Plant Biotechnology 24, 165–170 (2007)

Review

Plastidic proteins containing motifs of nuclear transcription factors

Yutaka Kodama*

Graduate School of Biological Sciences, Nara Institute of Science and Technology, Nara 630-0192, Japan
* E-mail: yu-kodam@bs.naist.jp Tel: +81-743-72-5653 Fax: +81-743-72-5659

Received November 27, 2006; accepted December 25, 2006 (Edited by Y. Hotta)

Abstract Plants are constantly exposed to environmental factors, including biotic and abiotic stresses, which may confer serious damage and affect survival. In order to cope with these stresses, plants have evolved a variety of defense systems. Despite intensive surveys, the molecular mechanisms are still not completely understood, and in particular, information on diversification of proteins is limited. In this article, we focus on examples of proteins changing both cellular localization and functions. Experimentally, six such proteins have so far been identified, all of them containing motifs of nuclear transcription factors, but localized in plastids: a 41-kDa protein from *Nicotiana tabacum* containing a zinc finger motif, shown to be a chloroplast nucleoid DNA binding protein (CND41); a plastid envelope DNA binding protein (PEND) from *Pisum sativum* possessing a basic domain plus leucine zipper motif; proteins designated as PD1 and PD3 also isolated from *P. sativum*, both having AT-hook motifs; a protein with a basic helix-loop-helix (bHLH) motif from *Arabidopsis thaliana* proposed to be a plastid transcription factor, and designated as PTF1; one wound-induced protein from *N. tabacum* (NtWIN4) shown to be a bHLH protein and exclusively localized into plastids. Judging from their structures, these proteins could have originated from eukaryotic ancestors. At the N-terminus, they possess clustered basic residues, which might constitute the key structure for conversion from nuclear transcription factors to plastid-resident proteins.

Key words: Chloroplast proteins, localization change, nuclear transcription factors, plastid localization, protein evolution.

Since emerging on earth about 500 million years ago, plants have continuously been exposed to biotic and abiotic environmental stresses, including pathogens, herbivores, competing other plants, wounding, low nutrients, water deficiency, extreme temperatures, and low and high light exposure. All these stresses can cause serious damage to individual plants, and occasionally cause extinction of species. To maintain life, therefore, plants have continually evolved to best fit diverse environmental conditions.

One process in plant evolution is considered to be diversification of proteins, and fine genomic analyses constitute a powerful approach to elucidate underlying mechanisms. Information is already available from three plant species, Arabidopsis, rice and poplar, whose genomic sequences have been completely decoded (The Arabidopsis Genome Initiative 2000; Goff et al. 2002; Tuskan et al. 2006). The numbers of genes encoding proteins greatly differ among the three plants; 25,554 for Arabidopsis, 37,544 for rice, and 45,555 for poplar. Such diversity in gene numbers may suggest that plants have independently evolved functional proteins to fit environmental conditions by increasing copy numbers of genes and/or chance protein modification. Data bases also indicate that relative frequencies of protein motifs are similar in each genome (Tuskan et al. 2006), suggesting that positions and/or combinations of motifs are important for protein evolution. However, it remains to be determined how functional proteins have actually evolved, and data on which proteins have actually converted to other functions during plant evolution are also limited.

This article attempts to shed light on this topic, by considering a group of proteins demonstrating change of cellular localization from the nucleus to plastids with concomitant functional alteration.

Plastidic proteins

The plastids are unique organelles of plants, comprising proplastids, chloroplasts, amyloplasts, leucoplasts, etioplasts and chromoplasts. They originated from prokaryotic cyanobacteria, and function in various regulatory systems including biosynthesis of

Abbreviations: bHLH, basic helix-loop-helix; bZIP, basic domain plus leucine zipper; CND41, 41-kDa chloroplast nucleoid DNA binding protein; GFP, green fluorescent protein; LRP, light-responsive promoter; PEND, plastid envelope DNA binding protein; PTF1, plastid transcription factor 1; WIN4, wound-induced clone 4.

