RARE PATERNAL PLASTID INHERITANCE IN ARABIDOPSIS

by

ARUN K. AZHAGIRI

A Dissertation submitted to the

Graduate School-New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Plant Biology

written under the direction of

Professor Pal Maliga

and approved by

_____

_____

_____

_____

_____

New Brunswick, New Jersey

[October, 2007]

ABSTRACT OF THE DISSERTATION

RARE PATERNAL PLASTID INHERITANCE IN ARABIDOPSIS

by Arun K. Azhagiri

Dissertation Director: Dr. Pal Maliga

Plastids and mitochondria, in a standard genetic cross, are transmitted to the seed progeny by the maternal parent in *Arabidopsis thaliana*. My objective was to test, if exceptional pollen transmission of plastid occurs in Arabidopsis, a species in which no plastid DNA could be detected in pollen or sperm cells by cytological methods. The maternal parent was the nuclear male sterile (*ms1-1/ms1-1*), spectinomycin sensitive Ler ecotype and the pollen parent was the male fertile RLD-Spc1 plant carrying a plastid-encoded spectinomycin resistance mutation. I selected for exceptional pollen transmission in the progeny by spectinomycin resistance encoded in the paternal plastid DNA. I found that plastids, in general, are inherited maternally in Arabidopsis and rare events of paternal plastid transmission to the seed progeny occurs at a low (3.9×10^{-5}) frequency. This observation extends previous reports in *Nicotiana tabacum* (family: Solanaceae) to a cruciferous species suggesting that low frequency paternal leakage of plastids *via* pollen may be universal in plants previously thought to exhibit strict maternal plastid inheritance.

Two components needed to accomplish this study were a plastid-encoded, spectinomycin resistant mutant as the pollen parent and plastid markers to identify the origin of the plastids in the hybrid seed progeny obtained from the crosses. To identify genetic markers in the *Arabidopsis thaliana* plastid genome (ptDNA) I amplified and sequenced the *rpl2-psbA* and *rbcL-accD* regions in 26 ecotypes. The two regions contained eight polymorphic sites including five insertions and/or deletions (indels) involving changes in the length of A or T mononucleotide repeats and three base substitutions. The 27 alleles provided a practical set of ptDNA markers for the commonly used RLD, Ler, Col and C24 ecotypes. I used a RLD-Spc1 line carrying a plastid-encoded spectinomycin resistant mutation as the pollen parent that was compatible with the Ler maternal parent.

ACKNOWLEDGEMENTS

I thank Dr. Pal Maliga for the advice and guidance he provided during my period as a graduate student in his laboratory. I am also grateful for the discussions regarding my work and the encouragement he provided, which helped me to persevere at times when my work seemed to progress slowly.

My tenure in this laboratory was made pleasant due to my friends, both past and present. I thank them for their camaraderie and conversations regarding the science pursued in the laboratory and on other sundry topics. My gratitude goes to Zora Svab, the “mother” of the laboratory, who takes all the students under her wing and provides them with encouragement and guidance during their stay. Finally, my special thanks to Kerry Lutz for stimulating discussions and efficient management of the laboratory.

I dedicate this dissertation to my Father who would have said, “Finally!”

To my wife, Mamatha, who stood by me during this time, gave me encouragement and never let me loose focus, I have no words that can express my feelings. To my brother, Ashok, and his family who helped keep me aware of the “real” world outside, for the conversations and late night coffees, I am grateful and indebted forever. I also take this opportunity to express my gratitude to my Mother and Anup, who held the fort back home.

TABLE OF CONTENTS

Appendix 3

Lists of tables

List of Figures

CHAPTER 1

Introduction

Natural containment due to lack of pollen transmission has been claimed to be one advantage of introducing transgenes into the plastid genome. However, doubts regarding the effectiveness of using plastids for transgene containment were raised. The basis for the doubts came from studies showing high frequency (up to 2.5% in tobacco) paternal plastid DNA transmission *via* pollen in species thought to inherit plastids strictly from the maternal parent. However, recent studies have shown that the results are line dependent and the frequencies are 100-1000 fold lower (Ruf et al., 2007; Svab and Maliga, 2007). My objective was to determine if rare paternal transmission of ptDNA *via* pollen occurs in *Arabidopsis thaliana* and the frequency at which paternal plastids are transmitted to the seed progeny. I hoped to achieve this by crossing a nuclear male sterile, spectinomycin sensitive L*er* line as the maternal parent and a male fertile, plastid-encoded spectinomycin resistant mutant RLD line as the pollen parent. Selection, in hybrid seed progeny, for paternal plastid-encoded spectinomycin resistance in tissue culture enabled the identification of rare events involving paternal transmission of plastids. The tools developed to carry out this study were; (a) Plastid genetic markers, to differentiate between the paternal (pollen) parent (RLD) and the maternal parent (L*er*), to determine the origin of the plastid DNA in the hybrid seed progeny. (b) Plastid-encoded spectinomycin resistant Arabidopsis mutant as pollen parent to enable selection for paternal plastid DNA (ptDNA) in progeny. While proceeding to isolate spectinomycin resistant mutants in the L*er* background using a post-transcriptionally regulated BABY BOOM (BBM), I made use of a plastid-spectinomycin resistant RLD line, available in the laboratory (Dr. Marina Skarjinskaia, unpublished data). This mutant was compatible with my experiment, since the maternal parent was from the L*er* ecotype and using a RLD spectinomycin resistant mutant as the pollen parent would aid to differentiate the plastids from the two parents. I also worked on isolating plastid-encoded spectinomycin resistant mutants in the Ws ecotype using a transcriptionally regulated WUSCHEL (*WUS*) gene (Zuo et al., 2002). However, in my experience the WUSHCEL system was not useful in maintaining embryogenic capacity in long term tissue culture. Therefore, I did not pursue the project to completion. Further, my interest in developing plastid

transformation in Arabidopsis gave rise to the data in Appendix 1, Appendix 2 and Appendix 3. These experiments were designed to develop alternate techniques that would increase plastid transformation efficiency in Arabidopsis.

Specific objectives

The specific objectives of my research were:

i. Determine if rare transmission of paternal plastids *via* pollen occurs, and if so at what frequency.

ii. Identify polymorphisms in the plastid genomes of Arabidopsis ecotypes that could be used as plastid genetic markers to distinguish between the maternal parent and the pollen parent.

iii. Isolate plastid-encoded spectinomycin resistant mutant to be used as a pollen parent.

Literature Review

Plastid inheritance in higher plants

I chose to study plastid inheritance in *Arabidopsis thaliana* because it is the most advanced plant model species due to its small size, rapid life cycle and its nuclear (Arabidopsis Genome Initiative, 2000), plastid (Sato et al., 1999) and mitochondrial (Unseld et al., 1997) genomes sequenced. Plastids and mitochondria of higher plants are inherited maternally, paternally or from both parents, dependent on the taxonomic group (Reboud and Zeyl, 1994; Mogensen, 1996; Hagemann, 2002). Solanaceous species, in general, inherit plastids from the maternal parent. In *Petunia hybrida* maternal inheritance of plastids was shown in crosses of 16 inbred lines. However, transmission of paternal plastids in 0.1 to 2% of the seed progeny could be readily detected in six out of 22 inbred lines indicating variability for the genetic control of paternal plastid inheritance. Paternal plastids were identified as green sectors on the maternal virescent mutant background in the seed progeny (Cornu and Dulieu, 1988; Derepas and Dulieu, 1992). Maternal inheritance of plastids was also shown in *Nicotiana tabacum* using plastid-encoded streptomycin and spectinomycin resistance mutations or a pigment mutation as genetic markers (Svab et al., 1990; Kanevski

and Maliga, 1994). However, when tissue culture selection was carried out for the paternal spectinomycin resistance marker, exceptional paternal pollen transmission events were found in 0.01% to 0.001% (10^{-4} to 10^{-5}) of the progeny (Medgyesy et al., 1986). Since recovery of paternal plastids was heavily biased by tissue culture selection, it is estimated that visible sectors in plants would appear at most in one out of 100,000 or 1,000,000 progeny (Ruf et al., 2007; Svab and Maliga, 2007). Paternal ptDNA in alloplasmic Nicotiana lines was transmitted via pollen at a higher (up to 2.5%) frequency (Medgyesy et al., 1986), (Horlow et al., 1990; Avni and Edelman, 1991). In *Arabidopsis thaliana* maternal inheritance of plastids was shown in a reciprocal cross between the L*er*-0 and Cvi-0 ecotypes by ptDNA markers (Martinez et al., 1997). Lack of ptDNA in generative cells of pollen is in accordance with maternal inheritance of plastids in *Arabidopsis* (Nagata et al., 1999). The numbers of individuals observed in these studies were few and if paternal inheritance is a rare event, then these events could have been missed. Current progress in Arabidopsis genetics enables the use of plastid genetic markers and plastid-encoded antibiotic resistant mutants to study adequate numbers and in a more stringent manner as described in Chapter 2.

Identification of ptDNA markers

In Chapter 3, I report the identification of ptDNA markers that enable the differentiation of plastids between the maternal L*er* and paternal RLD ecotypes. Evolution of DNA sequence within the protein coding regions is subject to functional constrains. However, simple mononucleotide repeats in the intergenic regions of the plastid genome arising due to replication slippage are potential sources of variability. This attribute was successfully used to reveal intraspecies ptDNA variation in Arabidopsis (Powell et al., 1995b; Provan et al., 2001). Earlier studies on ptDNA polymorphism in *Arabidopsis thaliana* focused on the evolutionary origin and expansion of Arabidopsis populations in Northern Europe but provided markers in only a few commonly studied ecotypes (Provan, 2000; Provan and Campanella, 2003) (Sall et al., 2003), (Martinez et al., 1997). It appears that the degree of ptDNA variation is dependent on the taxonomic group. A high degree of intra-species ptDNA variability was found using microsatellite markers in pines (Powell et al., 1995a) and soybean (Powell et al., 1995b). A high degree of ptDNA diversity was also found when comparing the entire 134.5-kb plastid genome sequences of the *indica* and *japonica*

subspecies of rice (Tang et al., 2004). In this study 72 single nucleotide polymorphisms and 27 insertions and deletions have been identified. In contrast, there is relative lack of polymorphism in the Solanaceae family, in which no nucleotide difference was found upon sequencing the 155.5-kb plastid genome of two tomato cultivars (Kahlau et al., 2006). Also, comparison of the 155.9 kb plastid genomes of *Nicotiana tabacum* and its maternal progenitor, *Nicotiana sylvestris*, revealed only seven polymorphic sites, four nucleotide substitutions and three indels (Yukawa et al., 2006). Variability in ptDNA could experimentally be detected in ~ 60 % of Oenothera lines carrying a nuclear gene inducing plastid mutations (plastome mutator), a frequency that is 200- to 1000-fold higher than natural variation (Sears and Sokalski, 1991; Stoike and Sears, 1998).

Isolation of fertile, plastid spectinomycin resistant mutant using a steroid-inducible BABY BOOM plant regeneration system

In Chapter 4, I describe the isolation of spectinomycin resistant mutants to enable me to select for rare paternal transmission events. Isolation of plastid spectinomycin resistant mutants has been shown in tobacco and *Solanum nigrum* (Fromm et al., 1987; Svab and Maliga, 1991; Kavanagh et al., 1994). Being aware that regeneration of fertile plants from tissue culture was a problem in Arabidopsis (Sikdar et al., 1998), I used a new plant regeneration system based on the BABY BOOM (BBM) transcription factor to isolate spectinomycin resistant mutants.

Constitutive expression of BABY BOOM (BBM), a member of the AP2/ERF family of transcription factors (Riechmann et al., 2000; Nole-Wilson et al., 2005), was reported to sustain spontaneous production of somatic embryos but did not yield fertile plants (Boutilier et al., 2002). Therefore regulating BABY BOOM function by fusing BBM with the glucocorticoid receptor steroid-binding domain could overcome the bottleneck. Although the fusion proteins are constitutively expressed, transcription of the downstream target genes is dependent on the supply of steroid hormones in the culture medium. In the absence of the steroid hormone, the receptor associates with cellular regulatory proteins, including Hsp90, and becomes anchored in the cytosol as a monomer. Association of steroid with the hormone-binding domain leads to the release of HSP90 from the receptor. The receptor subsequently

dimerizes, translocates into the nucleus and binds to the target DNA to activate transcription (Zuo and Chua, 2000).The use of such systems has been successfully employed to regulate expression various genes (Lloyd et al., 1994), (Wagner et al., 1999), (Gallois et al., 2002) (Hay et al., 2003); and recently BABY BOOM (Srinivasan et al., 2007). Such regulated expression of BBM would enable the production of fertile plastid-encoded spectinomycin transformants in Arabidopsis. BBM is not the only regulated plant regeneration system in Arabidopsis. The WUSCHEL gene, in a transcriptionally regulated system, promoted vegetative-to-embryonic transition and yielded fertile plants after the removal of the inducer (Zuo et al., 2002). LEC1 and LEC2 are two additional seed-expressed transcription factor genes that promote spontaneous embryo formation on vegetative tissues (Lotan et al., 1998; Stone et al., 2001). Interestingly, regulated over-expression of LEC1 cited in ref. (Zuo et al., 2002) and induction of a LEC2:GR fusion (Santos Mendoza et al., 2005) did not result in formation of embryo-like structures.

