TECHNICAL ADVANCE

Identification of a plastid intercistronic expression element (IEE) facilitating the expression of stable translatable monocistronic mRNAs from operons

Fei Zhou, Daniel Karcher and Ralph Bock*
Max-Planck-Institut für Molekulare Pflanzenphysiologie (MPI-MP), Am Mühlenberg 1, D-14476 Potsdam-Golm, Germany

Received 29 June 2007; accepted 12 July 2007.
*For correspondence (fax +49 331 567 8701; e-mail rbock@mpimp-golm.mpg.de).

Summary

Most plastid genes are part of operons and expressed as polycistronic mRNAs. Many primary polycistronic transcripts undergo post-transcriptional processing in monocistronic or oligocistronic units. At least some polycistronic transcripts are not translatable, and endonucleolytic processing may therefore be a prerequisite for translation to occur. As the requirements for intercistronic mRNA processing into stable monocistronic transcript are not well understood, we have sought to define minimum sequence elements that trigger processing and thus are capable of generating stable translatable monocistronic mRNAs. We describe here the in vivo identification of a small intercistronic expression element that mediates intercistronic cleavage into stable monocistronic transcripts. Separation of foreign genes by this element facilitates transgene stacking in operons, and thus will help to expand the range of applications of transplastomic technology.

Keywords: chloroplast, RNA processing, intercistronic processing, RNA cutting, polycistronic transcript, plastid transformation.

Introduction

In plastids (chloroplasts), primary transcripts undergo a complex series of mRNA maturation steps. These include processing of the 5' and 3' ends (RNA trimming), intron splicing, RNA editing, and cleavage of polycistronic precursor transcripts into monocistronic or oligocistronic mRNAs (RNA cutting). Most of these RNA processing steps exhibit prokaryotic features that have been retained from the cyanobacterial ancestor of all present-day plastids. 5' and 3' end processing in plastids is catalyzed by nucleus-encoded prokaryotic-type ribonucleases. Whereas 5' end maturation is catalyzed primarily by endoribonucleases, 3' end formation is mediated by the concerted action of endoribonucleases and 3' → 5' exoribonucleases (Sakamotu et al., 1994; Hayes et al., 1996; reviewed in Monde et al., 2000; Herrin and Nickelsen, 2004). Stem–loop-type RNA secondary structures within the 5' and 3' untranslated regions (UTRs) of plastid messenger RNAs provide important recognition elements for RNA processing enzymes, and, in addition, can serve as protective elements preventing rapid RNA degradation (Barkan and Goldschmidt-Clermont, 2000; Mayfield et al., 1995; Monde et al., 2000; Stern and Gruissem, 1987).

As in eubacteria, most genes in plastids are organized in polycistronic transcription units (operons). Transcription of bacterial operons usually gives rise to stable polycistronic mRNAs that are directly translated, although in some cases processing to monocistronic mRNAs is known to occur and is involved in gene regulation (e.g. Carpousis et al., 1989). In plastids, most polycistronic precursor transcripts are post-transcriptionally processed into monocistronic or...
oligocistronic units, presumably by specific endonucleolytic cleavage (Herrin and Nickelsen, 2004; Sugita and Sugiura, 1996; Westhoff and Herrmann, 1988). One of the exceptions is the psbE operon, which comprises four small genes for polypeptides of photosystem II (psbE, psbF, psbL and psbJ; Carrillo et al., 1986; Willey and Gray, 1989). The psbE operon is transcribed as a single 1.1 kb mRNA species that remains tetracistronic and is not processed further. Other examples of unprocessed polycistronic transcripts include the psaA/B transcript (Meng et al., 1988) and petA, which represents the last cistron of a large polycistronic transcript and is not cleaved off from the upstream open reading frame ycf10 (Willey and Gray, 1990). The transcripts from most other plastid operons undergo intercistronic processing (also referred to as RNA cutting; Sugiura, 1992), and, at least in some cases, cutting into monocistronic units is an essential processing step: while some polycistronic precursor transcripts can be translated (Barkan, 1988), others must be processed to become translatable or make translation more efficient. This is supported by the analysis of nuclear mutants defective in distinct intercistronic processing events, as well as by in vitro translation studies. For example, the maize crr1 mutant is defective in intercistronic processing between the petB and petD cistrons, which results in a concomitant loss of petD translation (Barkan et al., 1994; Fisk et al., 1999), suggesting that petD needs to be monocistronic to be translated. Similarly, defective processing of psbH mRNA from the pentacistronic primary transcript of the psbB operon leads to loss of psbH translation in the Arabidopsis hcf107 mutant (Felder et al., 2001). Another Arabidopsis mutant with impaired intercistronic RNA processing is crr2, in which endonucleolytic cleavage between the rps7 and ndhB cistrons does not occur (Hashimoto et al., 2003). This results in loss of the NDH complex, most probably because the unprocessed ndhB message cannot be translated (Hashimoto et al., 2003). mRNA secondary structure formation has been implicated in impaired translatability of unprocessed polycistronic precursors (Felder et al., 2001; Hirose and Sugiura, 1997). Direct evidence for this has come from in vitro translation experiments with ndhD transcripts, another plastid mRNA whose translation is dependent on prior intercistronic processing. Translation of the di-cistronic psaC-ndhD precursor transcript was shown to be impaired by RNA secondary structure formation between a short (8 nt) sequence within the psaC coding region and a complementary sequence in the 5’ UTR of the downstream ndhD (Hirose and Sugiura, 1997).