This article can be found at http://www.jspcmb.jp/

chlorophylls, carotenoids and fatty acids (Buchanan et al. 2000). Proteomic analysis has revealed that chloroplasts contains 3,500–4,000 polypeptides, the majority of them considered to be transported into chloroplasts from the cytoplasm, since the maximal number of chloroplast genes is estimated to be about 200 (Race et al. 1999; Leister 2003). In general, transported proteins are encoded by genes residing in the nucleus, and their N-terminal regions are critical for transportation into plastids (Buchanan et al. 2000). Termed transit peptides, they are characterized by a lack of positively-charged amino acids in the N-terminal proximal region and acidic residues in the central region (Soll and Tien 1998). The central region is also rich in hydroxylated amino acids (Soll and Tien 1998) but detailed features of transit peptides remain to be determined.

Eukaryotic ancestors

Among transported proteins, several DNA binding proteins have been identified, and their origins shown to be not only from prokaryotes but also in a few cases from eukaryotes (Sato 2001). Proteins of eukaryotic ancestors possess motifs of nuclear transcription factors, suggesting them to be converted from the latter during evolution. To date, six such proteins have been identified from plastids (Nakano et al. 1997; Sato et al. 1993, 1995, 1998; review in Sato 2001; Baba et al. 2001; Kodama and Sano 2006). One chloroplast nucleoid-DNA binding protein was isolated from *Nicotiana tabacum*. This protein of 41-kDa was designated as CND41, and shown to contain a zinc finger motif (Nakano et al. 1993). A plastid envelope DNA binding protein was isolated from *Pisum sativum*, named PEND (Sato et al. 1993) and found to possess a basic domain plus leucine zipper (bZIP) motif. Two additional proteins, PD1 and PD3, were also isolated from *P. sativum*, and shown to have two and five AT-hook motifs, respectively (Sato et al. 1995). A plastid transcription factor 1 (PTF1) equipped with a basic helix-loop-helix (bHLH) motif was similarly identified from *Arabidopsis thaliana* (Baba et al. 2001) and another bHLH protein, designated as wound-induced clone 4 (NtWIN4), was found in *N. tabacum* (Kodama and Sano 2006). These six plastid-resident proteins are thought to have been functionally converted from transcription factors during evolution (review in Sato 2001; Kodama and Sano 2006, 2007). The following section describes current knowledge of the individual proteins and a putative key structure for localization change from nucleus to plastids.

CND41

During purification of chloroplasts-nucleoids from *N. tabacum*, a 41-kDa nucleoid DNA binding protein, CND41, was isolated by two-dimensional polyacrylamide gel electrophoresis and southwestern analysis (Nakano et al. 1993) and found to be present in both chloroplasts and etioplasts. The calculated isoelectric point was basic (Nakano et al. 1993) and two distinct domains were identified, a helix-turn-helix motif in the lysine-rich N-terminal region of mature protein, and an aspartyl protease active site (Nakano et al. 1997). *In vitro* assays showed CND41 to bind to DNA non-specifically through its lysine-rich region (Table 1) (Nakano et al. 1997). CND41 has strong proteolytic activity at acidic pH (pH2–4) *in vitro* (Murakami et al. 2000), and transgenic tobacco plants expressing antisense CND41 clearly showed a role in controlling transcripts for *psbA*, *psbD/C*, *rbcL*, *16SrDNA* and *23SrDNA*, all encoded by chloroplast DNA (Kato et al. 2004; Kato et al. 2005). CND41 is also involved in degradation of ribose-1,5-bisphosphate carboxylase/oxygenase through post-translational regulation in senescent leaves (Kato et al. 2004; Kato et al. 2005). In addition, CND41 appears to regulate plastid development and gibberellin biosynthesis (Nakano et al. 2003). These results strongly suggested that CND41 plays a critical role in transcription of the chloroplast genome, and in regulation of senescence, plastid development and gibberellin biosynthesis. Structural analysis indicated that CND41 possesses a zinc finger motif (Cys-X4-Cys-X2-Leu-X9-Cys-X4-Cys-X-Tyr) at amino acid positions 210–235 (Figure. 1, Table 1). Zinc finger motifs have been identified in many eukaryotic proteins, and shown to interact with nuclear DNA (O'Halloran 1993). The motif found in CND41 closely resembles that of transcription factor IIIA from yeast (His-X4-His-X2-Leu/Met-X9-Cys-X4-Cys-X-Tyr/Phe)

