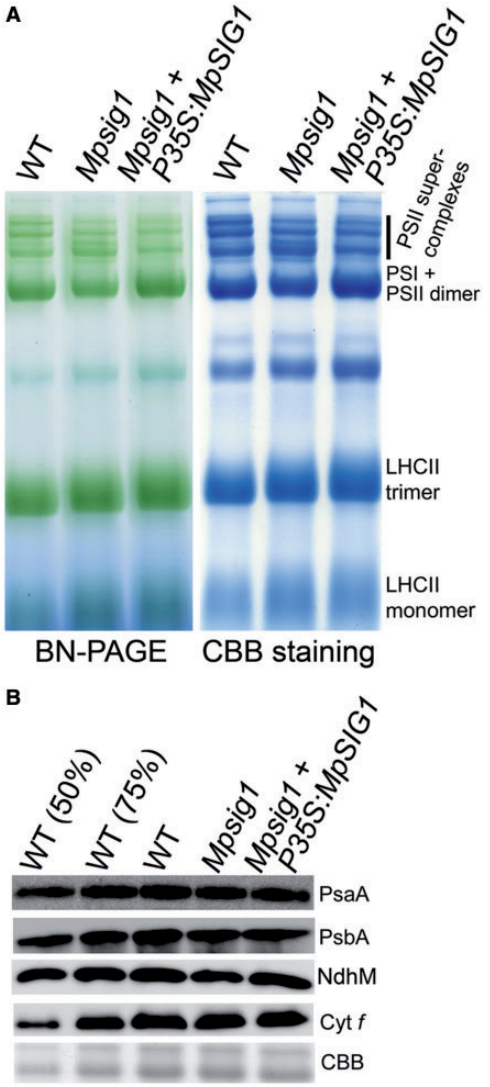

Fig. 4.—Protein accumulation in thylakoids compared between WT plants and the *Mpsig1* mutant plants. (*A*) BN-PAGE analysis. Samples containing 6 μg chlorophyll thylakoid proteins were loaded. (*B*) Immunoblot analysis. Samples containing 0.5 μg chlorophyll thylakoid proteins were loaded. The proteins were separated by sodium dodecyl sulphate (SDS)-PAGE, and the blots were probed with specific antibodies against PsaA (a subunit of photosystem I), PsbA (a subunit of photosystem II), NdhM (a subunit of the NAD(P)H dehydrogenase [NDH] complex), and Cyt *f* (a subunit of the cytochrome b_6f complex). Fifteen-day-old plants grown under continuous white light (40 μmol photons $m^{-2}s^{-1}$) at 20 °C were used for these analyses.

overexpressors (complemented lines), it is highly likely that MpSIG1 is involved in the transcription of a wide range of transcripts and that its function overlaps with and/or is complemented by those of other sigma factors. We could not eliminate the possibility that some of the mutant phenotypes observed in the RNA expression levels are due to a secondary effect from disturbance of the sigma factor network.

Accumulation of Thylakoid Membrane Proteins in the *Mpsig1* Mutant

To analyze the influence of lower mRNA levels on protein accumulation in the *Mpsig1* mutant, BN-PAGE analysis of the protein complexes in the thylakoid membrane was conducted for the *Mpsig1* mutant and WT plants. There was no significant difference between the *Mpsig1* mutant and WT plants (fig. 4*A*).

Similar results were obtained in the immunoblot analysis of PsaA (a subunit of photosystem I), PsbA (a subunit of photosystem II), NdhM (a subunit of the NADH dehydrogenase-like complex), and Cyt *f* (a subunit of the cytochrome b_6f complex) proteins. Consistent with the transcript abundance results and morphological phenotype (figs. 1*B*, 2, and 3), there was no significant difference in the accumulation level of these

proteins among WT plants, the complemented *Mpsig1* mutant and the *Mpsig1* mutant (fig. 4*B*).

Effect of *MpSIG2* Overexpression on Liverwort Plastid Gene Expression

The conserved intron–exon structure suggests that the common ancestor of angiosperm sigma factors was initially duplicated into SIG1 and SIG2 (Lysenko 2007). Transcription initiation activity of the *psbA* and *rbcL* promoters in *Arabidopsis* depends on sigma factors; SIG1 drives much less activity at these promoters than SIG2 and SIG3 (Hakimi et al. 2000; Privat et al. 2003). Next, we surveyed whether there are differences in target preference between the SIG1 and SIG2 proteins in liverwort, using overexpressed lines. An EST clone encoding *MpSIG2* was identified in an EST library (the library constructed by the Kohchi laboratory), and the deduced amino acid sequence of MpSIG2 clustered with orthologs of SIG2 proteins from land plants (supplementary fig. S3, Supplementary Material online).