References

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

Avni, A. and Edelman, M. (1991) Direct selection for paternal inheritance of chloroplasts in sexual progeny of *Nicotiana*. *Mol Gen Genet* **225,** 273-277.

Boutilier, K., Offringa, R., Sharma, V.K., Kieft, H., Ouellet, T., Zhang, L., Hattori, J., Liu, C.M., van Lammeren, A.A.M., Miki, B.L.A., Custers, J.B.M. and van Lookeren Campagne, M.M. (2002) Ectopic expression of BABY BOOM triggers a conversion of vegetative to embryonic growth. *Plant Cell* **14,** 1737-1749.

Cornu, A. and Dulieu, H. (1988) Pollen transmission of plastid DNA under genotypic control in *Petunia hybrida* Hort. *J Hered* **79,** 40-44.

Derepas, A. and Dulieu, H. (1992) Inheritance of the capacity to transfer plastids by the pollen parent in *Petunia hybrida* Hort. *J Hered* **83,** 6-10.

Fromm, H., Edelman, M., Aviv, D. and Galun, E. (1987) The molecular basis for rDNA-dependent spectinomycin resistance in *Nicotiana* chloroplasts. *EMBO J* **6,** 3233-3237.

Gallois, J.L., Woodward, C., Reddy, G.V. and Sablowski, R. (2002) Combined SHOOT MERITEMLESS and WUSCHEL trigger ectopic organogenesis in *Arabidopsis*. *Development* **129,** 3207-3217.

Hagemann, R. (2002) Milestones in plastid genetics of higher plants. *Prog Bot* **63,** 5-51.

Hay, A., Jackson, D., Ori, N. and Hake, S. (2003) Analysis of the competence to respond to KNOTTED1 activity in Arabidopsis leaves using a steroid induction system. *Plant Physiol* **131,** 1671-1680.

Horlow, C., Goujaud, J., Lepingle, A., Missonier, C. and Bourgin, J.P. (1990) Transmission of paternal chloroplasts in tobacco (*Nicotiana tabacum*). *Plant Cell Rep* **9,** 249-252.

Kahlau, S., Aspinall, S., Gray, J.C. and Bock, R. (2006) Sequence of the tomato chloroplast DNA and evolutionary comparison of solanaceous plastid genomes. *J Mol Evol* **63,** 194-207.

Kanevski, I. and Maliga, P. (1994) Relocation of the plastid *rbcL* gene to the nucleus yields functional ribulose-1,5-bisphosphate carboxylase in tobacco chloroplasts. *Proc Natl Acad Sci U S A* **91,** 1969-1973.

Kavanagh, T.A., O'Driscoll, K.M., McCabe, P.F. and Dix, P.J. (1994) Mutations conferring lincomycin, spectinomycin, and streptomycin resistance in *Solanum nigrum* are located in three different chloroplast genes. *Mol Gen Genet* **242,** 675-680.

Lloyd, A.M., Schena, M., Walbot, V. and Davis, R.W. (1994) Epidermal cell fate determination in Arabidopsis: patterns defined by a steroid-inducible regulator. *Science* **266,** 436-439.

Lotan, T., Ohto, M., Yee, K.M., West, M.A., Lo, R., Kwong, R.W., Yamagishi, K., Fischer, R.L., Goldberg, R.B. and Harada, J.J. (1998) Arabidopsis LEAFY COTYLEDON1 is sufficient to induce embryo development in vegetative cells. *Cell* **93,** 1195-1205.

Martinez, P., Lopez, C., Roldan, M., Sabater, B. and Martin, M. (1997) Plastid DNA of five ecotypes of *Arabidopsis thaliana*: sequence of *ndhG* gene and maternal inheritance. *Plant Sci* **123,** 113.

Medgyesy, P., Pay, A. and Marton, L. (1986) Transmission of paternal chloroplasts in *Nicotiana*. *Mol Gen Genet* **204,** 195-198.

Mogensen, H.L. (1996) The hows and whys of cytoplasmic inheritance in seed plants. *Amer J Bot* **83,** 383-404.

Nagata, N., Saito, C., Sakai, A., Kuroiwa, H. and Kuroiwa, T. (1999) The selective increase or decrease of organellar DNA in generative cells just after pollen mitosis one controls cytoplasmic inheritance. *Planta* **209,** 53-65.

Nole-Wilson, S., Tranby, T.L. and Krizek, B.A. (2005) *AINTEGUMENTA*-like (*AIL*) genes are expressed in young tissues and may specify meristematic or division-competent states. *Plant Mol Biol* **57,** 613-628.

Powell, W., Morgante, M., McDevitt, R., Vendramin, G.G. and Rafalski, J.A. (1995a) Polymorphic simple sequence repeat regions in chloroplast genomes: applications to the population genetics of pines. *Proc Natl Acad Sci U S A* **92,** 7759-7763.

Powell, W., Morgante, M., Andre, C., McNicol, J.W., Machray, G.C., Doyle, J.J., Tingey, S.V. and Rafalski, J.A. (1995b) Hypervariable microsatellites provide a general source of polymorphic DNA markers for the chloroplast genome. *Curr Biol* **5,** 1023-1029.

Provan, J. (2000) Novel chloroplast microsatellites reveal cytoplasmic variation in *Arabidopsis thaliana*. *Mol Ecol* **9,** 2183-2185.

Provan, J. and Campanella, J.J. (2003) Patterns of cytoplasmic variation in *Arabidopsis thaliana* (Brassicaceae) revealed by polymorphic chloroplast microsatellites. *Syst Bot* **28,** 578-583.

Provan, J., Powell, W. and Hollingsworth, P.M. (2001) Chloroplast microsatellites: new tools for studies in plant ecology and evolution. *Trends Ecol Evol* **16,** 142-147.

Reboud, X. and Zeyl, C. (1994) Organelle inheritance in plants. *Heredity* **72,** 132-140.

Riechmann, J.L., Heard, J., Martin, G., Reuber, L., Jiang, C.-Z., Keddie, J., Adam, L., Pineda, O., Ratcliffe, O.J., Samaha, R.R., Creelman, R., Pilgrim, M., Broun, P., Zhang, J.Z., Ghandehari, D., Sherman, B.K. and Yu, G.L. (2000) Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes. *Science***,** 2105-2110.

Ruf, S., Karcher, D. and Bock, R. (2007) Determining the transgene containment level provided by chloroplast transformation. *Proc Natl Acad Sci USA* **104,** 6998-7002.

Sall, T., Jakobsson, M., Lind-Hallden, C. and Hallden, C. (2003) Chloroplast DNA indicates a single origin of the allotetraploid *Arabidopsis suecica*. *J Evol Biol* **16,** 1019-1029.

Santos Mendoza, M., Dubreucq, B., Miquel, M., Caboche, M. and Lepiniec, L. (2005) LEAFY COTYLEDON 2 activation is sufficient to trigger the accumulation of oil and seed specific mRNAs in Arabidopsis leaves. *FEBS Lett* **579,** 4666-4670.

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

Sears, B.B. and Sokalski, M.B. (1991) The Oenothera plastome mutator: effect of UV irradiation and nitroso-methyl urea on mutation frequencies. *Mol Gen Genet* **229,** 245-252.

Sikdar, S.R., Serino, G., Chaudhuri, S. and Maliga, P. (1998) Plastid transformation in *Arabidopsis thaliana*. *Plant Cell Rep* **18,** 20-24.

Srinivasan, C., Liu, Z., Heidmann, I., Supena, E.D., Fukuoka, H., Joosen, R., Lambalk, J., Angenent, G., Scorza, R., Custers, J.B. and Boutilier, K. (2007) Heterologous expression of the BABY BOOM AP2/ERF transcription factor enhances the regeneration capacity of tobacco (*Nicotiana tabacum* L.). *Planta* **225,** 341-351.

Stoike, L.L. and Sears, B.B. (1998) Plastome mutator-induced alterations arise in Oenothera chloroplast DNA through template slippage. *Genetics* **149,** 347-353.

Stone, S.L., Kwong, L.W., Yee, K.M., Pelletier, J., Lepiniec, L., Fischer, R.L., Goldberg, R.B. and Harada, J.J. (2001) LEAFY COTYLEDON2 encodes a B3 domain transcription factor that induces embryo development. *Proc Natl Acad Sci U S A* **98,** 11806-11811.

Svab, Z. and Maliga, P. (1991) Mutation proximal to the tRNA binding region of the *Nicotiana* plastid 16S rRNA confers resistance to spectinomycin. *Mol Gen Genet* **228,** 316-319.

Svab, Z. and Maliga, P. (2007) Exceptional transmission of plastids and mitochondria from the pollen parent and its impact on transgene containment. *Proc Natl Acad Sci U S A* **104,** 7003-7008.

Svab, Z., Hajdukiewicz, P. and Maliga, P. (1990) Stable transformation of plastids in higher plants. *Proc Natl Acad Sci U S A* **87,** 8526-8530.

Tang, J., Xia, H., Cao, M., Zhang, X., Zeng, W., Hu, S., Tong, W., Wang, J., Wang, J., Yu, J., Yang, H. and Zhu, L. (2004) A comparison of rice chloroplast genomes. *Plant Physiol* **135,** 412-420.

Unseld, M., Marienfeld, J.R., Brandt, P. and Brennicke, A. (1997) The mitochondrial genome of *Arabidopsis thaliana* contains 57 genes in 366,924 nucleotides. *Nature Genet* **15,** 57-61.

Wagner, D., Sablowski, R.W. and Meyerowitz, E.M. (1999) Transcriptional activation of APETALA1 by LEAFY. *Science* **285,** 582-584.

Yukawa, M., Tsudzuki, T. and Sugiura, M. (2006) The chloroplast genome of *Nicotiana sylvestris* and *Nicotiana tomentosiformis*: complete sequencing confirms that the *Nicotiana sylvestris* progenitor is the maternal genome donor of *Nicotiana tabacum*. *Mol Genet Genomics* **275,** 367-373.

Zuo, J. and Chua, N.H. (2000) Chemical-inducible systems for regulated expression of plant genes. *Curr Opin Biotechnol* **11,** 146-151.

Zuo, J., Niu, Q.W. and Chua, N.H. (2002) The *WUSCHEL* gene promotes vegetative-to-embryonic transition in *Arabidopsis*. *Plant J* **30,** 349-359.

CHAPTER 2

Rare Paternal Plastid Inheritance in Arabidopsis

A study to detect rare events involving transmission of paternal plastid DNA (ptDNA) *via* pollen in *Arabidopsis thaliana* is reported here. 76,825 seed progeny from a cross between the nuclear male sterile, spectinomycin sensitive, L*er* maternal parent and fertile, spectinomycin resistant; RLD pollen parent were screened for the presence of paternal ptDNA. Rare events of plastid transfer through pollen to progeny were detected as green sectors in seedling calli on spectinomycin containing media. Based on ptDNA markers, all resistant clones carried RLD plastids indicating that spectinomycin resistance in the three lines is the result of pollen transmission of plastids. The possibility that the selected spectinomycin resistant lines were the result of contamination with RLD-Spc1 seed was excluded by showing that the clones were heterozygotes for nuclear markers. We report here that paternal plastids are transmitted to the seed progeny in Arabidopsis at a low (3.9×10^{-5}) frequency. This observation extends previous report in *Nicotiana tabacum* (family: Solanaceae) to a cruciferous species suggesting that low frequency paternal leakage of plastids *via* pollen may be universal in plants.

Results

Experimental design

To produce hybrid seed in sufficient numbers and avoid the need for hand emasculation, we chose a nuclear male sterile mutant *ms1-1/ms1-1* (spectinomycin sensitive) in the L*er* ecotype as maternal parent (L*er-ms1-1*) (van der Veen and Wirtz, 1968; Wilson et al., 2001) (Figure 2.1a). The paternal parent was the RLD-Spc1 spectinomycin resistant mutant carrying a point mutation in the 16S rRNA gene (*rrn16*) available in the RLD ecotype (Figure 2.1a). Hybrid seed was obtained by pollinating homozygous, male sterile L*er-ms1-1*flowers with RLD-Spc1 pollen (Figure 2.1b).