Table 1. Properties of plastid-resident eukaryotic ancestor proteins

Name	Motif	Bound sequence	Presequence	Recognition sequence
CND41	Zinc finger	Non-specific	120 aa[c]	nd
PEND	bZIP	TAAGAAGT	15 aa	~265 aa
PD1	AT-hook	psbM in cpDNA[a]	nd	nd
PD3	AT-hook	psbM in cpDNA[a]	nd	nd
PTF1	bHLH	ACC repeat	nd	nd
NtWIN4	bHLH	nd[b]	nd	67 aa

The presequence indicates the number of amino acids that are cleaved out and the recognition sequence gives those that are essential for plastid localization.

[a] cpDNA; chloroplast DNA, [b] nd.; not determined, [c] aa; amino acids.

(Archambault et al. 1992; Nakano et al. 1997), albeit with histidine residues in transcription factor IIIA substituted with cysteine residues in CND41. These observations suggest that a zinc finger protein initially located in the nucleus was converted to a chloroplast-localized DNA binding protease.

PEND

Basic domain plus leucine zipper (bZIP) proteins are eukaryotic transcription factors, having both DNA binding and dimerization motifs of approximately 60 amino acids (Bohmann et al. 1987; Landschulz et al. 1988; Struhl 1987; Vogt et al. 1987). The plastid envelope DNA binding protein (PEND) from *P. sativum* contains a distinct bZIP motif, an apparently unique sextuple repeat region and a membrane-spanning region (Figure. 1, Table 1) (Sato et al. 1993, 1998). PEND binds to a specific DNA sequence, TAAGAAGT, and is localized to the inner envelope membranes of chloroplasts (Table 1) (Sato et al. 1993, 1998; Sato and Ohta 2001), indicating a function in binding of chloroplast DNA to envelope membranes (Sato et al. 1993, 1998). PEND homologues were recently identified in angiosperms, and found to localize in chloroplasts, suggesting a common function among PEND homologues (Terasawa and Sato 2005). Biological activity of a PEND homologue, BnPEND, from *Brassica napus* was examined using transgenic over-expressing tobacco plants under control of the cauliflower mosaic virus (CaMV) 35*S* promoter (Wycliffe et al. 2005). Leaves of resulting transgenic tobacco were shown to develop a chlorophyll-deficit, and few palisade cells and chloroplasts. These observations point to a PEND evolution from a nuclear bZIP transcription factor to an envelope-resident DNA binding protein, and that it interferes with plastid development leading to a distorted leaf composition.

PD1 and PD3

During research into the PEND protein (Sato et al. 1993, 1998), two cDNAs for other plastidic DNA binding proteins were identified from *P. sativum* by southwestern analysis using the psbM region of chloroplast DNA as the probe (Table 1) (Sato et al. 1995). These were designated as *PD1* and *PD3* genes and PD1 and PD3 proteins possess AT-hook motifs, originally found in high mobility group I (Y) proteins, and could be shown to interact with the minor groove AT-rich regions of nuclear DNA (Grasser 1995) (Table 1). AT-hook is a small motif, which has a typical sequence pattern of centered glycine-arginine-proline (GRP) tri-peptides (Reeves and Nissen 1990). PD1 has two AT-hook motifs at the N-terminus (Sato et al. 1995) (Figure. 1). PD3 was previously reported to possess five AT-hook motifs (Sato et al. 1995), but scanning the latest PD3 amino acid sequence identified only four GRP tri-peptides (Figure. 1). PD3 was also shown to contain a jmjC domain at the C-terminus, this possibly being involved in chromatin organization by modulating heterochromatisation (described in the NCBI database; accession number CAA67296). Despite limited functional analyses, the structural information suggests that these proteins were converted to plastid DNA binding proteins from nuclear transcription factors containing AT-hook motifs.