As such long-range secondary structural interactions are not easily predictable, it is generally not possible to make educated guesses about the translatability of polycistronic transcripts in plastids. This is highly unfortunate, because simultaneous expression of multiple transgenes from operons is viewed as one of the unique attractions of chloroplast transformation technology (Bogorad, 2000; Daniell and Dhingra, 2002; Heifetz, 2000; Maliga, 2004). Expression of transgenes from polycistronic mRNAs has been successful in some cases (Quesada-Vargas et al., 2005; Staub and Maliga, 1995), but poor translation of polycistronic mRNAs is likely to be responsible for at least some cases where transgene expression was disappointingly low (Nakashita et al., 2001) or unsuccessful altogether (Magee et al., 2004). Clearly, processing of polycistronic transcripts into stable monocistronic mRNAs would greatly reduce the risk of failure of transgene expression from the plastid genome and thus make transplastomic experiments more predictable.

Here, we report a solution for this problem. We have identified a small sequence element, referred to as an intercistronic expression element (IEE), that mediates the efficient intercistronic cleavage of polycistronic mRNAs into stable monocistronic transcripts. We show that, while this element is not required for processing downstream of the first cistron to occur, it is essential to confer mRNA stability and translation of the second cistron. The identified IEE is small enough to serve as a universal tool for stacking of foreign genes in operons, and thus will help to extend the range of applications of transplastomic technology.

Results

Mapping of intercistronic mRNA processing sites in the tobacco psbB operon transcript

To identify sequence elements suitable for triggering processing of polycistronic transcripts into stable and translatable monocistronic mRNAs, we analyzed processing in the tobacco psbB operon (Figure 1a), which is one of the best characterized multi-gene operons in plastids (Felder et al.,...
2001; Meierhoff et al., 2003; Westhoff and Herrmann, 1988). The psbB operon consists of five genes, three of which encode photosystem II components (psbB, psbT and psbH), with the remaining two encoding subunits of the cytochrome b6f complex (petB and petD) (Figure 1a). The five genes are co-transcribed, giving rise to a long polycistronic precursor RNA, which is then cleaved into smaller units by a complex series of processing events (Westhoff and Herrmann, 1988). Not all final processing products are monocistronic: the small psbT cistron remains associated with the upstream psbB, forming a di-cistronic mature mRNA, and the two cytochrome b6f components, petB and petD, are only inefficiently processed into monocistronic mRNAs, leaving a large proportion of the transcripts di-cistronic (Felder et al., 2001; Westhoff and Herrmann, 1988). Stem–loop-type secondary structures are found upstream of most cleavage sites, suggesting that they stabilize the 3′ ends of the mature transcripts (Stern and Gruissem, 1987; Westhoff and Herrmann, 1988).

We decided to map the intercistronic cleavage sites upstream and downstream of the psbH cistron in tobacco, because psbH is efficiently cleaved into a monocistronic mRNA by two endonucleolytic cleavage events upstream and downstream (Felder et al., 2001; Westhoff and Herrmann, 1988). To precisely identify the cleavage sites, we employed an RNA circularization-based method by which the head-to-tail ligated 5′ and 3′ UTRs of the mRNA can be simultaneously analyzed (Zandueta-Criado and Bock, 2004). This analysis revealed a major cleavage site upstream of psbH (Figure 1b), two nucleotides away from the suggested processing site in Arabidopsis that was determined by S1 nuclease mapping (Felder et al., 2001). In the psbH-petB intergenic spacer, we identified one major and two minor cleavage sites (Figure 1c). The major site was found in four of the six clones sequenced, the minor sites in one clone each.

We next wished to determine whether RNA secondary structures are potentially involved in cleavage, for example whether they could mediate cleavage site recognition by a specific endoribonuclease. We therefore analyzed the nucleotide sequences surrounding the identified cleavage sites for their potential to fold into stable secondary structures. This was the case for both the processing site upstream and the site downstream of psbH(Figure 1d). Interestingly, in both structures, the cleavage site is localized approximately in the middle of the central loop of a stem–loop structure, possibly suggesting that cleavage site selection is aided by the formation of RNA secondary structures.

Integration of transgene operons with putative processing elements into the tobacco plastid genome