Paternal ptDNA transmitted by pollen may be present in any part of a seedling. We decided to identify spectinomycin resistant sectors in seedling callus because callus formation enables prolific cell division from all cell types in a seedling, facilitating the recovery of rare paternal ptDNA. Seedlings derived from the cross were allowed to germinate on a selective spectinomycin medium where sensitive seedlings form white callus due to selective inhibition of protein synthesis on the prokaryotic type plastid ribosomes and cells carrying exceptional paternal plastids form green sectors on the sensitive white seedling calli. We planned to distinguish paternal (RLD) ptDNA transmission events from spontaneous L*er* spectinomycin resistant mutants by the ecotype-specific plastid RFLP markers (Azhagiri and Maliga, 2007).

To verify that the protocol is suitable for the recovery of paternal spectinomycin resistant plastids, spectinomycin resistant mutants were selected in the hybrid background. In callus derived from 1,000 mutagenized hybrid L*er-ms1-1* x RLD-Spc1 seeds (2.5 mM N-nitroso-N-methyl-urea) five green calli were identified on a selective medium containing 100 mg/l spectinomycin (Figure 2.2a). To verify that the resistant clones were new mutants, a part of the *rrn16* gene, known to encode spectinomycin resistance (Svab and Maliga, 1991) was PCR amplified and sequenced. DNA sequencing identified nucleotide exchanges in the *rrn16* gene, which are distinct from the RLD-Spc1 mutation (C to T) (Figure 2.2c). All five mutations induced by NMU were G to A exchanges affecting an *Aat*II site. Bias in favor of G to A mutations is not surprising, since the main points of NMU alkylating action are the N7 and O6 positions of

guanidine (Hagemann, 1982). Thus, the feasibility of identifying paternal spectinomycin resistant plastids in the seed progeny was directly confirmed.

Selection for paternal spectinomycin resistance in the seed progeny

To identify seedlings that carry exceptional paternal plastids, hybrid seeds were surface sterilized, and germinated on ARMI callus induction medium (Marton and Browse, 1991) containing 100 mg/L spectinomycin. The calli were individually transferred to a fresh plate every fortnight three times. Arabidopsis callus on ARMI medium forms green callus in the absence of spectinomycin. Spectinomycin prevents greening, unless the paternal spectinomycin resistant plastids have been transmitted via pollen to the progeny (Figure 2.3a). Altogether three spectinomycin resistant calli were identified among 76,825 calli selected on spectinomycin medium. Two of these, At-PSpc1-1 and At-PSpc1-2, were found during the 3rd subculture (weeks 6-8) and the 3rd line, At-PSpc1-3, was noticed during the 4th subculture (weeks 8-10). Identification of spectinomycin resistant clones with a delay, after six to eight weeks callus induction, is compatible with both: paternal pollen transmission and identification of a new spontaneous mutant. If only a few paternal ptDNA copies are transmitted by pollen, it may take several rounds of ptDNA replication and plastid and cell division for a visible green sector to form (Figure 2.3a). Formation of green sectors is facilitated by the lack of exact duplication of the cytoplasm during cell division (Figure 2.3c) and preferential maintenance of resistant plastids with normal translation of mRNAs on the resistant ribosomes (Moll et al., 1990). Likewise, formation of mutant sectors from an initial mutant copy also takes time and involves replication and sorting of the mutant ptDNA over several cell generations.

Confirming paternal ptDNA transmission and hybrid nature of resistant progeny

Spectinomycin resistant lines identified in the seedling callus may be the result of pollen transmission of paternal ptDNA, or new mutations in the plastid genome. If spectinomycin resistance is the result of pollen transmission of paternal ptDNA, spectinomycin resistance should be due to a C to T mutation affecting the 1st C of the *Aat*II site, and carried on an RLD ptDNA. On the other hand, a spontaneous spectinomycin

resistant mutant should have L*er*-specific ptDNA markers. The RLD and L*er* ecotypes can be distinguished by an amplified length polymorphic (AFLP) marker in the *rpl2-psbA* intergenic regions and two RFLP markers (*Xba*I, *Mse*I) in the *rbcL-accD* intergenic region (Azhagiri and Maliga, 2007). All three lines carry RLD plastids (Figure 2.4a) indicating that spectinomycin resistance in the three lines is the result of pollen transmission of plastids. In accordance, the *Aat*II site is absent in the *rrn16* gene (Figure 2.4a) and DNA sequencing confirmed the presence of RLD-SpcI specific spectinomycin resistance mutation in all three clones (Figure 2.4b). One of the clones, At-PSpc1-2, carried both RLD- and L*er*-specific ptDNA markers indicating, that the DNA sample contained both plastid types indicating input from both the maternal and the paternal parents.

The possibility that the selected spectinomycin resistant lines were the result of contamination with RLD-Spc1 seed was excluded by showing that the clones were heterozygotes for markers GENEA, nga126, nga8, nga158 (www.arabidopsis.org) and the *ms1-1* mutation (Wilson et al., 2001) on chromosomes I, III, IV and V, respectively (Figure 2.4c).

Discussion

We report here that plastids are transmitted to the progeny by pollen at a low (3.9×10^{-5}) frequency in *Arabidopsis thaliana*. This observation extends previous reports in *Nicotiana tabacum* (family: Solanaceae) to a cruciferous species suggesting that low frequency paternal leakage of plastids *via* pollen may be universal in plants. The frequency of paternal pollen transmission is low, and probably involves only a very few ptDNA copies. These could be detected only because a sensitive tissue culture selection was employed for spectinomycin resistance encoded in the paternal ptDNA. Such rare ptDNA copies would have remained undetected in genetic crosses showing maternal inheritance of ptDNA in Arabidopsis (Martinez et al., 1997). These few copies are also below the detection limit of careful cytological observations, which indicated that plastid and mitochondrial DNA rapidly disappears in generative cells after pollen mitosis one, in line with the expectation of maternal inheritance of both plastids and mitochondria in *Arabidopsis thaliana* (Nagata et al., 1999). The low frequency paternal ptDNA transmission shown here in Arabidopsis and in normal (non-alloplasmic) tobacco lines (Ruf et al., 2007; Svab and Maliga, 2007) should be viewed as an exception to the rule, due to a malfunctioning of the biological mechanisms that exclude paternal pollen transmission in the majority of higher plant species. The mechanism of paternal inheritance in such a case is not necessarily the same as that functioning during high frequency biparental plastid inheritance. A mechanism for such exceptional cases could be rare inclusion of organelles in sperm cell nuclei (Yu and Russell, 1994).

In dicots such as Arabidopsis layer 2 of the shoot apex contributes the cells that form the germ layer (Poethig, 1989). Thus, paternal plastids in the hybrid will be inherited maternally only if they get established in the shoot apex and this depends on the anatomical position of the cells that carry them. The anatomical position of the carrier cells in the present study could not be determined. However, in a similar tobacco study, it was estimated that, at best, one in ten paternally transmitted plastids are transmitted by seed (Ruf et al., 2007; Svab and Maliga, 2007). Thus, the probability that paternal plastids will be transmitted to the seed progeny is very low.

Diversity in the mode of organelle inheritance within families (Havey et al., 1998), or even species (Derepas and Dulieu, 1992), suggests that organelle inheritance is amenable to genetic analyses.

Investigation of plastid inheritance with sensitive genetic markers and molecular tools in Arabidopsis is the first step towards the molecular genetic dissection of mechanisms controlling organelle inheritance.

Experimental procedures

Plant material

The maternal parent in the crosses was the L*er*-ms1-1 nuclear male sterile mutant (van der Veen and Wirtz, 1968). This mutant is characterized by the lack of viable pollen due to a G to A transition at exon2-intron2 junction in the *ms1* gene that leads to a mis-spliced transcript (Wilson et al., 2001). The seed (CS75) was obtained from the Arabidopsis Biological Research Center, Columbus, OH. The pollen parent At-RLD-Spc1 mutant is resistant to spectinomycin due to a mutation in the plastid 16S small rRNA gene (Marina Skarjinskaia, unpublished). The mutation is a C to T nucleotide exchange at position 102,151 in the Arabidopsis ptDNA (GenBank Accession No. AP000423). Hybrid seed was obtained by manual pollination of the sterile flowers on L*er-ms1-1/ms1-1* plants homozygous for the *ms1-1* mutation.

Tissue culture selection

The hybrid Arabidopsis seeds were vapor sterilized (Clough and Bent, 1998) and germinated on a selective ARMI medium containing 100 mg/L spectinomycin HCl. Resistance to spectinomycin manifested as green sectors on the white sensitive calli. The ARMI medium was essentially as described (Marton and Browse, 1991). The ARMI medium contains MS salts, 3% Sucrose, 0.8% agar, 1ml/L vitamin solution (200mg myo-inositol, 10mg vitamin B1, 1mg vitamin B6, 1mg nicotinic acid, 1mg glycine, 0.1mg biotin per ml), 3mg/L indolacetic acid, 0.15mg/L 2,4-D (2,4-dichlorophenoxyacetic acid), 0.6mg/L BA (Benzyladenine) and 0.3mg/L IPA (isopentyladenosine). The seeds spread on a solid ARMI medium were vernalized at 4°C for 4 days and then grown at 22 °C/18 °C (day/night temperature) and 16-hr light/8-hr dark cycle. The calli were transferred to new medium every two weeks. Plants were maintained on ARM5 medium (Sikdar et al., 1998).

PCR and sequencing

Genomic DNA was extracted from calli according to the CTAB protocol (Murray and Thompson, 1980) and PCR amplification was carried out using primers listed in Table 2.1. Amplification of 50 ng total DNA was carried out in a 50µl volume with Platinum ® *Pfx* DNA Polymerase (Invitrogen, Carlsbad, CA, USA) using the following conditions: 5 min 94°C, 35 cycles of 30 s 94°C, 30 s 55°C, 45 s 72°C and 1 cycle of 7 min 72°C. Amplified DNA fragments were purified using QIAquick Gel Extraction Kit (Qiagen Inc., Valencia, CA, USA) and both strands were directly sequenced with the primers used for amplification. DNA sequencing was performed with the BigDye Terminator v3.1 Cycle Sequencing Kit and purified samples were run on an ABI Prism 3700 96-capillary automated sequencer (Applied BioSystems, Foster City, CA, USA).

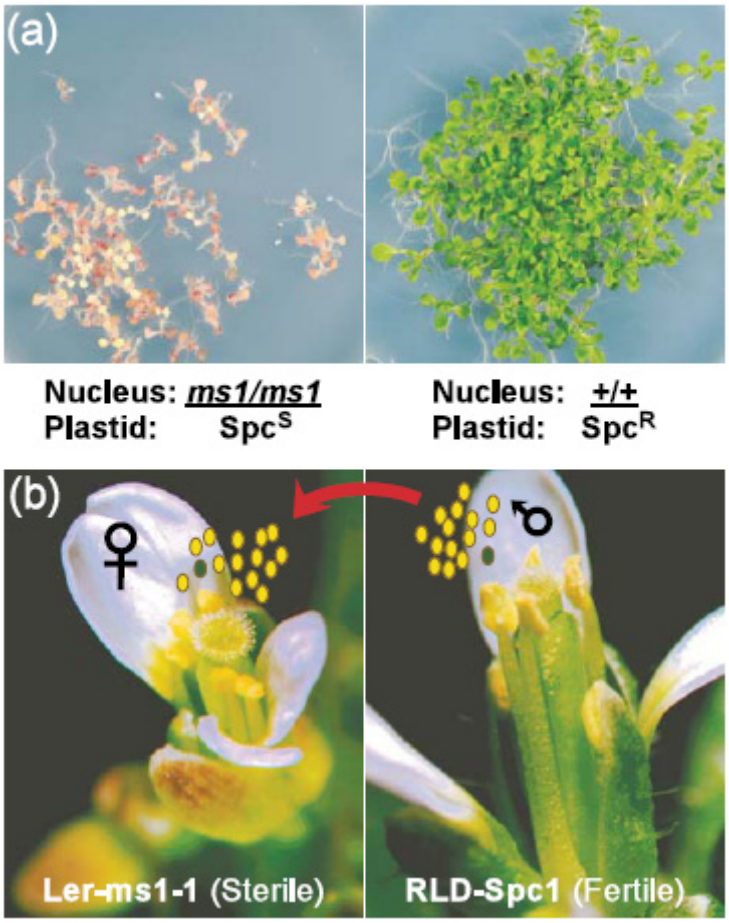


Figure 2.1. L*er ms1-1/ms1-1* maternal and RLD-Spc1 paternal parent.
(a) Maternal parent is spectinomycin sensitive (bleached), paternal parent is spectinomycin resistant (green) when germinated on ARM5 medium containing 100 mg/l spectinomycin HCl.(b) Flower morphology. Note lack of pollen on anthers of maternal and abundant pollen on the anthers of paternal parent