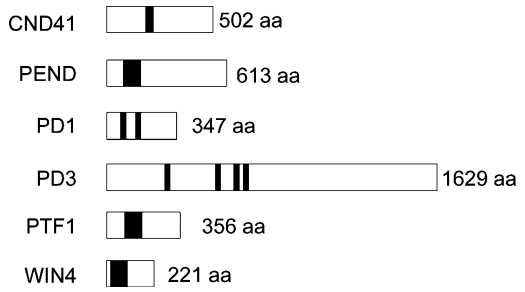

Figure 1. Schematic illustration of eukaryotic ancestor proteins in plastids. Closed boxes indicate motifs of nuclear transcription factors. Motifs for zinc finger and bZIP are shown in CND41 and in PEND, respectively. In PD1 and PD3, AT-hook motifs are indicated. bHLH motifs are present within PTF1 and NtWIN4 proteins. Lengths of amino acids are shown on the right side.

PTF1

Basic helix-loop-helix (bHLH) proteins possess a motif, consisting of a basic region to bind to the E-box sequence (CANNTG) and a helix-loop-helix region to function to form homo- and/or hetero-dimers (Murre et al. 1989; Ferre-D'Amare et al. 1994). A number of bHLH proteins have been identified in both plants and mammals, and these are thought to constitute a family of transcription factors located within nuclei (Garrell and Modolell 1990; Quail 2000). An exceptional case is a plastid-localized bHLH protein from *A. thaliana*. This protein was isolated by yeast one-hybrid screening methods using a region derived from *psbD* light-responsive promoter (LRP) from chloroplast DNA (Baba et al. 2001) (Figure. 1), and designated as plastid transcription factor 1 (PTF1). PTF1 was found to bind to the ACC repeat region of the *psbD* LRP sequence (Baba et al. 2001) (Table 1), and green fluorescent protein (GFP)-tagged PTF1 was shown to localize into chloroplasts of tobacco guard cells. A PTF1-deficient mutant, *ptf1*, showed several abnormalities, such as reduced activity of *psbD* LRP under continuous light conditions, and early bleaching, late flowering and dwarfism under short-day conditions. Based on these findings, PTF1 was proposed to regulate *psbD* LRP through its transcription in plastids. However, further studies gave rise to controversial arguments as to its function and cellular localization. For example, *psbD* LRP on chloroplast DNA was shown to be transcribed by

a nuclear-encoded transcription factor, AtSig5 (Tsunoyama et al. 2004). This was confirmed with a mutant, *sig5*, in which *psbD* LRP activity was completely eliminated, whereas in *ptf1* mutants the activity was not affected (Baba et al. 2001). The finding that light-dependent *psbD* transcription was not diminished in *ptf1* mutants suggests that PTF1 is unlikely to be involved in light signaling (Tsunoyama et al. 2004). We also established a GFP-tagged PTF1 to be clearly localized only in nuclei, and not in plastids (Kodama and Sano 2007), and that it exhibited repression activity on nuclear transcription by dual-luciferase assay (Kodama and Sano 2007). These observations suggest that PTF1 be carefully reevaluated as to biological functions and cellular localization in future.