To identify a minimum sequence element sufficient for triggering processing of polycistronic transcripts into stable and translatable monocistronic mRNAs, we decided to test sequences derived from the two major processing sites mapped upstream and downstream of psbH in vivo by chloroplast transformation. To this end, we constructed a plastid transformation vector with two transgenes linked together in an operon: the kanamycin resistance gene nptII and the gene for the yellow fluorescent protein, yfp (Figure 2a,b). The two coding regions are separated by a sequence encoding a stem–loop structure (TrbCl) (Figure 2b) to ensure transcript stability of the mRNA from the first cistron after processing, two restriction sites suitable for integrating potential intercistronic processing elements, and a Shine–Dalgarno sequence to mediate translation initiation at the second cistron. For both major processing sites (Figure 1b–d), we constructed two chloroplast transformation vectors (Table 1). Vectors pZF75 and pZF77 contain the complete secondary structures in which the cleavage sites are embedded. This corresponds to sequence elements from −25 to +25 with respect to the psbT–psbH processing site (Figure 1b,d; vector pZF75) and −14 to +14 with respect to the psbH–petB processing site (Figure 1c,d; vector pZF77). In addition, we used two shorter sequences that included only the stem–loop up to the first bulge in the stem structure (Figure 1d). These sequence elements correspond to nucleotide positions −15 to +15 with respect to the psbT–psbH processing site (Figure 1b,d; vector pZF74) and −8 to +8 with respect to the psbH–petB processing site (Figure 1c,d; vector pZF76). Finally, a fifth construct containing no putative processing element between the nptII and yfp cassettes was transformed as a control (pZF73; Table 1).

The constructs were introduced into the tobacco plastid genome by biolistic chloroplast transformation (Svab and Maliga, 1993). Two homologous recombination events in the regions flanking the three transgenes (the nptII–yfp operon and the selectable spectinomycin resistance gene aadA) (Figure 2a,b) incorporate the foreign genes into the plastid genome. Selection of bombarded leaf samples for resistance to spectinomycin conferred by the aadA marker gene yielded several transplastomic lines that were subjected to additional rounds of regeneration and selection to obtain homoplasmic tissue (Bock, 2001; Svab and Maliga, 1993). After three such rounds, plants were regenerated, rooted in sterile culture, transferred to soil and grown to maturity in the glasshouse. Plants from all transplastomic lines were phenotypically identical in that they were indistinguishable from wild-type plants (not shown). For each construct, two independently generated lines were selected for further analysis. The transplastomic lines will be subsequently referred to as Nt-pZF followed by the number of the construct and the number of the individual transplastomic line (e.g. Nt-pZF73-3, indicating transplastomic tobacco line 3 generated using construct pZF73).

To confirm correct integration of the transgenes into the plastid genome and to test for homoplasmity of the
transplastomic lines, RFLP analyses were performed (Figure 2c, and data not shown). These assays revealed, in addition to a strong band of the expected size for the transplastomic fragment, a faint hybridization signal that corresponded in size to the restriction fragment from the wild-type genome (Figure 2c). Persistence of a wild-type-like hybridization signal even after multiple rounds of selection and regeneration is often seen in transplastomic lines and usually is not caused by true heteroplasmy of the plastid transformants, but rather by the presence of promiscuous plastid DNA in one of the other two genomes of the plant cell. This is because, during evolution, large fragments of plastid DNA have integrated into the nuclear and mitochondrial genomes (for review, see Bock, 2006; Timmis et al.,

Table 1 Features of chloroplast transformation vectors and transplastomic lines generated in this study

| Vector | Transplastomic lines | Presence of a potential intercistronic expression element (IEE) | Origin of the potential IEE (intergenic spacer) | Size of the potential IEE |
|--------|----------------------|-------------------------------------------------|---------------------------------|-------------------------|
| pZF73  | Nt-pZF73            | –                  | –                                | 0                       |
| pZF74  | Nt-pZF74            | +                  | psbT-psbH                         | ±15                     |
| pZF75  | Nt-pZF75            | +                  | psbT-psbH                         | ±25                     |
| pZF76  | Nt-pZF76            | +                  | psbH-petB                         | ±8                      |
| pZF77  | Nt-pZF77            | +                  | psbH-petB                         | ±14                     |

© 2007 The Authors
Journal compilation © 2007 Blackwell Publishing Ltd, The Plant Journal, (2007), 52, 961–972
2004) as non-functional, so-called 'promiscuous DNA'. Our previous work has established that the wild-type-like bands in DNA gel-blot analyses of otherwise homoplasmic transplastomic lines originate from such promiscuous DNA (Figure 2c) (Hager et al., 1999; Ruf et al., 2000).

To ultimately confirm homoplasmy of the transplastomic lines, seeds were obtained from transplastomic plants that had either been selfed or reciprocally crossed to the wild-type. With these seeds, inheritance assays were performed, which represent the most sensitive available test to assess homoplasmy (Bock, 2001; Maliga, 2004). As expected, lack of segregation of spectinomycin resistance in the T1 generation demonstrated homoplasmy (Figure 3, and data not shown) and confirmed uniparentally maternal transgene inheritance, as is typical of a plastid-encoded trait.

Analysis of transcript accumulation and RNA processing in transplastomic lines

Having successfully generated homoplasmic transplastomic plants with all vectors, we next wished to compare the five constructs with respect to RNA processing and transcript accumulation for the two genes of the operon. We first analyzed transcript pattern and RNA accumulation for nptII, the first cistron of the operon. Surprisingly, when RNA gel blots were hybridized to an nptII-specific probe, identical transcript patterns were detected in all transplastomic lines (Figure 4a): a strongly hybridizing band corresponding in size to the monocistronic nptII message was seen in all lines. In addition, weakly hybridizing larger RNA species were detected, including a transcript of the size of the di-cistronic nptII-yfp RNA (Figure 4a). The same transcript pattern was also present in the Nt-pZF73 control lines that do not harbor a putative processing element, indicating that processing downstream of nptII does not require a specific sequence element. This may suggest that the transcript-stabilizing stem-loop structure downstream of the nptII coding region (taken from the rbcL 3' UTR) (Figure 2b) is sufficient to mediate faithful 3' end formation of the nptII mRNA.