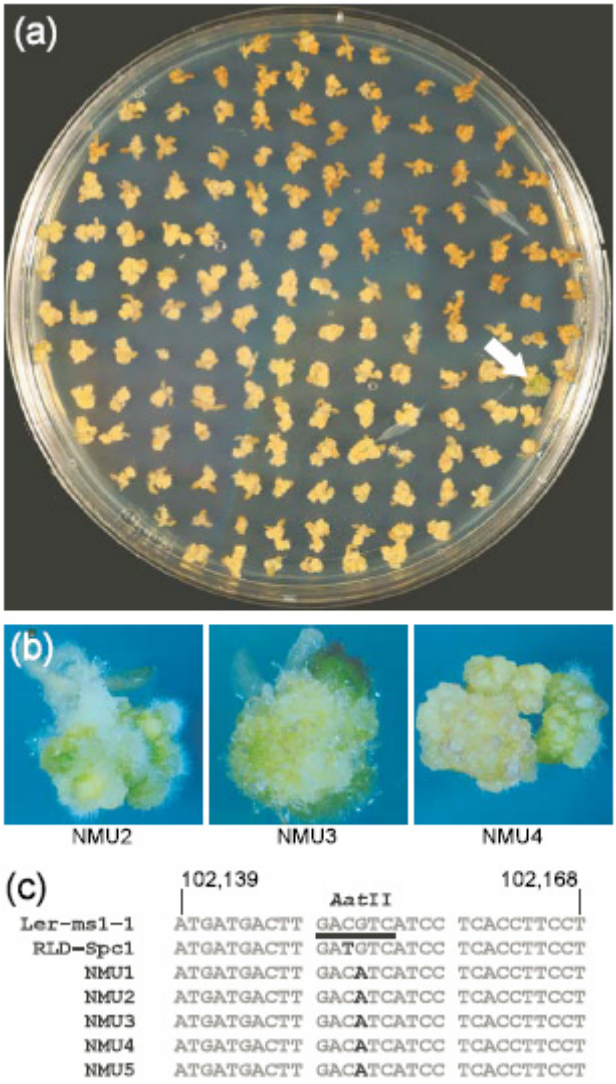


Figure 2.2. Isolation of spectinomycin resistant mutants in mutagenized (2.5 mM NMU) *Arabidopsis* seedling callus.
(a) Identification of spectinomycin resistant mutants on selective ARMI medium (100 mg/l spectinomycin HCl). White arrow points to spectinomycin resistant clone.
(b) Spectinomycin resistant clones with segregating green and white sectors.
(c) Mutations conferring spectinomycin resistance in the *rrn16* gene. The first nucleotide in the sequence corresponds to position 102,139 in the *Arabidopsis thaliana* ptDNA (GenBank Accession No. AP000423).

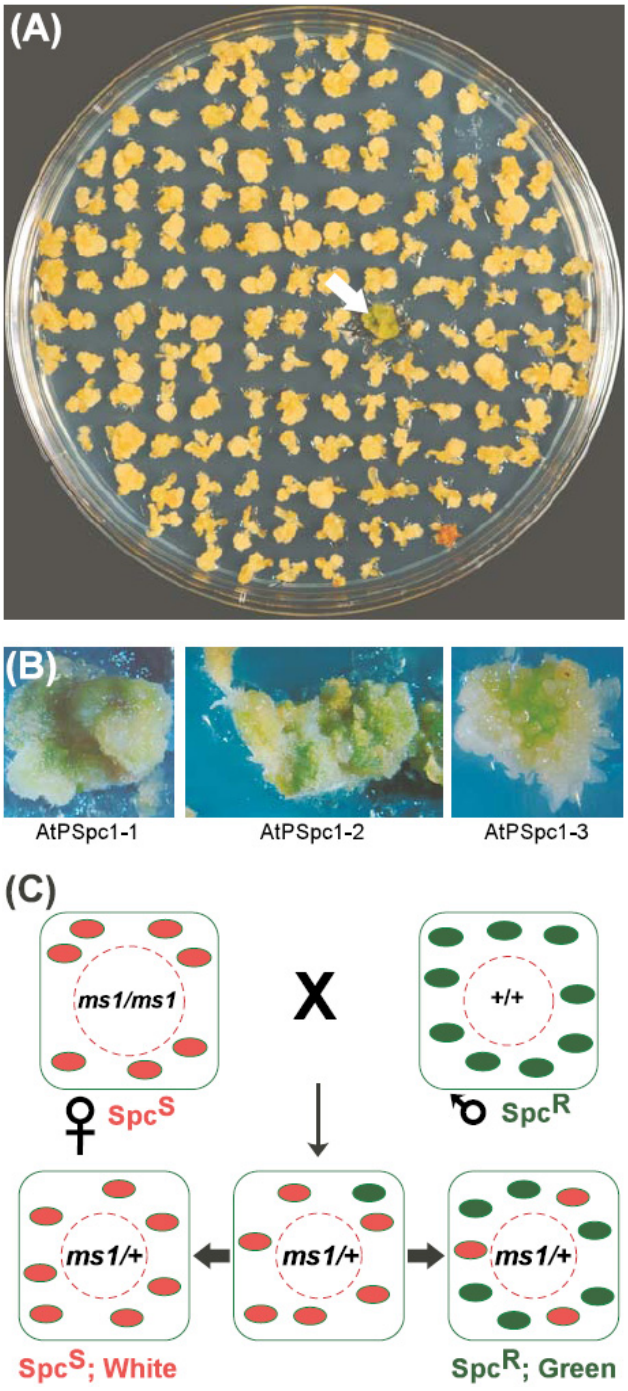


Figure 2.3. Identification of spectinomycin resistant lines in the seedling calli.
(a) Spectinomycin resistant clone appears after six weeks of cultivation on selective ARMI medium (100 mg/l spectinomycin HCl). Arrow points to green spectinomycin resistant clone.
(b) AtPT1–AtPT3 spectinomycin resistant clones with segregating green and white sectors.
(c) Plastid division and unequal division of the cytoplasm yields cells with sensitive (red organelles; on left) and mixed sensitive and resistant (red and green) plastids.

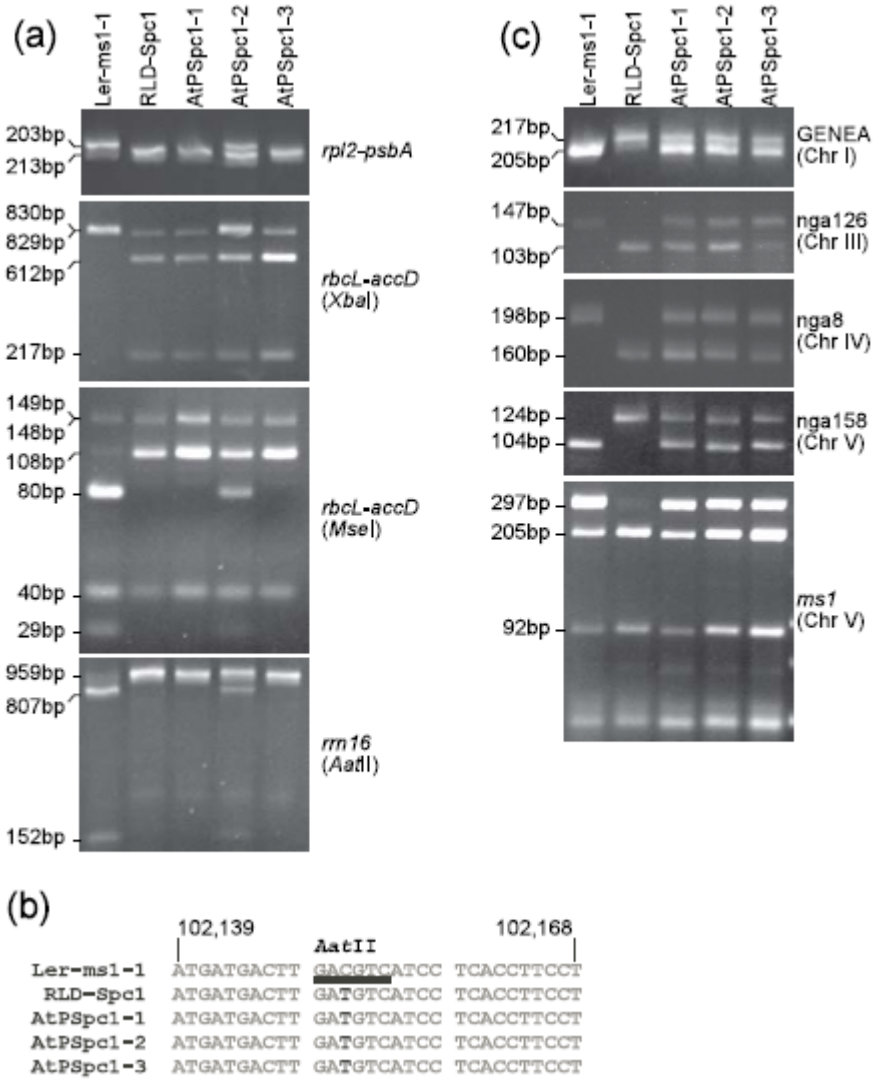


Figure 2.4. DNA markers confirm pollen transmission of ptDNA.
(a) Ecotype-specific plastid DNA markers. Data are shown for parental lines L*er-ms1-1* and RLD-Spc1, and spectinomycin resistant lines AtPT1-AtPT3. The markers are: *rpl2-psbA* AFLP; *rbcL-accD* RFLPs (*Xba*I, *Mse*I) and *rrn16* RFLP (*Aat*II).
(b) DNA sequence of the rrn16 region containing spectinomycin resistance mutations.
(c) Spectinomycin resistant lines are heterozygotes for AFLPs GENEA (Chr I), nga126 (Chr III), nga8 (Chr IV) and nga158 (Chr V) and the *ms1-1* RFLP (*Rsa*I) (Chr V).

Table 2.1. Primers to amplify plastid and nuclear markers

Marker	Product length (bp)	Restriction enzyme	Fragment length (bp)
Plastid marker: *rpl2-psbA* [a]			
L*er*	203		
RLD	213		
Primer 1: CCCTACGCTACTATCTATTC			
Primer 2: GAGCAAATTTTTATAGAGTATC			
Plastid marker: *rbcL-accD* [a]			
L*er*	830	*Xba*I	830
RLD	829	*Xba*I	217, 612
Primer 1: TGAACTAGCTGCTGCTTGTG			
Primer 2: GCTTTACTTAGCTCACCTCTG			
Plastid marker: *rbcL-accD* [a]			
L*er*	149	*Mse*I	29, 40, 80
RLD	148	*Mse*I	40, 108
Primer 1: CATAAGCGAAGAACAAATATTTC			
Primer 2: GAGTAAGCAAAACATATCGATGC			
Plastid marker: *rrn16*			
L*er*	959	*Aat*II	807, 152
RLD	959	*Aat*II	959
Primer 1: GGAGGAGTACGCCATGCTAATGTG			
Primer 2: GCGGACACAGGTGGTGCATGGCTG			
Nuclear marker: GENEA (Chr I) [b]			
L*er*	205		
RLD	217		
Primer 1: ACATAACCACAAATAGGGGTGC			
Primer 2: ACCATGCATAGCTTAAACTTCTTG			
Nuclear marker: nga126 (Chr III) [b]			
L*er*	147		
RLD	103		
Primer 1: CAAGAGCAATATCAAGAGCAGC			
Primer 2: GAAAAAACGCTACTTTCGTGG			
Nuclear marker: nga8 (Chr IV) [b]			
L*er*	198		
RLD	160		
Primer 1: TGGCTTTCGTTTATAAACATCC			
Primer 2: GAGGGCAAATCTTTATTTCGG			
Nuclear marker: nga158 (Chr V) [b]			
L*er*	104		
RLD	124		
Primer 1: ACCTGAACCATCCTCCGTC			
Primer 2: TCATTTTGGCCGACTTAGC			
Nuclear marker: ms1-1 (Chr V) [c]			
L*er*	279	*Rsa*I	297
RLD	279	*Rsa*I	205,92
Primer 1: CTCTCTCAACGGCATTGAAAC			
Primer 2: CAGCCTCAACTCCATTCCTT			

[a](Azhagiri and Maliga, 2007); [b]www.arabidopsis.org; [c] (Wilson et al., 2001)

References

Azhagiri, A. and Maliga, P. (2007) DNA markers define plastid haplotypes in *Arabidopsis thaliana. Curr Genet* **in press**.