MpSIG2 was overexpressed under the control of the CaMV 35S promoter in the *Mpsig1* mutant, and its effects on various transcript levels were analyzed by qRT-PCR. The overexpressing lines accumulated approximately a 3-fold higher level of *MpSIG2* transcripts than the WT plants and did not show any visible phenotypic difference. *MpSIG2* overexpression resulted in an effect similar to that of *MpSIG1* overexpression on the transcript levels of *psaA*, *psbE*, *psbK*, and *rps18*, suggesting that both sigma factors have roughly the same affinity for these promoters. In contrast, a drastic difference was observed in the effect on the *ndhF* transcript level. Although MpSIG1 highly up-regulated *ndhF* transcript accumulation, the effect was subtle in the MpSIG2 overexpression lines (fig. 5), suggesting a different promoter preference between the two sigma factors. Overexpression of *MpSIG2* upregulated the expression levels of *psbA* and *rbcL* slightly more than that of *MpSIG1* (fig. 5). Collectively, these results suggest a slight difference in the promoter preferences between MpSIG1 and MpSIG2; however, the lack of a clear mutant phenotype in the *Mpsig1* mutant implies that sigma factors may share many targets in liverwort and that each sigma factor may have acquired more specific functions during the further evolution of land plants.

Identification of Genes Regulated by *OsSIG1*

Our study revealed that MpSIG1 was involved in the expression of *ndhF*, *rps18* (*petL* operon: *petL/petG/psaJ/rpl33/rps18*), and *psbK* (*psbK* operon: *psbK/psbI*). In a previous study, the effect of OsSIG1 on the expression of these three genes was not reported (Tozawa et al. 2007). To test whether land plant SIG1 proteins generally regulate the expression of these three genes, their transcript levels were analyzed in an *Ossig1* mutant (a mutant with an insertion of the retrotransposon *Tos17* at the *OsSIG1* locus: NE8184) by qRT-PCR analysis

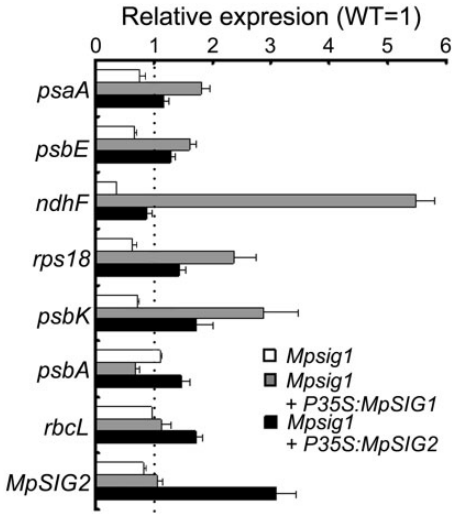

Fig. 5.—The effect of overexpressing liverwort sigma factor 2 (*MpSIG2*) on plastid transcript abundance in the *Mpsig1* mutant. *MpSIG2* was overexpressed under the control of the CaMV 35S promoter in the *Mpsig1* mutant. Fifteen-day-old plants grown under continuous white light (40 μmol photons $m^{-2}s^{-1}$) at 20 °C on agar plates were used. The means are depicted by white, gray, and black bars for the *Mpsig1* mutant (*Mpsig1*), plants overexpressing *MpSIG1* (*Mpsig1* + *P35S:MpSIG1*), and plants overexpressing *MpSIG2* (*Mpsig1* + *P35S:MpSIG2*), respectively. The standard errors ($n = 3$, n stands for technical replicates) are indicated by lines extending from the bars. Each mean represents the ratio of the expression level of transcripts in the experimental plants compared with that of WT plants.

(fig. 6). Gene expression of *psaA* in the *Ossig1* mutant was previously reported to be 11–36% compared with the WT plants (Tozawa et al. 2007). In our study, *psaA* transcripts in the *Ossig1* mutant were 38% of the transcript levels of the *Ossig1* mutant (fig. 6*D*; supplementary table S6, Supplementary Material online). Thus, we confirmed that the *psaA* transcript level in the *Ossig1* mutant under our growth conditions was reduced to the same extent as previously reported.

Next, transcript levels of the *psbK* and *ndhF* genes in the *Ossig1* mutant were evaluated by qRT-PCR and found to be 30% and 21% lower than WT, respectively (fig. 6*D*). This result suggests that MpSIG1 and OsSIG1 proteins essentially share their target promoters (*psaA*, *psbB*, *psbE*, and *psbK* operons; and *ndhF*), but OsSIG1 is no longer the predominant regulator of *ndhF* expression in contrast to MpSIG1.

Ossig1 and *Mpsig1* mutants showed different phenotypes in the abundance of *rps18* transcripts. A 1.2-fold increase in the *rps18* transcript level was observed in the *Ossig1* mutant. *rps18* is classified as a Class II gene regulated by NEP and PEP according to Ishizaki et al. (2005). The enhanced accumulation of *rps18* transcript might be induced by NEP (see the Discussion section).