NtWIN4

NtWIN4 is a novel plastid-resident bHLH type protein from *N. tabacum* (Kodama and Sano 2006). The gene encoding this protein was initially isolated as a wound-responsive gene by the fluorescent differential display method, and designated as *N. tobacum* wound-induced clone 4 (NtWIN4). *NtWIN4* is up-regulated not only by wounding, but also by pathogen attack and NtWIN4 possesses a bHLH motif at the N-terminus, and exhibits nuclear transcription repression activity (Figure. 1, Table 1). It is localized in chloroplasts of tobacco leaf cells and fine dissection assays using GFP showed a polypeptide of 221 amino acids, initiated from the second in-frame AUG, to be localized into plastids, while a polypeptide of 247 amino acids, translated from the first AUG, was localized to cytoplasm and nucleus (Kodama and Sano 2006). Further investigations on mRNA population and protein synthesis *in vitro* using the wheat germ translation system indicated that a plastid-type polypeptide is synthesized from a short mRNA lacking the first AUG (unpublished observation). These results suggest that NtWIN4 is mainly translated *in planta* from a short mRNA lacking the first AUG codon. Transgenic tobacco plants constitutively over-expressing NtWIN4 under control of the CaMV 35*S* promoter are abnormal in phenotype, showing albinism and slow growth with ultimate death. On transient expression after agro-infiltration, leaves showed severe chlorosis and transgenic tobacco plants with reduced NtWIN4 due to RNA interference exhibited delayed cell death when inoculated with pathogens. Based on these observations, we have concluded that NtWIN4 is converted from a nuclear bHLH-type transcription repressor to a plastid-resident regulatory factor, and that it is involved in hypersensitive responses through chloroplast disruption.

Basic region for plastid localization

Plastid proteins are initially synthesized as precursors containing the N-terminal transit peptide, generally defined as a removable polypeptide which is essential for plastid localization of proteins imported into plastids (Buchanan et al. 2000). In this article, transit peptide is defined as possessing two features; the recognition sequence as the plastid localization signal, and the presequence as the removable sequence.

The presequence of CND41 was predicted by direct amino acid sequencing of the N-terminus of mature protein, revealing that 120 amino acids at the N-terminus are cleaved out when CND41 is incorporated in chloroplasts (Nakano et al. 1997). It is not clear whether the presequence alone is sufficient for plastid localization. In the case of NtWIN4, a clear recognition sequence could be identified although a presequence was not determined (Kodama and Sano 2006). The C-terminal region with 15 amino acids of the recognition sequence was found to partly overlap with a basic region of the bHLH motif, which serves as a DNA binding domain. This indicates that the overlapping sequence is critical for DNA binding and plastid localization. However, since the native molecular mass of NtWIN4 is 17 kDa, it is highly probable that this region is cleaved out, resulting in NtWIN4 that lacks DNA binding activity *in planta*. PEND has an N-terminal presequence of 15 amino acids, which is cleaved out when PEND is translocated into the chloroplast envelope (Sato et al. 1998; Sato and Ohta 2001). However, a GFP fusion protein containing only the presequence was not translocated into plastids, in contrast to a fusion protein containing the basic region of bZIP (Ohki and Sato 2000), suggesting this latter to be prerequisite for the recognition sequence of PEND. PEND homologues were also shown to contain a basic region at the N-terminus and were localized to plastids, suggesting a common localization system for this family (Terasawa and Sato 2005). A basic region was also found in the N-terminus of CND41 at amino acid positions 121–138 in its precursor protein (Nakano et al. 1997). Within N-termini of PD1 and PD3, basic regions could also be identified at around amino acid positions 100 and 35, respectively (Figure. 2). Thus, a basic region at the N-terminus might be a key structure, which enables a protein to convert from a nuclear transcription factor to a plastid-resident protein.

Concluding remarks

This article documents the conversion of proteins containing motifs of eukaryotic nuclear transcription factors into plastid-resident functional proteins. Six such proteins were here exemplified, and one-by-one case

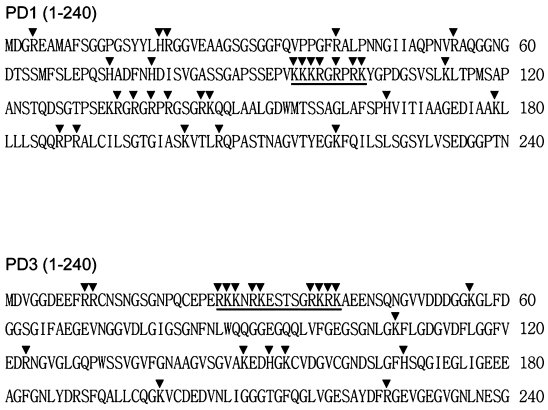

Figure 2. Basic regions within N-terminuses of PD1 and PD3. The N-terminal 240 amino acids of PD1 (upper panel) and PD3 (lower panel) are shown. Arrowheads indicate basic residues (H, K and R) and underlining indicates the region of clustered basic residues.