Clough, S.J. and Bent, A.F. (1998) Floral dip: a simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. *Plant J* **16,** 735-743.

Derepas, A. and Dulieu, H. (1992) Inheritance of the capacity to transfer plastids by the pollen parent in *Petunia hybrida* Hort. *J Hered* **83,** 6-10.

Hagemann, R. (1982) Induction of plastome mutations by nitroso-urea-compounds. In *Methods in chloroplast molecular biology* (Edelman, M., Hallick, R.B. and Chua, N.H., eds). Amsterdam: Elsevier Biomedical Press, pp. 119-127.

Havey, M.J., McCraight, J.D., Rhodes, B. and Taurick, G. (1998) Differential transmission of the *Cucumis* organellar genomes. *Theor Appl Genet* **97,** 122-128.

Martinez, P., Lopez, C., Roldan, M., Sabater, B. and Martin, M. (1997) Plastid DNA of five ecotypes of *Arabidopsis thaliana*: sequence of *ndhG* gene and maternal inheritance. *Plant Sci* **123,** 113.

Marton, L. and Browse, J. (1991) Facile transformation of Arabidopsis. *Plant Cell Rep* **10,** 235-239.

Moll, B., Polsby, L. and Maliga, P. (1990) Streptomycin and lincomycin resistances are selective plastid markers in cultured *Nicotiana* cells. *Mol Gen Genet* **221,** 245-250.

Murray, M.G. and Thompson, W.F. (1980) Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res* **8,** 4321-4325.

Nagata, N., Saito, C., Sakai, A., Kuroiwa, H. and Kuroiwa, T. (1999) The selective increase or decrease of organellar DNA in generative cells just after pollen mitosis one controls cytoplasmic inheritance. *Planta* **209,** 53-65.

Poethig, S. (1989) Genetic mosaics and cell lineage analysis in plants. *Trends Genet* **5,** 273-277.

Ruf, S., Karcher, D. and Bock, R. (2007) Determining the transgene containment level provided by chloroplast transformation. *Proc Natl Acad Sci USA* **104,** 6998-7002.

Sikdar, S.R., Serino, G., Chaudhuri, S. and Maliga, P. (1998) Plastid transformation in *Arabidopsis thaliana. Plant Cell Rep* **18,** 20-24.

Svab, Z. and Maliga, P. (1991) Mutation proximal to the tRNA binding region of the *Nicotiana* plastid 16S rRNA confers resistance to spectinomycin. *Mol Gen Genet* **228,** 316-319.

Svab, Z. and Maliga, P. (2007) Exceptional transmission of plastids and mitochondria from the pollen parent and its impact on transgene containment. *Proc Natl Acad Sci U S A* **104,** 7003-7008.

van der Veen, J.H. and Wirtz, P. (1968) EMS-induced genic male sterility in *Arabidopsis thaliana*: A model selection experiment. *Euphytica* **17,** 371-377.

Wilson, Z.A., Morroll, S.M., Dawson, J., Swarup, R. and Tighe, P.J. (2001) The Arabidopsis MALE STERILITY1 (MS1) gene is a transcriptional regulator of male gametogenesis, with homology to the PHD-finger family of transcription factors. *Plant J* **28,** 27-39.

Yu, H.S. and Russell, S.D. (1994) Occurrence of mitochondria in the nuclei of tobacco sperm cells. *Plant Cell* **6,** 1477-1484.

Author Contributions to manuscript on which Chapter 2 is based

Arun Azhagiri and Prof. Pal Maliga designed research; AA performed research; AA and PM analysed data and wrote the paper.

CHAPTER 3

DNA polymorphisms to identify the ptDNA of *Arabidopsis thaliana* ecotypes

Arabidopsis thaliana is the most advanced plant model species with its nuclear (Arabidopsis Genome Initiative, 2000), plastid (Sato et al., 1999) and mitochondrial (Unseld et al., 1997) genomes sequenced. Genomic resources enable rapid progress in understanding inheritance of plastids, plastid function and interaction of plastid and nuclear genes (Nott et al., 2006). Although Arabidopsis would be the natural choice for studies in plastid genetics this is difficult due to the lack of useful genetic markers and problems with tissue culture. This study was conducted to identify ptDNA markers that would enable the identification of plastids from various Arabidopsis ecotypes. Towards this end I studied ptDNA variation in the *rpl2-psbA* and *rbcL-accD* intergenic regions in 26 accessions. The two regions contained eight polymorphic sites including five insertions and/or deletions (indels) involving changes in the length of A or T mononucleotide repeats and three base substitutions. The 27 alleles defined 15 plastid haplotypes, providing a practical set of ptDNA markers for the Landsberg *erecta*, Columbia and Wassilewskija ecotypes that are commonly used in genetic studies and also for the C24 and RLD ecotypes that are the most amenable for cell culture manipulations (Valvekens et al., 1988; Marton and Browse, 1991).

Results

We amplified a 680-bp fragment with primers complementary to the *rpl2* and *psbA* coding regions and sequenced the 548 nucleotides between the plastid *rpl2* and *psbA* coding regions from 26 *A. thaliana* ecotypes (Table 3.1). The sequence between the *rpl2* and *psbA* genes contains the *trnH* gene (Figure 3.1). Polymorphic sequences were found only in the *trnH-psbA* intergenic region. Shown in Figure 3.1 are DNA sequences of the three variant indel regions, differing with respect to the number of A or T mononucleotide repeats. The variant indel loci are identified by the position of the first A or T nucleotide of the repeat in the Col-0 ptDNA sequence; the alleles are distinguished by the number of mononucleotide repeats. Loci AtPt112, AtPt173 and AtPt289 contain five, eight and two alleles; respectively (Table 3.3). No polymorphisms in individuals of the same ecotype were detected in the 26 accessions.

Sequencing of the 678 bp *rbcL* and *accD* intergenic region revealed two indels consisting of A or T mononucleotide repeats and 3 nucleotide substitutions constituting a total of 12 alleles (Figure 3.2, Table 3.3). The indels at positions AtPt56758 and AtPt56839 have two and four alleles, respectively. The nucleotide substitutions create 2 alleles each at positions AtPt56533 (T or C), AtPt56730 (T or G) and AtPt56897 (A or G). The mutation at AtPt56533 creates an *Xba*I restriction site and a second mutation at AtPt56897 eliminates an *MseI* restriction site relative to the Col-0 ecotype. Indels and substitutions representing 27 alleles at eight polymorphic loci from the two intergenic regions were combined to calculate the pairwise genetic distance between the ecotypes using the GenAlEx program (data not shown). The analysis defined 15 haplotypes (Table 3.4) in the 26 ecotypes with each possessing a unique sequence. The ecotypes from Asia (Kondara, Hodja and Dijon-M) are significantly separated genetically from those from Southern Europe and Africa (Ag-0, Bl-1 and Ita-0). Ecotypes from Central and Northern Europe, such as Col-0, L*er*-0, Estland, S96 and Es-0 tend to have markers from both the Asian and Southern European accession, supporting proposed postglacial colonization of central and northern Europe from Asia and the Mediterranean region (Sharbel et al., 2000).

Particularly useful to distinguish Arabidopsis plastid haplotypes are point mutations that create or eliminate restriction sites. Examples for such RFLP markers are the *Xba*I and *Mse*I restriction sites present in 11 of the 15 haplotypes (AtPt56533, AtPt56897; this study) and a *Bam*HI site that can be used to

distinguish Cvi-0 from Col-0, L*er*-0, Le-0 and the St-0 ecotypes (Martinez et al., 1997) (Table 3.5). Comparison of ptDNA sequences in the study of Sall et al. (Sall et al., 2003) with the Col-0 ecotype identified loss or creation of restriction sites in six of the twelve studied regions. Presence or absence of these restriction sites was tested in six commonly used ecotypes and is listed in Table 3.5. Absence of a unique *Hin*fI site in the Col-0 ecotype enabled distinction of Col-0 from L*er*-0 ecotype in group F. Interestingly, variation common in the Scandinavian ecotypes (Sall et al., 2003) was absent in the ecotypes we studied (data not shown).

Discussion

We identified eight polymorphic loci in 26 *Arabidopsis thaliana* ecotypes including five mononucleotide A or T repeats and three nucleotide substitutions that represent a significant degree of ptDNA variability in a 1.2-kb region. Readily accessible information for the use of these polymorphic loci is listed in Table 3.6. Polymorphisms were found at loci with at least ten A or T nucleotides at a locus (Table 3.3). Polymorphic plastid indel loci, although termed "hypervariable" because they arise and persist in populations of the same species (Golenberg et al., 1993; Powell et al., 1995a; Provan et al., 2001), are relatively stable. Important in this regard is that in this study no variability was detected between seedlings from the same accession. The degree of ptDNA variability observed in this study is comparable to that detected in 25 Scandinavian Arabidopsis accessions (Sall et al., 2003) where, in a 4.2-kb sequence, 12 substitutions and eight indels were identified. None of the reported polymorphisms overlap with the loci described here. Thus, there is sufficient, readily identifiable polymorphism in the Arabidopsis ptDNA to distinguish each accession by sequencing additional loci.

It appears that the degree of ptDNA variation is dependent on the taxonomic group. A high degree of intra-species ptDNA variability was found using microsatellite markers in pines (Powell et al., 1995a) and soybean (Powell et al., 1995b). A high degree of ptDNA diversity was also found when comparing the entire 134.5-kb plastid genome sequences of the *indica* and *japonica* subspecies of rice (Tang et al., 2004). In this study 72 single nucleotide polymorphisms and 27 insertions and deletions have been identified. In contrast, there is a relative lack of polymorphisms in the Solanaceae family, in which no nucleotide difference was found upon sequencing the 155.5-kb plastid genome of two tomato cultivars (Kahlau et al., 2006). Also, comparison of the 155.9 kb plastid genomes of *Nicotiana tabacum* and its maternal progenitor, *Nicotiana sylvestris*, revealed only seven polymorphic sites, four nucleotide substitutions and three indels (Yukawa et al., 2006). Variability in ptDNA could experimentally be detected in ~ 60 % of Oenothera lines carrying a nuclear gene inducing plastid mutations (plastome mutator), a frequency that is 200- to 1000-fold higher than natural variation (Sears and Sokalski, 1991; Stoike and Sears, 1998).

Study of rare paternal pollen transmission (Medgyesy et al., 1986; Avni and Edelman, 1991) and experimental approaches to plastid genome evolution, (Stoike and Sears, 1998; Huang et al., 2003;

Stegemann et al., 2003; Timmis et al., 2004) thus far have been limited to species in the genus Nicotiana and Oenothera. Identification of intraspecies ptDNA markers reported here will facilitate the extension of these studies to *Arabidopsis thaliana.*

Experimental procedures

Plant material and growth conditions

The 26 Arabidopsis ecotypes (Table 3.1) were obtained from the Arabidopsis Biological Resource Center, Ohio State University, Columbus, Ohio. The seedlings were germinated under sterile conditions and grown as described by Valvekens et al. (Valvekens et al., 1988). Briefly, the seeds were surface sterilized, plated on germination medium (GM), vernalized for 7 days at 4 °C, transferred to 22 °C and grown at 22 °C under 16-hr light/8-hr dark cycle.

Amplification and sequencing of plastid intergenic regions

We amplified and sequenced the *rpl*2-*psbA* and *rbcL-accD* intergenic regions with primers listed in Table 3.2. The primers were designed based on the published Col-0 plastid genome sequence (Accession No. AP000423). Genomic DNA was extracted from leaf tissue according to the CTAB protocol (Murray and Thompson, 1980) from three seedlings of the 26 Arabidopsis ecotypes. Target DNA of each seedling (14days old) was amplified in three independent samples and sequenced. Amplification from 50 ng total DNA was carried out in a 25µl volume with Platinum ® *Pfx* DNA Polymerase (Invitrogen, Carlsbad, CA, USA) using touchdown PCR (Don et al., 1991) with the following program: 5 min 94°C, 3 cycles of 30 s 94°C, 30 s 68°C, 45 s 72°C, 3 cycles of 30 s 94°C, 30 s 64°C, 45 s 72°C, 3 cycles of 30 s 94°C, 30 s 60°C, 45 s 72°C, 20 cycles of 30 s 94°C, 30 s 58°C, 45 s 72°C and 1 cycle of 7 min 72°C. Amplified DNA fragments were purified using QIAguick Gel Extraction Kit (Qiagen Inc., Valencia, CA, USA) and both strands were directly sequenced with the primers used for amplification. DNA sequencing was performed with the BigDye Terminator v3.1 Cycle Sequencing Kit and purified samples were run on an ABI Prism 3700 96-capillary automated sequencer (Applied BioSystems, Foster City, CA, USA). DNA sequences have been deposited in GenBank under accession numbers EF032559-EF032584 and EF032533-EF032558.