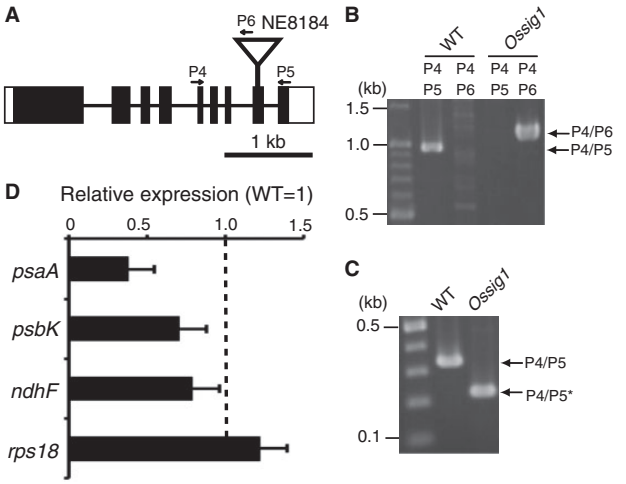

Fig. 6.—Plastid transcript levels in the *Ossig1* mutant. (*A*) The exon–intron structure of *OsSIG1*. The black boxes, white boxes, and horizontal lines indicate exons, untranslated regions, and introns, respectively. The black arrows show the positions of the primers designed for genotyping. The position of a *tos17* insertion in the OsSIG1 locus of line NE8184 is shown with a triangle. The black arrows and the triangle are not to scale. (*B*) Genotyping in the *Ossig1* mutant. The P4 and P5 (P4/P5 in the panel) primer pair detects WT plants, and the P4 and P6 (P4/P6 in the panel) primer pair detects the *tos17* insertion. The primer sequences are shown in supplementary table S2, Supplementary Material online. (*C*) Expression analysis of *OsSIG1* by RT-PCR in plants. A miss-spliced RT-PCR product was detected in the *Ossig1* mutant as previously reported (Tozawa et al. 2007). (*D*) The abundance of mRNAs was measured in the *Ossig1* mutant using qRT-PCR.

Identification of the *ndhF* Transcription Initiation Sites by CR-RT-PCR in Liverwort and Rice

In addition to rice, the transcription of *ndhF* was not predominantly regulated by SIG1 in *Arabidopsis* because *ndhF* transcription mainly depends on AtSIG4 (Favory et al. 2005). Why has SIG1 specifically lost its ability to control the transcription of *ndhF*?

The rapid evolution of *ndhF* promoter sequences in flowering plants has been already reported (Seliverstov et al. 2009), suggesting a possible explanation that the occurrence of extreme divergences in the *ndhF* promoter region during the evolution of land plants resulted in the loss of SIG1's regulation of *ndhF* expression and led to AtSIG4-specific regulation in *Arabidopsis*. To test this possibility, CR-RT-PCR was conducted to identify the unknown transcription initiation site of *ndhF* in liverwort and rice.

First, a control experiment was performed to check the efficacy of first-strand cDNA preparation for CR-RT-PCR. The transcription initiation sites of *psaA* in liverwort and rice are surrounded by consensus PEP recognition motifs that are highly conserved among land plants. The 5′-end of the *psaA* CR-RT-PCR product was mapped to 148-nt and 129-nt upstream from the translation start codon of *psaA* in liverwort and rice, respectively (figs. 7*C* and *D*). The −35 and −10 consensus motifs are highly conserved in this region, suggesting that the promoter is conserved between liverwort and flowering plants and validating our CR-RT-PCR (fig. 7*E*).

Next, we used CR-RT-PCR to evaluate liverwort and rice *ndhF*. In liverwort, two 3′-ends and a single 5′-end of *ndhF* mRNA were identified (fig. 7*A*). The 5′-end of *ndhF* CR-RT-PCR products was mapped to only 135-nt upstream of the translation start codon, although the 3′-ends were located at 580- and 316-nt downstream of the stop codon (fig. 7*C*). In rice, the CR-RT-PCR for *ndhF* revealed multiple 3′- and 5′-ends of *ndhF* mRNA (figs. 7*B* and *D*). Among them, convincing −10 and −35 element-like motifs were identified from only the 5′-end of *ndhF* mRNA that mapped to 617-nt upstream of the translation start codon (figs. 7*D* and *E*). In Poaceae, the putative promoter sequences of *ndhF* are highly conserved. Four representative Poaceae species are shown in figure 7*F* and 24 Poaceae species whose chloroplast genome sequences are publically available are shown in supplementary figure S4 (Supplementary Material online). The other 5′-ends of *ndhF* mRNA might be derived from either a NEP-dependent transcript or a processed form. The CR-RT-PCR experiments revealed that the primary structure surrounding transcription initiation sites of *ndhF* were different from those of *ndhF* mRNA in *Arabidopsis*, liverwort, and rice (fig. 7*E*). This result, based mainly on CR-RT-PCR, indicates that the transcription initiation sites of *ndhF* are highly diversified among land plants. Additional experiments using other techniques, such as 5′ RACE or next generation RNA sequencing, would be required to confirm this possibility.