Next, we investigated transcript pattern and mRNA accumulation of yfp, the second cistron of the operon. Interestingly, high amounts of monocistronic yfp mRNA accumulated only in the Nt-pZF75 lines harboring the complete stem-loop structure surrounding the psbT–psbH intercistronic processing site. All other lines had at best small amounts of yfp mRNA that corresponded in size to monocistronic message (Figure 4b). Remarkably, lack of accumulation of monocistronic yfp mRNA in these lines was not accompanied by increased accumulation of the di-cistronic precursor RNA. This, together with the presence of comparable amounts of monocistronic nptII mRNA in all lines, suggests that the presence of a putative processing element does not influence formation of the nptII 3' end. Instead, it appears to serve as a critical mRNA stability determinant for the second cistron, yfp (Figure 4a,b).

In order to exclude the possibility that intercistronic processing in the Nt-pZF75 lines did not occur within the psbT–psbH intercistronic processing site, but instead was caused by sequence context-dependent processing at the upstream TrbCL sequence (Figure 2b) (Chakrabarti et al., 2006; Staub and Maliga, 1995), we mapped the mRNA 5'
end of the monocistronic yfp mRNA. RNA circularization and analysis of the head-to-tail ligated 5’ and 3’ UTRs allowed us to determine the termini of both mRNA ends. Sequencing of ten individual clones revealed that, in all cases, 5’ processing occurred faithfully within the psbT–psbH intercistronic processing element, confirming that this sequence element indeed serves as an intercistronic processing signal triggering the cleavage of polycistronic into monocistronic mRNAs. All ten clones had an identical 5’ end to the psbH mRNA (Figures 1b and 4c). In contrast, the 3’ end was slightly variable, as has been observed previously for many other plastid mRNAs. All 3’ ends mapped 3–10 nucleotides downstream of a putative transcript-stabilizing stem-loop-type RNA secondary structure within the rps16 3’ UTR of the chimeric yfp (Figure 4c).

Accumulation of NptII and YFP proteins expressed from operon constructs

Having established that monocistronic nptII transcripts accumulate in all transplastomic lines, whereas stable monocistronic yfp mRNA accumulates only in the Nt-pZF75 lines, we next wished to investigate the correlation between RNA abundance and protein accumulation. The high-level kanamycin resistance of all transplastomic lines tentatively indicated that NptII protein accumulates to reasonably high levels (Figure 3, and data not shown). This was confirmed by Western blot analysis with a specific anti-NptII antibody: all lines accumulated similarly high levels of NptII protein (Figure 5a), as expected from the accumulation of similar amounts of monocistronic nptII message in all

Figure 4. RNA accumulation in transplastomic lines harboring various candidate processing elements between the nptII and yfp cistrons.
(a) Accumulation of nptII mRNA. All transplastomic lines accumulate predominantly monocistronic nptII mRNA, in addition to small amounts of di-cistronic nptII-yfp transcripts and other minor RNA species that were not further characterized. Accumulation of monocistronic nptII message in the Nt-pZF73 control lines demonstrates that 3’ processing of nptII mRNA is independent of the presence of the putative processing sequences.
(b) Accumulation of yfp mRNA. Significant amounts of (monocistronic) yfp mRNA accumulated only in the Nt-pZF75 lines that harbor the s25 IEE from the psbT–psbH intergenic spacer. Note that, in addition to the monocistronic yfp and di-cistronic nptII–yfp transcripts, two minor RNA species also accumulate. One of them is apparently larger than the monocistronic yfp message, and the other is approximately 200 bp larger than the di-cistronic nptII–yfp transcript. These minor RNA species were not further characterized, but the most probable explanation is that they originate from read-through transcription through trnfM (see Figure 1b), whose antisense transcript can also fold into a stable cloverleaf-like secondary structure and thus act as an RNA processing signal.
(c) Mapping of the 5’ and 3’ ends of the monocistronic yfp mRNA in Nt-pZF75 plants. The sequence of a cDNA clone derived from head-to-tail ligated yfp mRNA is shown. The ligation site (i.e. the border between the 3’ end and the 5’ end of the circularized mRNA) is indicated by the dotted vertical line. Note that the 5’ end generated by processing within the psbT–psbH spacer element is identical in all ten cDNA clones and corresponds precisely to the 5’ end of the psbH mRNA (Figure 1b,d). Alternative 3’ ends are indicated by arrowheads, and the number of clones in which the respective termini were found is indicated. A putative transcript-stabilizing stem-loop-type RNA secondary structure within the rps16 3’ UTR is marked by horizontal arrows (interruptions indicate unpaired nucleotides). The stop codon, Shine–Dalgarno sequence and start codon are boxed.
transplastomic lines (Figure 4a). In contrast, when the blots were probed with an anti-GFP antibody (which also recognizes YFP, because YFP is a mutant GFP variant), protein accumulation was only detected in the Nt-pZF75 lines, correlating with the accumulation of stable monocistronic mRNA only in these lines. We therefore conclude that the complete stem–loop structure \((\pm 25)\) surrounding the psbT-psbH intercistronic processing site (as present in the Nt-pZF75 lines) represents a suitable sequence element to confer stable expression of downstream cistrons in multigene operons, and thus can serve as a genuine IEE.