Data analysis

Sequences for each region were edited using Chromas Lite 2.0 (http://www.technelysium.com.au/chromas.html) and aligned using BESTFIT (GCG Wisconsin Package). Multiple sequence alignments were made using PILEUP (GCG Wisconsin Package). Sequences were trimmed to remove the terminal coding regions in all analyses. The sequences were aligned to the *rpl*2-*psbA* and *rbcL*-*accD* sequences of the Col-0 ptDNA (AP000423). Gaps in the multiple sequence alignments were positioned to minimize the number of nucleotide differences. Phylogenetic analysis was performed using the GenAlEx software (Peakall and Smouse, 2006) to calculate the pairwise genetic distance between the ecotypes and to define the haplotypes. Both SNPs and insertion/deletions were included in the analysis.

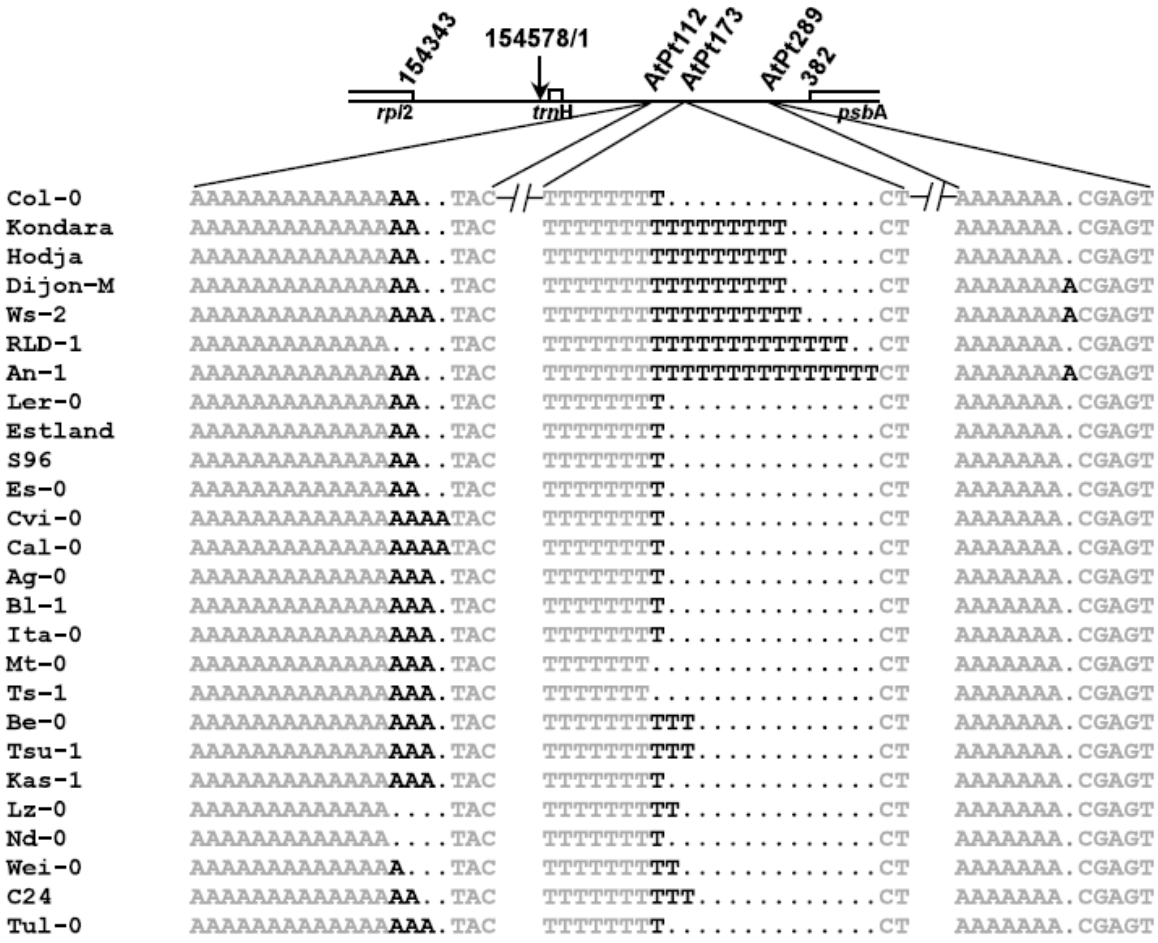


Figure 3.1. Schematic map and multiple sequence alignment of the Arabidopsis plastid *rpl2-psba* region. The positions of the polymorphic loci and gene names are shown on the map. Sequences in Figure 1 have been deposited in GenBank under accession numbers EF032559-EF032584.

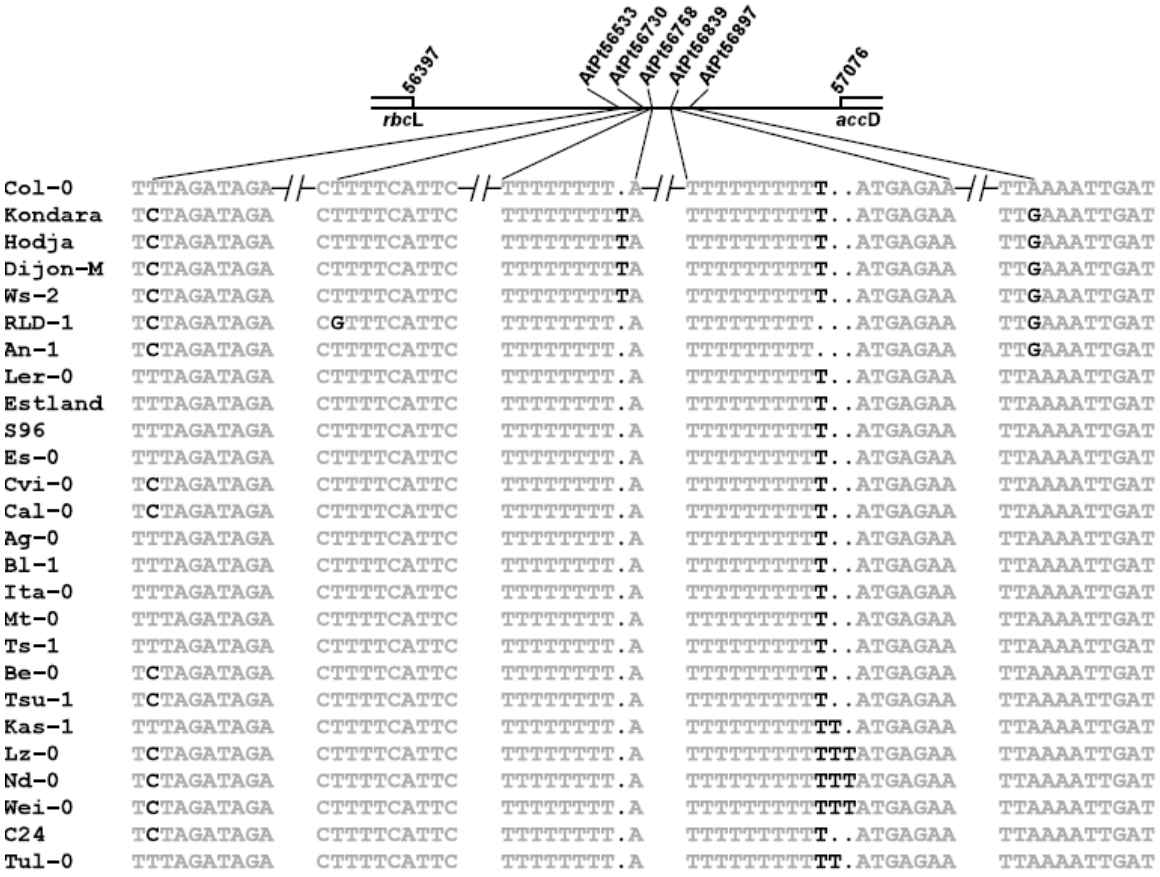


Figure 3.2. Schematic and multiple sequence alignment of the Arabidopsis plastid *rbcL-accD* intergenic region. The positions of the polymorphic loci and gene names are shown on the map. Sequences in Figure 3.2 have been deposited in GenBank under accession numbers EF032533-EF032558.

Table 3.1. List of *Arabidopsis thaliana* ecotypes investigated in this study, their commonly used abbreviations and seed stock numbers.

No.	ABRC Stock No.	Ecotype
1	CS20	L*er*-0
2	CS902	Cvi-0
3	CS911	Estland
4	CS913	RLD-1
5	CS914	S96
6	CS916	Kondara
7	CS919	Dijon-M
8	CS922	Hodja
9	CS937	Ag-0
10	CS944	An-1
11	CS964	Be-0
12	CS969	Bl-1
13	CS1062	Cal-0
14	CS1145	Es-0
15	CS1245	Ita-0
16	CS1265	Kas-1
17	CS1354	Lz-0
18	CS1381	Mt-0
19	CS1390	Nd-0
20	CS1552	Ts-1
21	CS1640	Tsu-1
22	CS2360	Ws-2
23	CS3110	Wei-0
24	CS1092	Col-0
25	CS1570	Tul-0
26	CS906	C24

Table 3.2. PCR primers for the amplification of Arabidopsis ptDNA[a]

Location	Primers (5'-3')	Position in AtPtDNA	Length of Product
rpl2-psbA	f: gagactttgattcttcgtcgc	154279	680 bp
	r: aaccgtgctaaccttggtatg	481	
rbcL-accD	f: tgaactagctgctgcttgtg	56315	830 bp
	r: gctttacttagctcacctctg	57145	

[a]The position of the primers in the published Columbia plastid DNA sequence (AP000423) and the length of the products are given. Abbreviations: f, forward primer; r, reverse primer.

Table 3.3. Polymorphisms in the Arabidopsis plastid *trnH-psbA* and *rbcL-accD* intergenic regions[a]

Locus	Intergenic region	Mutation type	No. of Alleles	Alleles
AtPt112	*trnH-psbA*	SSR	5	At112-A_{13}
				At112-A_{14}
				At112-A_{15}
				At112-A_{16}
				At112-A_{17}
AtPt173	*trnH-psbA*	SSR	8	At173-T_7
				At173-T_8
				At173-T_9
				At173-T_{10}
				At173-T_{16}
				At173-T_{17}
				At173-T_{20}
				At173-T_{22}
AtPt289	*trnH-psbA*	SSR	2	At289-A_7
				At289-A_8
AtPt56533	*rbcL-accD*	SNP	2	At56533-T
				At56533-C
AtPt56730	*rbcL-accD*	SNP	2	At56730-T
				At56730-G
AtPt56758	*rbcL-accD*	SSR	2	At56758-T_9
				At56758-T_{10}
AtPt56839	*rbcL-accD*	SSR	4	At56839-T_9
				At56839-T_{10}
				At56839-T_{11}
				At56839-T_{12}
AtPt56897	rbcL-accD	SNP	2	At56897-A
				At56897-G

[a]The variant simple sequence repeat (SSR) loci are identified by the position of the first A or T nucleotide of the repeat region in the Col-0 ptDNA sequence (AP000423) or the position of the single nucleotide polymorphism (SNP).

Table 3.4. Plastid haplotypes detected in the 26 *Arabidopsis thaliana* ecotypes defined by SSRs and SNPs at the eight variable loci

Haplotype	Ecotypes
A	Hodja, Kondara
B	Dijon-M
C	Ws-2
D	RLD-1
E	An-1
F*	Col-0, L*er*-0, Estland, Es-0, S96
G	Ag-0, Bl-1, Ita-0,
H	Kas-1, Tul-0
I	Mt-0, Ts-1,
J	Cvi-0, Cal-0
K	Be-0, Tsu-1
L	C24
M	Lz-0,
N	Wei-0
O	Nd-0

*Col-0 and L*er*-0 can be distinguished by marker AtPt69343, a *Hin*fI site (Table 3.5). Since the other ecotypes were not tested with *Hin*fI, the group is not subdivided.

Table 3.5. RFLP markers in the plastid genome of six commonly used Arabidopsis ecotypes[a]

Loci	*rpl20-clpP*		*rbcL-accD*		*ndhE-ndhI*
Ecotype	*Hin*fI (69346)	*Apo*I (69368)	*Xba*I (56532)	*Mse*I (56895)	*Bam*HI (118857)
Col-0	-	+	-	-	+
L*er*-0	+	+	-	-	+
Ws-2	+	+	+	+	+
RLD-1	+	+	+	+	+
C24	+	+	+	-	+
Cvi-0	+	+	+	-	-

[a]Nucleotide position of restriction sites is in brackets.