Discussion

Subfunctionalization of SIG1 during Land Plant Divergence

The duplication–degeneration–complementation (DDC) model implies that the usual mechanism of duplicated gene preservation is the partitioning of ancestral functions (subfunctionalization). As a result, both duplicated genes are required for the ancestral function when subfunctionalization is completed (Force et al. 1999). This model is well suited to the expansion of the sigma factor family during the evolution of land plants. In this study, we detailed the process of *SIG1* subfunctionalization during land plant evolution, noting that the subfunctionalization of *SIG1* is ongoing in a bryophyte (liverwort). The more advanced subfunctionalization of sigma factor genes seems to facilitate gene-specific regulation in angiosperm chloroplast genomes.

The DDC model was originally elaborated to explain nuclear gene duplication, followed by degenerative mutations in nuclear regulatory elements. Our analysis showed that this model could also explain the subfunctionalization of sigma

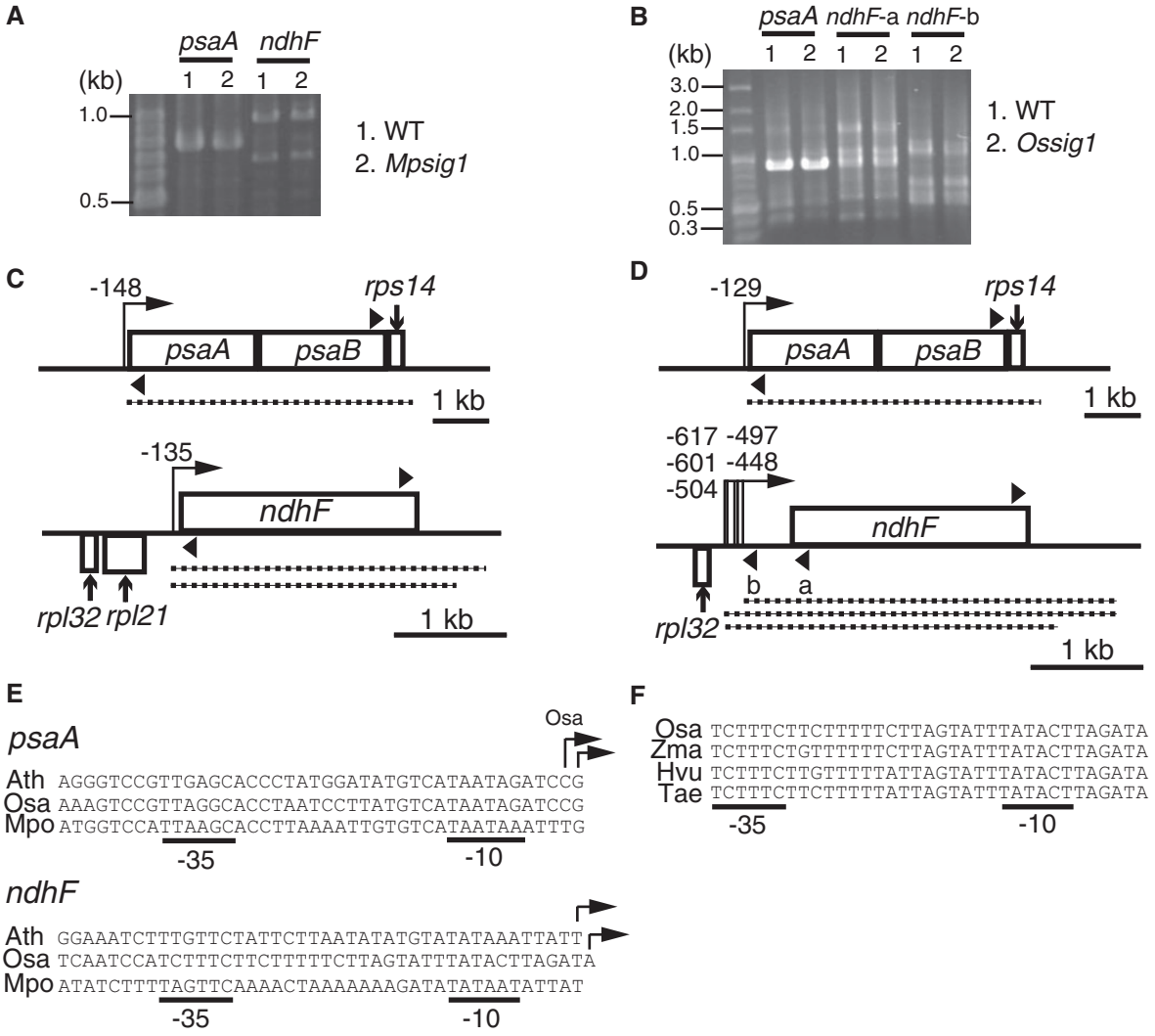

Fig. 7.—CR-RT-PCR experiments on the *psaA* and *ndhF* genes in the liverwort and rice chloroplast genomes. CR-RT-PCR amplification of the chloroplast-encoded *psaA* and *ndhF* genes in liverwort (*A*) and rice (*B*). A schematic representation of the CR-RT-PCR results in liverwort (*C*) and rice (*D*). Arrows with angled lines and arrowheads with dotted lines show the transcription initiation sites and transcribed regions, respectively, for *psaA* and *ndhF* determined by CR-RT-PCR. The black triangles, which are not to scale, show the primer positions for CR-RT-PCR. The primer sequences are shown in supplementary table S2, Supplementary Material online. (*E*) An alignment of the *psaA* and *ndhF* regions upstream from the 5′-ends of the transcripts. The transcription initiation sites for angiosperm *psaA* and *Arabidopsis* were obtained from Tozawa et al. (2007) and Favory et al. (2005), respectively. Underlined letters indicate the −10 and −35 putative consensus motifs. (*F*) Comparison of nucleotide sequences in the promoter regions of *ndhF* in four representative Poaceae species. Ath, *Arabidopsis thaliana*; Hvu, *Hordeum vulgare*; Mpo, *Marchantia polymorpha*; Osa, *Oryza sativa*; Tae, *Triticum aestivum*; Zma, *Zea mays*.

factor protein genes during plant evolution. In the case of sigma factor protein genes, however, the DCC model should explain the alteration of promoter elements in the plastid genome, adding some complexities to the model. Indeed, genomic deletions in the promoter region of *ndhF* were observed that should affect this regulatory element (fig. 7). Plastid-specific processes should be taken into account in a modified DDC model to explain the evolution of sigma factor protein genes.