Expression of yfp in the Nt-pZF75 lines, but not in the other transplastomic lines, was finally confirmed by confocal laser-scanning microscopy (Figure 5b, and data not shown). When YFP accumulation was compared with GFP accumulation in transplastomic control plants harboring a gfp gene under the control of the ribosomal RNA operon promoter, the GFP signal in the control plants was significantly stronger than the YFP signal in the Nt-pZF75 lines (Figure 5c). This may be due to weaker recognition of YFP by the anti-GFP antibody and/or lower stability of YFP in plastids.

**Discussion**

Stacking of multiple transgenes is often desirable in both basic research and plant biotechnology. As most plastid genes are arranged in operons and expressed as polycistronic transcripts, easy transgene pyramiding by co-expression of genes from operons is considered to be an attractive advantage of chloroplast transformation over conventional nuclear transformation. Therefore, understanding the rules that govern the efficient expression of multiple linked (trans)genes is a prerequisite for the successful engineering and rational design of complex operons. Earlier work had established that, unlike in eubacteria, post-transcriptional processing of polycistronic mRNAs into monocistronic units is often required for gene expression in plastids (Barkan et al., 1994; Fisk et al., 1999; Hashimoto et al., 2003; Hirose and Sugiura, 1997). However, very little is known about the signals at the RNA level that trigger intercistronic processing into stable and translatable monocistronic mRNAs.

In this work, we have attempted to identify a minimum sequence element, referred to as an IEE, that is suitable to direct cleavage of a polycistronic precursor transcript into stable monocistronic units. We have mapped two intercistronic cleavage sites in the tobacco psbB operon and found that they are located in the central position of the loop domain of a putative stem–loop-type RNA secondary structure. *In vivo* testing of these elements in a synthetic operon of two transgenes revealed that only the complete stem–loop structure surrounding the psbT-psbH processing site conferred expression of the downstream cistron and thus provides a functional IEE. Surprisingly, inclusion of an IEE between the two cistrons of the operon was not required to trigger generation of monocistronic mRNA for the 5′ cistron...
Experimental procedures

Plant material and growth conditions

Tobacco plants (Nicotiana tabacum cv. Petit Havana) for chloroplast isolation were grown under standard greenhouse conditions. Sterile tobacco plants were grown on agar-solidified MS medium containing 30 g L⁻¹ sucrose (Murashige and Skoog, 1962). Regenerated shoots from transplastomic lines were rooted and propagated on the same medium. Rooted homoplasmic plants were transferred to soil and grown to maturity in the glasshouse under standard conditions.

RNA extraction and circularization, cDNA synthesis and PCR

Chloroplast isolation and RNA extraction from purified chloroplasts were carried out as described previously (Bock, 1998). Total chloroplast RNA (10 µg) was self-ligated at 37°C for 1 h with 20 units of T4 RNA ligase (New England Biolabs; http://www.neb.com) in a final reaction volume of 100 µL. A 1 µl aliquot of the circularized RNA and 38 ng of oligonucleotide PpsbH-RT (5’-TTCCCCACCCAGAGCTA-3’) were denatured at 70°C for 5 min and used for cDNA synthesis, which was performed as described previously (Zandueta-Criado and Bock, 2004). The resulting cDNA was used directly as a template for PCR reactions, and amplified according to standard protocols (30 sec at 93°C, 90 sec at 52°C, 90 sec at 72°C; 30 cycles). Primers PpsbH-p1 (5’-AAATCTCTTACCAGATTC-3’) and PpsbH-p2 (5’-GGGAGAGGATCTTCTTTC-3’) were used for specific amplification of the head-to-tail ligated 5’ and 3’ UTRs. The mRNA 5’ and 3’ ends of the monocistronic yfp mRNA were mapped by the same method using 10 µg total RNA from an Nt-pZF75 plant and 38 ng of oligonucleotide Pyfp-RP (5’-CGTGGTGGACAGATGAACTTC-3’) for priming of the cDNA synthesis. Primers Pyfp-p1 (5’-TGCGGCTTTACTGCAGCC-3’) and Pyfp-p2 (5’-ACATGGCTCGCTGAACGGTC-3’) were used for the specific amplification of head-to-tail ligated yfp transcripts.

Cloning and DNA sequencing

Amplified PCR products were cloned into the pCR2.1-TOPO vector using the TOPO TA cloning kit (Invitrogen, http://www.invitrogen.com/). Individual clones were sequenced using M13 reverse primer (MWG-BIOTECH AG; http://www.mwg-biotech.com).