Table 3.6. Arabidopsis ptDNA markers

Genetic marker: Atpt173 (Simple Sequence Length Polymorphism or SSLP)

Species variant name	Product
Nd-0	201
Lz-0	202
Col-0; Estland; Es-0; L*er*-0; Mt-0; S96; Ts-1; Wei-0	203
Ag-0; Bl-1; Ita-0; Kas-1; Tul-0	204
C24; Cal-0; Cvi-0	205
Be-0; Tsu-1	206
Dijon-M; Hodja; Kondara	211
RLD-1; Ws-2	213
An-1	217

Primer 1: CCCTACGCTACTATCTATTC
Primer 2: GAGCAAATTTTTATAGAGTATC

Genetic marker: Atpt56533 (Cleaved Amplified Polymorphic Sequence or CAPS)

Species variant name	Product length	Restriction enzyme	Number of sites	Fragment length (bp)
An-1; RLD-1	829	*Xba*I	1	217, 612
Be-0; C24; Cal-0; Cvi-0; Tsu-1	830	*Xba*I	1	217, 613
Ag-0; Bl-1; Col-0; Es-0; Estland; Ita-0; L*er*-0; Mt-0; Ts-1; S96	830			830
Dijon-M; Hodja; Kondara; Ws-2	831	*Xba*I	1	217, 614
Kas-1; Tul-0	831			831
Lz-0; Nd-0; Wei-0	832	*Xba*I	1	217, 615

Primer 1: TGAACTAGCTGCTGCTTGTG
Primer 2: GCTTTACTTAGCTCACCTCTG

Genetic marker: Atpt56897 (CAPS)

Species variant name	Product length	Restriction enzyme	Number of sites	Fragment length (bp)
An-1; RLD-1	148	*Mse*I	1	40, 108
Ag-0; Be-0; Bl-1; C24; Cal-0; Col-0; Cvi-0; Es-0; Estland; Ita-0; L*er*-0; Mt-0; S96; Ts-1; Tsu-1	149	*Mse*I	2	29, 40, 80
Dijon-M; Hodja; Kondara; Ws-2	149	*Mse*I	1	40, 109
Kas-1; Tul-0	150	*Mse*I	2	29, 40, 81
Lz-0; Nd-0; Wei-0	151	*Mse*I	2	29, 40, 82

Primer 1: CATAAGCGAAGAACAAATATTTC
Primer 2: GAGTAAGCAAAACATATCGATGC

Genetic marker: Atpt69346 (CAPS)

Species variant name	Product length	Restriction enzyme	Number of sites	Fragment length (bp)
Col-0;	393			393
Cvi-0; L*er*-0; Ws-2; RLD-1; C24;	393	*Hin*fI	1	41, 352

Primer 1: GTCATTTACCCTGTTAGTCCG
Primer 2: GAAATACAAGACAGCCAATCC

Genetic marker: Atpt118857 (CAPS)

Species variant name	Product length	Restriction enzyme	Number of sites	Fragment length (bp)
Col-0; L*er*-0; Ws-2; RLD-1; C24;	1679	*Bam*HI	1	969, 710
Cvi-0;	1679			1679

Primer 1: GCTTCAGCGGCTGCAATTG
Primer 2: GCTTGTGAAGTATGTGTTCG

References

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

Avni, A. and Edelman, M. (1991) Direct selection for paternal inheritance of chloroplasts in sexual progeny of *Nicotiana*. *Mol Gen Genet* **225,** 273-277.

Don, R.H., Cox, P.T., Wainwright, B.J., Baker, K. and Mattick, J.S. (1991) 'Touchdown' PCR to circumvent spurious priming during gene amplification. *Nucleic Acids Res* **19,** 4008.

Golenberg, E.M., Clegg, M.T., Durbin, M.L., Doebley, J. and Ma, D.P. (1993) Evolution of a noncoding region of the chloroplast genome. *Mol Phylogenet Evol* **2,** 52-64.

Huang, C.Y., Ayliffe, M.A. and Timmis, J.N. (2003) Direct measurement of the transfer rate of chloroplast DNA into the nucleus. *Nature* **422,** 72-76.

Kahlau, S., Aspinall, S., Gray, J.C. and Bock, R. (2006) Sequence of the tomato chloroplast DNA and evolutionary comparison of solanaceous plastid genomes. *J Mol Evol* **63,** 194-207.

Martinez, P., Lopez, C., Roldan, M., Sabater, B. and Martin, M. (1997) Plastid DNA of five ecotypes of *Arabidopsis thaliana*: sequence of *ndhG* gene and maternal inheritance. *Plant Sci* **123,** 113.

Medgyesy, P., Pay, A. and Marton, L. (1986) Transmission of paternal chloroplasts in *Nicotiana*. *Mol Gen Genet* **204,** 195-198.

Murray, M.G. and Thompson, W.F. (1980) Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res* **8,** 4321-4325.

Nott, A., Jung, H.S., Koussevitzky, S. and Chory, J. (2006) Plastid-to-nucleus retrograde signaling. *Annu Rev Plant Biol* **57,** 739-759.

Peakall, R. and Smouse, P.E. (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol Ecol Notes* **6,** 288-295.

Powell, W., Morgante, M., McDevitt, R., Vendramin, G.G. and Rafalski, J.A. (1995a) Polymorphic simple sequence repeat regions in chloroplast genomes: applications to the population genetics of pines. *Proc Natl Acad Sci U S A* **92,** 7759-7763.

Powell, W., Morgante, M., Andre, C., McNicol, J.W., Machray, G.C., Doyle, J.J., Tingey, S.V. and Rafalski, J.A. (1995b) Hypervariable microsatellites provide a general source of polymorphic DNA markers for the chloroplast genome. *Curr Biol* **5,** 1023-1029.

Provan, J., Powell, W. and Hollingsworth, P.M. (2001) Chloroplast microsatellites: new tools for studies in plant ecology and evolution. *Trends Ecol Evol* **16,** 142-147.

Sall, T., Jakobsson, M., Lind-Hallden, C. and Hallden, C. (2003) Chloroplast DNA indicates a single origin of the allotetraploid *Arabidopsis suecica*. *J Evol Biol* **16,** 1019-1029.

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

Sears, B.B. and Sokalski, M.B. (1991) The Oenothera plastome mutator: effect of UV irradiation and nitroso-methyl urea on mutation frequencies. *Mol Gen Genet* **229,** 245-252.

Sharbel, T.F., Haubold, B. and Mitchell-Olds, T. (2000) Genetic isolation by distance in *Arabidopsis thaliana*: biogeography and postglacial colonization of Europe. *Mol Ecol* **9,** 2109-2118.

Stegemann, S., Hartmann, S., Ruf, S. and Bock, R. (2003) High-frequency gene transfer from the chloroplast genome to the nucleus. *Proc Natl Acad Sci U S A* **100,** 8828-8833.

Stoike, L.L. and Sears, B.B. (1998) Plastome mutator-induced alterations arise in Oenothera chloroplast DNA through template slippage. *Genetics* **149,** 347-353.

Tang, J., Xia, H., Cao, M., Zhang, X., Zeng, W., Hu, S., Tong, W., Wang, J., Wang, J., Yu, J., Yang, H. and Zhu, L. (2004) A comparison of rice chloroplast genomes. *Plant Physiol* **135,** 412-420.

Timmis, J.N., Ayliffe, M.A., Huang, C.Y. and Martin, W. (2004) Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. *Nat Rev Genet* **5,** 123-135.

Unseld, M., Marienfeld, J.R., Brandt, P. and Brennicke, A. (1997) The mitochondrial genome of *Arabidopsis thaliana* contains 57 genes in 366,924 nucleotides. *Nature Genet* **15,** 57-61.

Valvekens, D., Van Montagu, M. and Van Lusebettens, M. (1988) *Agrobacterium tumefaciens*-mediated transformation of *Arabidopsis thaliana* root explants by using kanamycin selection. *Proc Natl Acad Sci U.S.A.* **85,** 5536-5540.

Yukawa, M., Tsudzuki, T. and Sugiura, M. (2006) The chloroplast genome of *Nicotiana sylvestris* and *Nicotiana tomentosiformis*: complete sequencing confirms that the *Nicotiana sylvestris* progenitor is the maternal genome donor of *Nicotiana tabacum*. *Mol Genet Genomics* **275,** 367-373.

Author Contributions to paper on which Chapter 3 is based

Arun Azhagiri and Prof. Pal Maliga designed research; AA performed research; AA and PM analysed data and wrote the paper.

CHAPTER 4

Isolation of fertile, plastid spectinomycin resistant mutant in *Arabidopsis thaliana* using steroid-inducible plant regeneration systems

I needed selectable plastid DNA markers in the Arabidopsis paternal pollen parent. Spectinomycin resistant mutants have been isolated in tobacco and *Solanum nigrum* (Fromm et al., 1987; Svab and Maliga, 1991; Kavanagh et al., 1994) and were shown to be useful to identify rare paternal transmission in tobacco. Therefore, I set out to isolate spectinomycin resistant mutants in *Arabidopsis thaliana*. In Arabidopsis, regeneration of fertile plants from tissue culture is a problem. Therefore, the selection of spectinomycin resistant mutants was carried out in the new plant regeneration system based on the BABY BOOM (BBM) transcription factor.

The BABY BOOM (BBM) transcription factor was translationally fused with the green fluorescent protein (GFP) and the glucocorticoid receptor (GR) steroid-binding domain. Plants carrying the fusion protein were indistinguishable from non-transformed plants in the absence of the inducer. However, when the synthetic steroid dexamethasone was included in the culture medium, prolific shoot regeneration on Arabidopsis leaf segments and formation of somatic embryos on seedlings was obtained after extended periods of tissue culture. The utility of the system was shown by recovery of fertile plants of the L*er*-SPC2 line carrying a plastid-encoded spectinomycin-resistant mutation in the 16S rRNA.

Results

Construction of Arabidopsis plants with an inducible BBM *gene*

To construct a steroid-inducible BBM, the *BBM* coding region was translationally fused with the coding region of *gfp*, the *Aequorea victoria* green fluorescent protein (GFP) and the ligand-binding domain of the rat glucocorticoid receptor (GR). The BBM:GFP:GR fusion protein was cloned into a Cauliflower Mosaic Virus 35S expression cassette in a derivative of the pPZP200 binary plasmid carrying gentamycin resistance as a plant marker gene to create plasmid pKO216 (Figure 4.1). *Arabidopsis thaliana* RLD and Landsberg *erecta* (L*er*) plants were transformed with Agrobacterium strain EHA101 carrying binary plasmid pKO216 by the floral dip protocol (Clough and Bent, 1998). Eight independent insertion events were identified by selection of seedlings on medium containing 25 mg l^{-1} gentamycin. The gentamycin resistant seedlings were transferred to pots where they developed into normal plants, flowered and produced seed. Lines were designated by the ecotype, the plasmid name and a serial number, for example L*er*-pKO216-2.

Induction of somatic embryos from seed in the absence of plant hormones

BBM was shown to induce formation of somatic embryos in the absence of plant hormones (Boutilier et al., 2002). To test if the BBM:GFP:GR construct is biologically active, we germinated seed on a hormone-free plant maintenance medium (ARM) containing DEX (5 μM) and gentamycin (50 mg l^{-1}). We have found that seedlings in six out of the seven independently transformed lines tested formed clusters of somatic embryos and cotyledons indicating that the fusion proteins are biologically active (Figure 4.2a). Interestingly, not all gentamycin resistant seedlings responded by forming embryos indicating that the transgene phenotype does not have 100% penetrance. We therefore decided to test if a rapid, uniform tissue culture response can be obtained in BBM:GFP:GR plants in the absence of 2,4-D. The experiments were carried out with non-segregating, gentamycin resistant progenies identified in two of the lines, L*er*-pKO216-2-31 (L*er*-216-2) and RLD-pKO216-22-2 (RLD-216-22).

BBM potentiates shoot regeneration from leaves

Capacity for plant regeneration in *Arabidopsis thaliana* is ecotype dependent and depends on explant type: ecotype RLD is easier to regenerate than Landsberg *erecta* (L*er*); roots regenerate plants faster than leaves (Valvekens et al., 1988; Marton and Browse, 1991). Cell division and efficient shoot regeneration from both leaves (Feldmann and Marks, 1986) and roots (Valvekens et al., 1988; Marton and Browse, 1991) was dependent on short-term cultivation of explants on a medium containing 2,4-D, a synthetic auxin. Since BBM is known to induce somatic embryogenesis in the absence of plant hormones (Boutilier et al., 2002), we cultured leaf sections with and without the inducer, 5 μM dexamethasone (DEX), on ARM I medium lacking 2,4-D (ARM-B medium) to test if BBM expression can replace the short-term 2,4-D treatment. In one month most leaf sections of wild-type L*er* turned brown and died on the ARM-B medium (ARM medium containing IAA, 2IP and BA, but not 2,4-D), whether or not the inducer was included in the culture medium (Figure 4.2b). In dramatic contrast, L*er*-pKO216-2 leaf sections produced prolific shoots and some callus in the presence of DEX. BBM expression is somewhat leaky, since some shoot regeneration is also seen in the absence of the inducer (-DEX, L*er*-216-2; Figure 4.2b). The BBM response in the RLD background is similar, but reflects an overall better potential for shoot regeneration from leaf sections. Thus, significantly enhanced, uniform shoot regeneration could be obtained in both ecotypes in the presence of the inducer indicating that BBM expression could replace induction by 2,4-D.