Different phenotypes have been observed among the *sig1* mutants in liverwort, rice, and *Arabidopsis* although they essentially share the same targets (Tozawa et al. 2007; Hanaoka et al. 2012), which suggests different levels of *SIG1* subfunctionalization arose during the divergence of land plants. For example, in liverwort, the *Mpsig1* mutant showed no visible abnormalities (fig. 1*B*), but its transcript levels were altered, especially in some plastid-encoded genes (*ndhF* and genes in the *psaA*, *psbB*, *psbE*, *psbK*, and *petL* operons). This finding suggests that *MpSIG1* is involved in the transcription of these genes, although the subfunctionalization of SIG1 for their gene expression has not been completed and its function may overlap with those of other sigma

factors in liverwort. In bryophytes, the *SIG1*, *SIG2*, and *SIG5* genes appear to be encoded in the nucleus (Shiina et al. 2009). The phenotype of the *Mpsig1* mutant suggests that the function of *MpSIG1* is at least partly complemented by *MpSIG2* and *MpSIG5*, and also possibly by NEP. The *Atsig1* mutant showed no visible abnormalities either (Lai et al. 2011; Woodson et al. 2012). On the other hand, the rice *Ossig1* mutant could not transcribe the *psaA*, *psbB*, and *psbE* operons, resulting in a pale green phenotype (Tozawa et al. 2007). Transcription of these genes relies on *SIG1* in rice, in particular genes included in the *psaA* operon, indicating that profound subfunctionalization of *SIG1* occurred in rice.

Posttranscriptional mechanisms for organellar gene expression that are mediated by numerous factors affecting RNA stability, processing or translation have been more popular subjects for investigation (Schmitz-Linneweber and Small 2008; Woodson and Chory 2008; Stern et al. 2010; Barkan 2011). However, our study suggests that subfunctionalization of sigma factor protein that act in transcription has also contributed to gene-specific regulation during the evolution of land plants.

A Transcription Network Mediated by Sigma Factors and NEP in Land Plants

Because most plastid genes appear to be controlled by several sigma factors with overlapping functions (Liere et al. 2011), *MpSIG1* deficiency may be compensated by other sigma factors. Furthermore, NEP also participates in the transcription network in angiosperm plastids. In sigma factor mutants such as *Ossig1* and *Atsig6* in angiosperms, the defect upregulates the expression of NEP-dependent genes (e.g., *rpoB* and *rps18*; Ishizaki et al. 2005; Tozawa et al. 2007). Although an increase in transcript abundance of *rpoB* was observed, the expression of *rps18* was reduced in the *Mpsig1* mutant (figs. 2*B* and *E* and 3). We could not conclude that NEP upregulates *rpoB* expression in the *Mpsig1* mutant because plastid-localized NEP has not yet been reported in liverwort. Further investigation on NEP is needed to comprehensively understand the transcription network in liverwort plastids.

Did Gene Loss Give Rise to a Novel Promoter for *ndhF* during Land Plant Evolution?

Seliverstov et al. (2009) have already analyzed the *ndhF* promoter sequences in flowering plants in detail. These investigators suggested that the angiosperm *ndhF* promoter sequences rapidly evolved. Our study concurs with their study and suggests dynamic reassignment of the *ndhF* promoter during the evolution of land plants. Why did land plants diversify in the *ndhF* promoter region?