study provided not only novel ideas on multiple unique functions of proteins, but also clues as to their evolution. Basic regions at the N-terminus might allow proteins to change their localization from nuclei to plastids and DNA binding proteins such as transcription factors might originally possess such regions at high probability. However, whether or not the basic regions are recognized as transit peptides may depend on properties of individual proteins. Whatever the mechanism, the present survey points to two novel features concerning protein localization. First, proteins containing a transcription factor motif are not always nuclear transcription factors. Particularly, careful examination is needed if a basic region exists at the N-terminus of a given protein. Second, studies on such proteins may provide evidence for the hypothesis that plants have developed the ability to convert protein motifs, thereby acquiring a variety of functional proteins to best respond to severe environmental stresses.

Acknowledgements

The author thanks Prof. H. Sano (Nara Institute of Science and Technology) for valuable advice throughout the work and preparation of the manuscript, and Dr M. Moore (Intermal, Nagoya) for critical reading of the manuscript. This work was partly supported by a grant in-aid to Y. K. from the Scientific Research for Plant Graduate Students from Nara Institute of Science and Technology, sponsored by the Ministry of Education, Culture, Sports, Science and Technology, Japan.

References

Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408: 796–815

Archambault J, Milne CA, Schappert KT, Baum B, Friesen JD, Segall J (1992) The deduced sequence of the transcription factor TFIIIA from Saccharomyces cerevisiae reveals extensive divergence from Xenopus TFIIIA. *J Biol Chem* 267: 3282–3288

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

Bohmann D, Bos TJ, Admon A, Nishimura T, Vogt PK, Tjian R (1987) Human proto-oncogene c-jun encodes a DNA binding protein with structural and functional properties of transcription factor AP-1. *Science* 238: 1386–1392

Buchanan BB, Gruissem W, Jones RL (2000) *Biochemistry & Molecular Biology of Plants*. American Society of Plant Physiologists, Rockville, Maryland

Ferre-D'Amare AR, Pognonec P, Roeder RG., Burley SK (1994) Structure and function of the b/HLH/Z domain of USF. *EMBO J* 13: 180–189

Garrell J, Modolell J (1990) The Drosophila extramacrochaetae locus, an antagonist of proneural genes that, like these genes, encodes a helix-loop-helix protein. *Cell* 61: 39–48

Goff SA, Ricke D, Lan TH, Presting G, Wang R, Dunn M, Glazebrook J, Sessions A, Oeller P, Varma H, Hadley D, Hutchison D, Martin C, Katagiri F, Lange BM et al. (2002) A draft sequence of the rice genome (Oryza sativa L. ssp. japonica). *Science* 296: 92–100

Grasser KD (1995) Plant chromosomal high mobility group (HMG) proteins. *Plant J* 7: 185–192

Kato Y, Murakami S, Yamamoto Y, Chatani H, Kondo Y, Nakano T, Yokota A, Sato F (2004) The DNA-binding protease, CND41, and the degradation of ribulose-1,5-bisphosphate carboxylase/oxygenase in senescent leaves of tobacco. *Planta* 220: 97–104

Kato Y, Yamamoto Y, Murakami S, Sato F (2005) Post-translational regulation of CND41 protease activity in senescent tobacco leaves. *Planta* 222: 643–651

Kodama Y, Sano H (2006) Evolution of a basic helix-loop-helix protein from the transcriptional repressor to a plastid-resident regulatory factor—Involvement in hypersensitive cell death in tobacco plants. *J Biol Chem* 281: 35369–35380