Construction of plastid transformation vectors

The plastid transformation vectors constructed in this study are based on the previously described vector pRBS9 (Ruf et al., 2001). To facilitate analysis of intercistronic processing, two reporter genes were introduced into the polylinker of pRBS9: nptII and yfp. The nptII gene is driven by the tobacco plastid rRNA operon promoter (Prrn) fused with the 5’ UTR of gene 10 from Escherichia coli phage T7 and the downstream box (Kuroda and Maliga, 2001). The terminator is derived from rbcL (Kuroda and Maliga, 2001). The yfp gene was fused with the Shine–Dalgarno sequence from the rbcL gene, and is terminated by the rps16 terminator (Staub and Maliga, 1994; Wurbs et al., 2007). Between these two genes, various putative processing sequences were integrated (Figures 1 and 2). The coding region of yfp was PCR-amplified from plasmid pEYPFP-C1 (Clontech, http://www.clontech.com/) using primers Pyfp5’ (5’-TTTTGTGACGAGGAGGAGATTCTCATGAGCAAGGGCGAGGACC-3’) and Pyfp3’ (5’-TTTTCTCGAGTTACTTGATACGTCGTCCATT-3’).
With these primer sequences, the rbcL leader sequence containing the Shine–Dalgarno sequence (in italics in the Pyfp5′ sequence) and the restriction sites SalI and XhoI (underlined sequences) were introduced into the PCR product. The resulting amplification product was digested with SalI and XhoI, and ligated into the corresponding sites in the polylinker of a pBluescript vector, generating plasmid pBSyfp. The rps16 terminator was cloned by PCR amplification from tobacco DNA using the primer pair Rps16-5′ (5′-TTTTCTGAGTAGAAATTTTTTAAAG-3′) and Rps16-3′ (5′-TTTTCAATCCATAAGLTGAAGC-3′) and introducing XhoI and KpnI restriction site (underlined) into the PCR product. The PCR product was digested with the enzymes XhoI and KpnI, and cloned downstream of the yfp coding region into vector pBSyfp, generating plasmid pBSyfpT. Subsequently, the nptII expression cassette from phy2000 (Kuroda and Maliga, 2001) was cloned into the polylinker of pBSyfpT as a SalI/HindIII fragment. Finally, the two expression cassettes were excised from the pBluescript vector as a 2.2 kb SalI/KpnI fragment, and ligated into the similarly digested plastid transformation vector pR895, producing vector p2F73. The four different putative processing sequences were inserted into HindIII/SalI-digested p2F73 as annealed complementary synthetic oligonucleotides that contained single-stranded overhangs for ligation into HindIII and SalI sites. The following oligonucleotides were used (compatible overhangs are underlined): P5′-psbII±15-5′, 5′-AGCTTATTTAACAAGGAAATTTATACAAATGCT-CAAAG-3′; P5′-psbII±15-3′, 5′-TCGACATGATTTCTGATACATTTCCGGTTGAAATA-3′; 5′-psbII±25-5′, 5′-AGCTTTGACAGTTTATTTACACCCGAATTTTGGTTTTGGG-3′; P5′-psbII±25-3′, 5′-TCGACTTTAGACATGATTTCTGATACATTTCCGGTTGAAATA-3′; P5′-psbII±25-5′, 5′-AGCTTTGACAGTTTATTTACACCCGAATTTTGGTTTTGGG-3′; P5′-psbII±25-3′, 5′-AGCTTATTTAACAAGGAAATTTATACAAATGCT-CAAAG-3′; P5′-psbII±15-3′, 5′-TCGACATGATTTCTGATACATTTCCGGTTGAAATA-3′; P5′-psbII±15-5′, 5′-AGCTTTGACAGTTTATTTACACCCGAATTTTGGTTTTGGG-3′; P5′-psbII±15-3′, 5′-TCGACATGATTTCTGATACATTTCCGGTTGAAATA-3′. These cassettes were excised from the pBluescript vector as a 2.2 kb SacI/HindIII/I fragment, and ligated into the similarly digested plastid containing medium to enrich the transplastome and select for homoplasmic tissue. Construct, several independent transplastomic lines were subjected to PCR. The resistant lines were selected on regeneration medium containing spectinomycin and streptomycin (500 mg l–1). Primary spectinomycin-plasmid-coated 0.6 mm gold particles using a PDS1000He biolistic gun (Bio-Rad, http://www.bio-rad.com/). Young leaves from sterile tobacco plants were bombarded with plasmid-coated 0.6 μm gold particles using a PDS1000He biolistic gun (Bio-Rad, http://www.bio-rad.com/). Primary spectinomycin-resistant lines were selected on regeneration medium containing 500 mg l–1 spectinomycin (Svab and Maliga, 1993). Spontaneous spectinomycin-resistant plants were eliminated by double selection on medium containing spectinomycin and streptomycin (500 mg l–1 each) (Bock, 2001; Svab and Maliga, 1993). For each transformation construct, several independent transplastomic lines were subjected to three to four additional rounds of regeneration on spectinomycin-containing medium to enrich the transplastome and select for homoplasmic tissue.