Considering the reorganization of gene content upstream of *ndhF*, gene loss from the chloroplast genome provides a plausible explanation for the divergence of the *ndhF* promoter in land plants. *rpl21* is encoded in the flanking region of *ndhF* in liverwort (Ohyama et al. 1986) and has been generally lost from angiosperm chloroplast genomes (Martin et al. 2002). Despite the loss of *rpl21* from angiosperm chloroplast genomes, the length of the intergenic region between the *ndhF* and *rpl32* genes where *rpl21* is encoded is relatively conserved among land plant chloroplast genomes (e.g., *Arabidopsis*, 811 nt; liverwort, 707 nt; and rice, 716 nt). Generally, *ndhF* promoters are found up to position −195 upstream of the translation start codon of *ndhF* in angiosperms, as in liverwort (Seliverstov et al. 2009). On the other hand, *ndhF* promoters are located at around position −320 upstream of its translation start codon in members of the Brassicaceae family (Favory et al. 2005; Seliverstov et al. 2009). In the Poaceae, *ndhF* promoters were found further upstream of their translation start codon (fig. 7*D*). The original *ndhF* promoters have been displaced by novel promoters in the Brassicaceae and Poaceae (supplementary fig. S5, Supplementary Material online). These results suggest that the course of decay of *rpl21* coding sequences and the retained intergenic region between the *ndhF* and *rpl32* genes might have provided an opportunity to generate novel promoters for *ndhF* and allowed another sigma factor protein such as SIG4 or NEP to regulate the expression of *ndhF*.

Influence of Phosphorylation Status on Promoter-Binding Specificity of Sigma Factor Proteins

In *Arabidopsis*, the AtSIG1-dependent expression of *psaA* plays an important role in controlling the balance of photosynthetic efficiency via phosphorylation (Shimizu et al. 2010). The phosphorylation of AtSIG1 controlled by the chloroplast sensor kinase suppresses the expression of genes for photosystem I to adjust photosystem stoichiometry when the plastoquinone (PQ) pool becomes oxidized (Puthiyaveetil et al. 2010, 2013). In addition, sigma factor proteins contribute to circadian modulation of photosynthesis gene expression (references in Noordally et al. 2013). For example, 70% of protein-coding genes encoded by the chloroplast genome can be circadian regulated in *Arabidopsis* (Noordally et al. 2013). Phosphorylation plays an important role in the plant circadian system (Kusakina and Dodd 2012). Experimental data and bioinformatic predictions indicate that there are at least 20 chloroplast-localized rhythmic protein kinases that are encoded by the nuclear genome (Bayer et al. 2012). Various kinases might modify the phosphorylation status of SIG1 under different environmental conditions, and the phosphorylation status of SIG1 might change its specificity in promoter binding and recognition. In our study, the phosphorylation status of MpSIG1 was not tested. Analysis of the promoter recognition specificity of MpSIG1 under different phosphorylation conditions induced by redox shifts in the PQ pool or different light/dark cycles will facilitate our understanding of how chloroplast gene expression mediated by sigma factor proteins evolved in land plants.

Why Did *SIG4* Diverge from *SIG1*?

SIG4 is present in the Brassicaceae, including *Arabidopsis*, but not in monocotyledonous plants, such as rice and *Zea mays*. *SIG4* seems to have evolved after the divergence of the mono- and dicotyledonous plants and to be specifically retained in the Brassicaceae lineage because *SIG4* becomes a pseudogene in *Populus* sp. that are also dicotyledonous plants (Seliverstov et al. 2009).

Our study suggests that *AtSIG4* was functionally partitioned from the ancestral *SIG1*. What was the reason for the functional partitioning of *SIG4* from *SIG1*? In *Arabidopsis*, the transcription of *ndhF* depends almost exclusively on *AtSIG4* (Favory et al. 2005). The *ndhF* gene encodes a subunit of the chloroplast NADH dehydrogenase-like complex (NDH) that functions in photosystem I cyclic electron transport (Peng et al. 2011). An attractive idea is that the NDH level has to be regulated independently of the levels of the photosystems and that *SIG4* is involved in this process. However, chloroplast NDH consists of more than 25 subunits (Ifuku et al. 2011) and it is unclear whether the transcription of *ndhF* limits the accumulation of NDH.

Our results suggest expansion of the sigma factor family during the evolution of land plants, facilitating a variety of transcriptional regulatory processes in plastids. However, the physiological requirement for this evolution is not fully understood. Further analysis of the function and evolution of *MpSIG1* would provide deeper insight into this intriguing question.

Supplementary Material

Supplementary figures S1–S5 and tables S1–S7 are available at *Genome Biology and Evolution* online (http://www.gbe.oxfordjournals.org/).