Kodama Y, Sano H (2007) A comparative analysis of basic helix-loop-helix proteins, AtPTF1 and NtWIN4, with reference to plastid localization. *Plant Biotechnol* (in press)

Murakami S, Kondo Y, Nakano T, Sato F (2000) Protease activity of CND41, a chloroplast nucleoid DNA-binding protein, isolated from cultured tobacco cells. *FEBS Lett* 468: 15–18

Murre C, McCaw PS, Baltimore D (1989) A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD and myc proteins. *Cell* 56: 777–783

Landschulz WH, Johnson PF, McKnight SL (1988) The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. *Science* 240: 1759–1764

Leister D (2003) Chloroplast research in the genomic age. *Trends Genet* 19: 47–56

Nakano T, Sato F, Yamada Y (1993) Analysis of nucleoid-proteins in tobacco chloroplasts. *Plant Cell Physiol* 34: 873–880

Nakano T, Murakami S, Shoji T, Yoshida S, Yamada Y, Sato F (1997) A novel protein with DNA binding activity from tobacco chloroplast nucleoids. *Plant Cell* 9: 1673–1682

Ohki Y, Sato N (2000) Analysis of chloroplast targeting sygnals of the PEND protein. *Plant Cell Physiol* (supplement) 40: s222

O'Halloran TV (1993) Transition metals in control of gene expression. *Science* 261: 715–725

Quail PH (2000) Phytochrome interacting factors. Semin. *Cell Dev Biol* 11: 457–466

Race HL, Herrmann RG, Martin W (1999) Why have organelles retained genomes? *Trends Genet* 15: 364–370

Reeves R, Nissen MS (1990) The A·T-DNA-binding domain of mammalian high mobility group I chromosomal proteins. A novel peptide motif for recognizing DNA structure. *J Biol Chem* 265: 8573–8582

Sato N, Albrieux C, Joyard J, Douce R, Kuroiwa T (1993) Detection and characterization of a plastid envelope DNA-binding protein which may anchor plastid nucleoids. *EMBO J* 12: 555–561

Sato N (2001) Was the evolution of plastid genetic machinery discontinuous? *Trends Plant Sci* 6: 151–155

Sato N, Mitsumi O, Joyard J, Douce R (1995) DNA-binding proteins mediate interaction of nucleoids with envelope membrane in developing plastids. In *Photosynthesis: from Light to Biosphere* (Mathis, P. ed.) Vol. III, Kluwer Academic Publishers, Dordrecht, pp 635–638

Sato N, Ohshima K, Watanabe A, Ohta N, Nishiyama Y, Joyard J, Douce R (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

Sato N, 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

Soll J, Tien R (1998) Protein translocation into and across the chloroplastic envelope membranes. *Plant Mol Biol* 38: 191–207

Struhl K (1987) The DNA-binding domains of the jun oncoprotein and the yeast GCN4 transcriptional activator protein are functionally homologous. *Cell* 50: 841–846

Terasawa K, Sato N (2005) Occurrence and characterization of PEND proteins in angiosperms. *J Plant Res* 118: 111–119

Tsunoyama Y, Ishizaki Y, Morikawa K, Kobori M, Nakahira Y, Takeba G, Toyoshima Y, Shiina T (2004) Blue light-induced transcription of plastid-encoded *psbD* gene is mediated by a nuclear-encoded transcription initiation factor, AtSig5. *Proc Natl Acad Sci USA* 101: 3304–3309

Tuskan GA, Difazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Putnam N, Ralph S, Rombauts S, Salamov A et al. (2006) The genome of black cottonwood, Populus trichocarpa (Torr. & Gray). *Science* 313: 1596–1604

Vogt PK, Bos TJ, Doolittle RF (1987) Homology between the DNA-binding domain of the GCN4 regulatory protein of yeast and the carboxyl-terminal region of a protein coded for by the oncogene jun. *Proc Natl Acad Sci USA* 84: 3316–3319

Wycliffe P, Sitbon F, Wernersson J, Ezcurra I, Ellerstrom M, 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