Isolation of nucleic acids and hybridization procedures

Total plant DNA was isolated from fresh leaf tissue by a rapid cetyltrimethylammoniumbromide-based mini-prep procedure (Doyle and Doyle, 1990). RNA was extracted using the PeqGOLD TriFast™ reagent (Peqlab; http://www.peqlab.com) according to the manufacturer’s protocol. For Southern blot analysis, DNA samples (5 μg total DNA) were digested with the restriction enzyme BamHI, separated by gel electrophoresis on 0.8% agarose gels, and transferred onto Hybond XL membranes (Amersham, http://www5.amershambiosciences.com) by capillary blotting using standard protocols. A 283 bp PCR product generated by amplification of the ycf9 coding region using primers Pycf9a (5′-GGCAGGAATGGATCAAAT-3′) and Pycf9b (5′-GGCTATTGGTTTGTTTTGG-3′) was used as an RFLP probe to verify chloroplast transformation. Total cellular RNA samples (10 μg total RNA) were electrophoresed in formaldehyde-containing 1% agarose gels and blotted onto Hybond XL membranes. For detection of nptII and yfp transcripts, the complete coding regions of the respective genes were excised from plasmid clones. All hybridization probes were purified by agarose gel electrophoresis following extraction of the DNA fragments of interest from excised gel slices using the Nucleospin Extract II kit (Macherey-Nagel; http://www.macherey-nagel.com). Probes were radiolabeled with 32P-dCTP using the Megaprime kit (Amersham). Hybridizations were performed at 65°C in Rapid-Hyb buffer (Amersham) according to the manufacturer’s protocol.

Protein extraction and immunoblot analyses

Total soluble protein was extracted from leaf samples homogenized in a buffer containing 50 mM HEPES-KOH (pH 7.5), 10 mM KAC, 5 mM MgAc, 1 mM EDTA, 1 mM DTT, 2 mM PMSF and 1% β-mercaptoethanol. Samples representing 10 μg of extracted proteins were separated by electrophoresis in 15% SDS–polyacrylamide gels, and subsequently transferred to polyvinylidene fluoride membranes (Amersham). NptII protein was detected with a specific anti-NptII antibody generated in rabbits (Linaris GmbH; http://www.linaris.de), and the GFP protein was detected with a monoclonal mouse anti-GFP antibody (JL-8, Clontech). Immunobiochemical detection was performed using the ECL Plus detection system (Amersham) according to the manufacturer’s instructions. Purified recombinant GFP protein (rGFP, BD Biosciences; http://www.bdbiosciences.com) served as a standard.

Microscopy

Subcellular localization of YFP fluorescence was determined by confocal laser scanning microscopy (TCS SP2; Leica: http://www.leica.com) with an argon laser (488 nm). YFP fluorescence was visualized using a 514 nm excitation and a 527 nm emission filter. Chlorophyll fluorescence was detected using a 670–750 nm filter. All images were acquired using a 63 x objective lens.

Acknowledgements

We thank Dr Stephanie Ruf for help with chloroplast transformation, Annett Kassner and Sandra Stegemann for technical assistance, and the MPI-MP Green Team for plant care and cultivation. This research was financed by the Max Planck Society and a grant from the European Union (FP6 Plastomics).

References

Barkan, A. (1988) Proteins encoded by a complex chloroplast transcription unit are each translated from both monocistronic and polycistronic mRNAs. EMBO J. 7, 2637–2644.
Barkan, A. and Goldschmidt-Clermont, M. (2000) Participation of nuclear genes in chloroplast gene expression. Biochimie, 82, 559–572.
Barkan, A., Walker, M., Nolasco, M. and Johnson, D. (1994) A nuclear mutation in maize blocks the processing and translation of several chloroplast mRNAs and provides evidence for the differential translation of alternative mRNA forms. EMBO J. 13, 3170–3181.
Bock, R. (1998) Analysis of RNA editing in plastids. Methods, 15, 75–83.
Bock, R. (2001) Transgenic chloroplasts in basic research and plant biotechnology. J. Mol. Biol. 312, 425–438.
Bock, R. (2006) Extranuclear inheritance: gene transfer out of plastids. Prog. Bot. 67, 75–98.
Bock, R. (2007) Plastid biotechnology: prospects for herbicide and insect resistance, metabolic engineering and molecular farming. Curr. Opin. Biotechnol. 18, 100–106.

Bogorad, L. (2000) Engineering chloroplasts: an alternative site for foreign genes, proteins, reactions and products. Trends Biotechnol. 18, 257–263.

Carposius, A.J., Mudd, E.A. and Krisc, H.M. (1989) Transcription and messenger RNA processing upstream of bacteriophage T4 gene 32. Mol. Gen. Genet. 219, 39–48.

Carrillo, N., Seyer, P., Tyagi, A. and Herrmann, R.G. (1986) Cytochrome b-559 genes from Oenothera hookeri and Nicotiana tabacum show a remarkably high degree of conservation as compared to spinach. Curr. Genet. 10, 619–624.

Chakrabarti, S.K., Lutz, K.A., Lertwiriyawong, B., Svab, Z. and Carpousis, A.J., Mudd, E.A. and Krisch, H.M. (2001) Transgenic chloroplasts in basic research and plant biotechnology. EMBO J. 16, 6804–6811.

Hamtham, S. and Day, A. (2000) Removal of antibiotic resistance genes from transgenic tobacco plastids. Nature Biotechnol. 18, 1172–1176.

Kuroda, H. and Maliga, P. (2001) Complementarity of the 16S rRNA penultimate stem with sequences downstream of the AUG destabilizes the plastid mRNAs. Nucleic Acids Res. 29, 970–975.