Acknowledgments

We thank Dr M. Takemura and Dr K. Ohyama in Ishikawa Prefecture University for the tips on cultivation techniques of liverwort. We are also grateful for their kind distributions of chlorosulfuron, pGWB gateway vectors, and *tos17* insertion mutant rice seeds from Dupont and Dr T. Nakagawa in Shimane University, and the National Institute of Agrobiological Sciences (NIAS) in Japan, respectively. This work was supported by grant number 08062811 from the PRESTO (Precursory Research for Embryonic Science and Technology), the Global Center of Excellence Program "Formation of a Strategic Base for Biodiversity and Evolutionary Research: from Genome to Ecosystem" of MEXT, grant number GS015 from the Funding Program for Next Generation World-Leading Researchers (NEXT Program) (to Y.N.), and grant number 21-764 from the research fellowship of the Japan Society for the Promotion of Science for young scientists (to M.U.).

Literature Cited

Barkan A. 2011. Expression of plastid genes: organelle-specific elaborations on a prokaryotic scaffold. Plant Physiol. 155:1520–1532.

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

Bayer RG, et al. 2012. Chloroplast-localized protein kinases: a step forward towards a complete inventory. J Exp Bot. 63:1713–1723.

Bohne AV, Irihimovitch V, Weihe A, Stern DB. 2006. *Chlamydomonas reinhardtii* encodes a single sigma70-like factor which likely functions in chloroplast transcription. Curr Genet. 49:333–340.

Carter ML, Smith AC, Kobayashi H, Purton S, Herrin DL. 2004. Structure, circadian regulation and bioinformatic analysis of the unique sigma factor gene in *Chlamydomonas reinhardtii*. Photosynth Res. 82: 339–349.

Chiyoda S, Ishizaki K, Kataoka H, Yamato KT, Kohchi T. 2008. Direct transformation of the liverwort *Marchantia polymorpha* L. by particle bombardment using immature thalli developing from spores. Plant Cell Rep. 27:1467–1473.

Favory JJ, et al. 2005. Specific function of a plastid sigma factor for *ndh*F gene transcription. Nucleic Acids Res. 33:5991–5999.

Felsenstein J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39:783–791.

Force A, et al. 1999. Preservation of duplicate genes by complementary, degenerative mutations. Genetics 151:1531–1545.

Forner J, Weber B, Thuss S, Wildum S, Binder S. 2007. Mapping of mitochondrial mRNA termini in *Arabidopsis thaliana*: t-elements contribute to 5′ and 3′ end formation. Nucleic Acids Res. 35:3676–3692.

Fujiwara M, Nagashima A, Kanamaru K, Tanaka K, 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.

Gray MW. 1992. The endosymbiont hypothesis revisited. Int Rev Cytol. 141:233–357.

Hakimi MA, Privat I, Valay JG, Lerbs-Mache S. 2000. Evolutionary conservation of C-terminal domains of primary sigma(70)-type transcription factors between plants and bacteria. J Biol Chem. 275: 9215–9221.

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

Hanaoka M, Kato M, Anma M, 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.

Hara K, Sugita M, 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.

Hara K, et al. 2001. Characterization of two genes, *Sig1* and *Sig2*, encoding distinct plastid sigma factors(1) in the moss *Physcomitrella patens*: phylogenetic relationships to plastid sigma factors in higher plants. FEBS Lett. 499:87–91.

Ichikawa K, Sugita M, Imaizumi T, Wada M, 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.

Ifuku K, Endo T, Shikanai T, Aro EM. 2011. Structure of the chloroplast NADH dehydrogenase-like complex: nomenclature for nuclear-encoded subunits. Plant Cell Physiol. 52:1560–1568.

Ishizaki K, Chiyoda S, Yamato KT, Kohchi T. 2008. *Agrobacterium*-mediated transformation of the haploid liverwort *Marchantia polymorpha* L., an emerging model for plant biology. Plant Cell Physiol. 49: 1084–1091.

Ishizaki Y, 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.

Isono K, et al. 1997. Leaf-specifically expressed genes for polypeptides destined for chloroplasts with domains of sigma70 factors of bacterial RNA polymerases in *Arabidopsis thaliana*. Proc Natl Acad Sci U S A. 94: 14948–14953.

Kanamaru K, et al. 2001. An *Arabidopsis* sigma factor (SIG2)-dependent expression of plastid-encoded tRNAs in chloroplasts. Plant Cell Physiol. 42:1034–1043.

Kasai K, et al. 2004. Differential expression of three plastidial sigma factors, *OsSIG1*, *OsSIG2A*, and *OsSIG2B*, during leaf development in rice. Biosci Biotechnol Biochem. 68:973–977.

Kubota Y, et al. 2007. Two novel nuclear genes, *OsSIG5* and *OsSIG6*, encoding potential plastid sigma factors of RNA polymerase in rice: tissue-specific and light-responsive gene expression. Plant Cell Physiol. 48:186–192.