Magee, A.M., Horvath, E.M. and Kavanagh, T.A. (2004) Pre-screening plastid transgene expression cassettes in Escherichia coli may be unreliable as a predictor of expression levels in chloroplast-transformed plants. Plant Sci. 166, 1605–1611.

Maliga, P. (2004) Plastid transformation in higher plants. Annu. Rev. Plant Biol. 55, 289–313.

Mayfield, S.P., Yohn, C.B., Cohen, A. and Danon, A. (1995) Regulation of plastid gene expression. Annu. Rev. Plant Physiol. Plant Mol. Biol. 46, 147–166.

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

Meng, B.Y., Tanaka, M., Wakisugi, T., Ohme, M., Shinozaki, K. and Sugii, M. (1988) Cotranscription of the genes encoding two P700 chlorophyll a apoproteins with the gene for ribosomal protein CS14: determination of the transcriptional initiation site by in vitro capping. Curr. Genet. 14, 395–400.

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

Murashige, T. and Skoog, F. (1962) A revised medium for rapid growth and bio assays with tobacco tissue culture. Physiol. Plant. 15, 473–497.

Nakashita, H., Arai, Y., Shikanai, T., Doi, Y. and Yamaguchi, I. (2001) Introduction of bacterial metabolism into higher plants by polycistronic transgene expression. Biosci. Biotechnol. Biochem. 65, 1688–1691.

Nickelsen, J. (1999) Transcripts encoding the 5′ untranslated regions of the plastid genes psbA and psbB from higher plants are unstable in Chlamydomonas reinhardtii chloroplasts. Mol. Gen. Genet. 262, 768–771.

Nickelsen, J., van Dillewijn, J., Rahire, M. and Rochaix, J.-D. (1994) Determinants for stability of the chloroplast psbD RNA are located within its short leader region in Chlamydomonas reinhardtii. EMBO J. 13, 3182–3191.

Nickelsen, J., Fleischmann, M., Boudreau, E., Rahire, M. and Rochaix, J.-D. (1999) Identification of cis-acting RNA leader elements required for chloroplast psbD gene expression in Chlamydomonas. Plant Cell 11, 957–970.

Oesada-Vargas, T., Ruiz, O.N. and Daniell, H. (2005) Characteriza- tion of heterologous multigene operons in transgenic chloro- plastics. Transcription, processing, and translation. Plant Physiol. 138, 1746–1762.

Rogalski, M., Ruf, S. and Bock, R. (2006) Tobacco plastid ribosomal protein S18 is essential for cell survival. Nucleic Acids Res. 34, 4537–4545.

Ruf, S., Biehler, K. and Bock, R. (2000) A small chloroplast-encoded protein as a novel architectural component of the light-harvesting antenna. J. Cell Biol. 149, 369–371.

Ruf, S., Hermann, M., Berger, I.J., Carrer, H. and Bock, R. (2001) Stable genetic transformation of tomato plastids and expression of a foreign protein in fruit. Nature Biotechnol. 19, 870–875.

Sakamotu, W., Sturm, N.R., Kindle, K.L. and Stern, D.B. (1994) petD mRNA maturation in Chlamydomonas reinhardtii chloroplasts: role of 5′ endonucleolytic processing. Mol. Cell. Biol. 14, 6180–6186.

© 2007 The Authors
Journal compilation © 2007 Blackwell Publishing Ltd, The Plant Journal, (2007), 52, 961–972
Translation of the psbA mRNA is regulated by light via the 5'-untranslated region in tobacco plastids. *Plant J.* 6, 547–553.

Expression of a chimeric uidA gene indicates that polycistronic mRNAs are efficiently translated in tobacco plastids. *Plant J.* 7, 845–848.

Control of plastid gene expression: 3' inverted repeats act as mRNA processing and stabilizing elements but do not terminate transcription. *Cell* 51, 1145–1157.

Regulation of gene expression in chloroplasts of higher plants. *Plant Mol. Biol.* 32, 315–326.

The chloroplast genome. *Plant Mol. Biol.* 19, 149–168.

High-frequency plastid transformation in tobacco by selection for a chimeric aadA gene. *Proc. Natl Acad. Sci. USA* 90, 913–917.

Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. *Nature Rev. Genet.* 5, 123–136.

Complex RNA maturation in chloroplasts. *Eur. J. Biochem.* 171, 551–564.

Two small open reading frames are co-transcribed with the pea chloroplast genes for the polypeptides of cytochrome b-559. *Curr. Genet.* 15, 213–220.

An open reading frame encoding a putative haem-binding polypeptide is cotranscribed with the pea chloroplast gene for apocytochrome f. *Plant Mol. Biol.* 15, 347–356.

contained metabolic engineering in tomatoes by expression of carotenoid biosynthesis genes from the plastid genome. *Plant J.* 49, 276–288.

Engineering the provitamin A (β-carotene) biosynthetic pathway into (carotenoid-free) rice endosperm. *Science* 287, 303–305.

Surprising features of plastid ndhD transcripts: addition of non-encoded nucleotides and polysome association of mRNAs with an unedited start codon. *Nucleic Acids Res.* 32, 542–550.