Kuhn J, Binder S. 2002. RT-PCR analysis of 5′ to 3′-end-ligated mRNAs identifies the extremities of *cox2* transcripts in pea mitochondria. Nucleic Acids Res. 30:439–446.

Kusakina J, Dodd AN. 2012. Phosphorylation in the plant circadian system. Trends Plant Sci. 17:575–583.

Lai Z, et al. 2011. *Arabidopsis* sigma factor binding proteins are activators of the WRKY33 transcription factor in plant defense. Plant Cell 23: 3824–3841.

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.

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

Liu YG, Mitsukawa N, Oosumi T, Whittier RF. 1995. Efficient isolation and mapping of *Arabidopsis thaliana* T-DNA insert junctions by thermal asymmetric interlaced PCR. Plant J. 8:457–463.

Lohse M, Drechsel O, Bock R. 2007. OrganellarGenomeDRAW (OGDRAW): a tool for the easy generation of high-quality custom graphical maps of plastid and mitochondrial genomes. Curr Genet. 52:267–274.

Lysenko EA. 2007. Plant sigma factors and their role in plastid transcription. Plant Cell Rep. 26:845–859.

Martin W, 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.

Nagashima A, 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.

Noordally ZB, et al. 2013. Circadian control of chloroplast transcription by a nuclear-encoded timing signal. Science 339:1316–1319.

Ohyama K, et al. 1986. Chloroplast gene organization deduced from complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA. Nature 322:572–574.

Peng L, Fukao Y, Fujiwara M, Takami T, Shikanai T. 2009. Efficient operation of NAD(P)H dehydrogenase requires supercomplex formation with photosystem I via minor LHCI in *Arabidopsis*. Plant Cell 21: 3623–3640.

Peng L, Yamamoto H, Shikanai T. 2011. Structure and biogenesis of the chloroplast NAD(P)H dehydrogenase complex. Biochim Biophys Acta. 1807:945–953.

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

Puthiyaveetil S, Ibrahim IM, Allen JF. 2013. Evolutionary rewiring: a modified prokaryotic gene-regulatory pathway in chloroplasts. Philos Trans R Soc Lond B Biol Sci. 368:20120260.

Puthiyaveetil S, et al. 2010. Transcriptional control of photosynthesis genes: the evolutionarily conserved regulatory mechanism in plastid genome function. Genome Biol Evol. 2:888–896.

Rubinstein CV, Gerrienne P, de la Puente GS, Astini RA, Steemans P. 2010. Early middle ordovician evidence for land plants in Argentina (eastern Gondwana). New Phytol. 188:365–369.

Saitou N, Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol. 4:406–425.

Schmitz-Linneweber C, Small I. 2008. Pentatricopeptide repeat proteins: a socket set for organelle gene expression. Trends Plant Sci. 13: 663–670.

Seliverstov AV, Lysenko EA, Lyubetsky VA. 2009. Rapid evolution of promoters for the plastome gene *ndhF* in flowering plants. Russ J Plant Physiol. 56:838–845.

Shiina T, Ishizaki Y, Yagi Y, Nakahira Y. 2009. Function and evolution of plastid sigma factors. Plant Biotechnol. 26:57–66.

Shimizu M, et al. 2010. Sigma factor phosphorylation in the photosynthetic control of photosystem stoichiometry. Proc Natl Acad Sci U S A. 107:10760–10764.

Steemans P, et al. 2009. Origin and radiation of the earliest vascular land plants. Science 324:353.

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

Sugiura M. 1992. The chloroplast genome. Plant Mol Biol. 19:149–168.

Sugiura C, Kobayashi Y, Aoki S, Sugita C, Sugita M. 2003. Complete chloroplast DNA sequence of the moss *Physcomitrella patens*: evidence for the loss and relocation of *rpoA* from the chloroplast to the nucleus. Nucleic Acids Res. 31:5324–5331.

Tamura K, Dudley J, Nei M, Kumar S. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol. 24:1596–1599.

Tanaka 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.

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

Tozawa Y, Tanaka K, Takahashi H, Wakasa K. 1998. Nuclear encoding of a plastid sigma factor in rice and its tissue- and light-dependent expression. Nucleic Acids Res. 26:415–419.

Tozawa Y, et al. 2007. The plastid sigma factor SIG1 maintains photosystem I activity via regulated expression of the *psaA* operon in rice chloroplasts. Plant J. 52:124–132.

Tsunoyama Y, 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.

Ueda M, et al. 2012. Composition and physiological function of the chloroplast NADH dehydrogenase-like complex in *Marchantia polymorpha*. Plant J. 72:683–693.

Vassylyev DG, et al. 2002. Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 A resolution. Nature 417:712–719.

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

Woodson JD, Perez-Ruiz JM, Schmitz RJ, Ecker JR, Chory J. 2012. Sigma factor-mediated plastid retrograde signals control nuclear gene expression. Plant J. 73:1–13.

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

Associate editor: Takashi Gojobori