Identification of spectinomycin resistant mutants in embryogenic cultures

Selection of transplastomic clones by spectinomycin resistance is anticipated in one of two systems: in cultures initiated from seed or in embryogenic roots. We tested the utility of these systems by selection of spectinomycin resistant mutants using the steroid inducible *BBM:GFP:GR* gene. Spectinomycin resistant mutants in seedling-derived cultures were selected in mutagenised seed populations (2.5 mM NMU). Seeds were selected on agar-solidified ARM-B medium containing 5 μM DEX and spectinomycin (25 and 100 mg l^{-1}). The resistant mutants were identified as green sectors on ~ 0.5 % of the seedling calli (Table 4.1),

which were white due to inhibition of protein synthesis on the prokaryotic-type plastid ribosomes (Figure 4.3a).

Selection for spectinomycin resistance was also carried out in embryogenic Arabidopsis root cultures in a liquid ARM-B medium containing 5 μM DEX and spectinomycin (25 mg l^{-1} and 100 mg l^{-1}). One of the cultures contained a large green embryogenic callus (Figure 4.3b), which was transferred to DEX-free medium to facilitate shoot regeneration. The At-L*er*-SPC2 shoots were transferred to hormone-free ARM medium where they flowered and set seed in a Magenta box. The seed was sown on a hormone-free ARM medium in the presence of spectinomycin (100 mg l^{-1}) (Figure 4.3c). Spectinomycin resistant green L*er*-SPC2 seedlings confirmed inheritance of the plastid-encoded spectinomycin resistance mutation (see below).

Spectinomycin resistance in tobacco and *Solanum nigrum* was shown to be due to mutations in the plastid small 16S rRNA subunit (Fromm et al., 1987; Svab and Maliga, 1991; Kavanagh et al., 1994). We therefore PCR-amplified the *rrn16* gene (encoding 16S rRNA) from four mutants derived from the mutagenised seedlings and the L*er*-SPC2 spontaneous mutant. Direct sequencing of PCR-amplified DNA identified three point mutations. Two of the mutations are known to confer spectinomycin resistance in tobacco: C to T at nucleotide position 102,151 and G to A at position 102,152 in the *A. thaliana* ptDNA, and eliminate an *AatII* restriction site. A third mutation, a C to T exchange at nucleotide position 102,026, has been described in *E. coli* (Johanson and Hughes, 1995) but not yet in plastids. This mutant carries a second point mutation, a G to A exchange at position 102,217, which is not known to be involved in spectinomycin resistance. Spectinomycin resistance mutations identified in this study are listed in Figure 4.4. We maintain the L*er*-SPC2 spontaneous mutant by seed. No attempt was made to obtain plants from the mutagenised seedlings since NMU mutagenesis normally yields plants with fertility problems.

Discussion

We report here post-transcriptionally regulated plant regeneration in *Arabidopsis thaliana* based on a BBM:GFP:GR fusion protein. Plants expressing the *BBM:gfp:GR* gene from a constitutive promoter are normal in the absence of DEX, the GR ligand. BBM:GFP:GR biological activity in the presence of the inducer DEX was demonstrated in two systems: in seed culture and in leaf culture. Seeds carrying a *BBM:gfp:GR* gene formed somatic embryos in the absence of plant hormones as seen in seedlings constitutively expressing BBM (Boutilier et al., 2002). The response was not uniform as only some of the seeds formed embryo clusters indicating less than a 100% phenotypic penetrance of the *BBM:gfp:GR* gene. Interestingly, enhanced shoot regeneration on hormone-containing medium in BBM:GFP:GR leaves was uniform, enabling plant regeneration. Different Arabidopsis ecotypes regenerate plants at different efficiencies (Schmidt and Willmitzer, 1988; Luo and Koop, 1997). Thus, one use of the BBM system will be to boost morphogenic response in ecotypes which are recalcitrant to plant regeneration.

Experiments reported here were carried out with a protein obtained by fusing three proteins: BBM, GFP and GR. Although the protein contained GFP, no GFP fluorescence was detectable in the cells (data not shown). Plants expressing a BBM-GR fusion protein obtained by transformation with vector pKO214 behaved similarly: had no phenotype in the absence of the inducer and formed embryogenic clusters in the absence of hormones from seed (data not shown). While this manuscript was in preparation, a similar BBM:GR fusion protein was described, and shown to enhance plant regeneration in tobacco (Srinivasan et al., 2007).

BBM is not the only regulated plant regeneration system in Arabidopsis. The WUSCHEL gene, when expressed in a transcriptionally regulated system, promoted vegetative-to-embryonic transition and yielded fertile plants after the removal of the inducer (Zuo et al., 2002). LEC1 and LEC2 are two additional seed-expressed transcription factor genes which, when expressed constitutively, promoted spontaneous embryo formation on vegetative tissues (Lotan et al., 1998; Stone et al., 2001). Interestingly, regulated over-expression of LEC1 cited in ref. (Zuo et al., 2002) and induction of a LEC2:GR fusion (Santos Mendoza et al., 2005) did not result in formation of embryo-like structures.

When we transferred embryos from the embryogenic clusters to inducer-free medium, the embryos grew into fertile plants and set viable seed in Magenta boxes on ARM medium (data not shown). This system was useful to obtain fertile plants from the spontaneous spectinomycin resistant mutant L*er*-SPC2, which was grown in a selective spectinomycin medium for three months in the presence of the inducer. Since the Arabidopsis ecotypes have distinct plastid DNA markers (Azhagiri and Maliga, 2007), this mutant represents a valuable resource that enables tracking of plastids in crosses of different ecotypes. The new steroid-inducible BBM system will facilitate the recovery of fertile transplastomic plants.

Selection of spectinomycin mutants using the estradiol inducible WUSCHEL plant regeneration system

In Arabidopsis, several genes that promote adventitious shoot regeneration in culture have been identified. These include, the *WUSCHEL* gene which was shown to promote vegetative-to-embryonic transition in an inducible manner and yielded fertile plants after the removal of the inducer (Zuo et al., 2002). Two other genes are the LEC1 and LEC2 transcription factors that are expressed in seeds which, when expressed constitutively, promoted spontaneous embryo formation on vegetative tissues (Lotan et al., 1998; Stone et al., 2001).

Transgenic PGA6-1 (Plant Growth Activator 6) plants carrying an estradiol-inducible XVE-*WUSCHEL* (*WUS*) transgene formed highly embryogenic callus when induced in the presence of auxin (IAA, 0.15 mgl^{-1}). The components of this system are: a chimeric transcription activator, XVE, which is expressed constitutively but not imported into the nucleus unless the inducer 17-β-estradiol is present in the cytoplasm. Estradiol-activated XVE stimulates transcription from a chimeric promoter consisting of eight copies of the LexA operator fused with the -46 35S minimal promoter upstream of *WUS*. The somatic embryos, induced in culture, were able to germinate directly without any further treatment and develop into fertile adult plants (Zuo et al., 2002). However, these conditions were not conducive to maintaining the cultures for extended periods in tissue culture required for plastid transformation. The utility of the system for plant regeneration was tested by selecting spectinomycin resistant mutants (Table 4.2). The seeds were treated with the chemical mutagen, N-nitroso-N-methylurea (NMU) and germinated the seedlings on ARMI inductive medium supplemented with 100 mgl^{-1} spectinomycin. Spectinomycin resistant clones

were identified as green of sectors in seed-derived calli (Figure 4.5). PCR amplification of the *16Srrn* gene region from seven resistant mutants, followed by restriction digestion with *Aat*II confirmed that the mutation was in the *Aat*II site shown to confer spectinomycin resistance (Svab and Maliga, 1991). Direct sequencing of the PCR products identified the mutations as G to A at position 102,152 in the *A. thaliana* ptDNA These resistant calli were maintained on ARMI tissue culture medium and shoot formation was induced in ARM medium (Marton and Browse, 1991). Although, shoots were regenerated from these calli, none of them were fertile and further experiments were discontinued.

Experimental procedures

Construction of BBM:GFP:GR plants

To construct a steroid-inducible BBM, the *BBM* coding region was translationally fused with the coding region of *gfp*, the *Aequorea victoria* green fluorescent protein, and the ligand-binding domain of the rat glucocorticoid receptor (GR). The fusion protein is contained in a *Bam*HI-*Xba*I fragment. The BBM coding region was PCR amplified from cDNA (Genbank Accession Number AF317907). The *gfp* gene is a derivative of plasmid *psmGFP*, encoding the soluble modified version of GFP (GenBank Accession No. U70495) obtained under order number CD3-326 from the Arabidopsis Biological Resource Center, Columbus, OH (Davis and Vierstra, 1998). The ligand-binding domain of the rat glucocorticoid receptor (508-795) derives from plasmid pBI-ΔGR constructed by Alan Lloyd (Lloyd et al., 1994). Plasmid pKO216 is a pPZP200 binary plasmid derivative (Hajdukiewicz et al., 1994) in which the plant gentamycin resistance marker is expressed in a 2' promoter and *nos* terminator cassette, and the BBM:GFP:GR coding region is cloned as a *Bam*HI-*Xba*I fragment in the Cauliflower Mosaic Virus 35S promoter-terminator cassette derived from plasmid pFF19G (Timmermans et al., 1990). Plasmid pKO216 was introduced into the RLD and L*er* ecotypes of *Arabidopsis thaliana* by the floral dip protocol (Clough and Bent, 1998). Eight independent insertion events were identified by selecting seedlings on medium containing 25 mg l^{-1} gentamycin, a relatively low concentration because in this vector variant the *aacC1* gene is poorly expressed.

Tissue culture media

The tissue culture protocols were adopted from (Marton and Browse, 1991; Czako et al., 1993). The *Arabidopsis* tissue culture media (ARM) are derivatives of the Murashige & Skoog MS medium (Murashige and Skoog, 1962). ARM medium: MS salts, 3% sucrose, 0.8% agar (A7921; Sigma, St. Louis, MO), 200 mg myo-inositol, 0.1 mg l^{-1} biotin (1 ml l^{-1} of 0.1 mg m l^{-1} stock), 1 ml vitamin solution (10 mg vitamin B1, 1 mg vitamin B6, 1 mg nicotinic acid, 1 mg glycine per ml), pH 5.8. ARM-B medium: ARM

medium containing 3 mg indolacetic acid (IAA), 0.6 mg benzyladenine (BA) and 0.3 mg isopentenyladenine (IPA) per liter. The *Arabidopsis* shoots were rooted on ARM medium. *Arabidopsis* seed culture (ARM5) medium: ARM medium supplemented with 5% sucrose (Sikdar et al., 1998). The stocks of plant hormones were filter sterilized, and added to media cooled to 45 ^{o}C after autoclaving. Dexamethasone (DEX) (D4902) purchased from Sigma, St. Louis, MO, was used at 5 μM final concentration in the culture medium. The 10 mM stock solution was prepared by dissolving 4 mg DEX in 1 ml DMSO and filter sterilized using DMSO Safe Acrodisc Syringe Filter (Pall Corporation, Ann Arbor, MI).

Selection of spectinomycin resistant mutants

Vapor sterilize seed in a 1.5 ml eppendorf tube overnight in a desiccator with a flask containing 100 ml bleach and 3 ml of concentrated (37.8%) HCl (Clough and Bent, 1998). Flasks should be placed in the desiccator immediately after bleach and HCl are mixed. Sprinkle seed onto ARM plates, cold treat seed by incubating plates in cold (4 ^{o}C) for three days, then transfer to culture room (16 hrs light cycle; 25 ^{o}C).

Mutagenesis of sterile seed was carried out by incubating the seed for 6 hrs at 25 ^{o}C in 2.5 mM nitrosomethyl urea (NMU) (Sigma, St. Louis, MO) dissolved in citric acid –Na_2PO_4 buffer, pH 5.0 (Hagemann, 1982). The seeds were rinsed in sterile distilled H_2O, and plated on selective (25 mg l^{-1} spectinomycin HCl) ARM I (L*er*-216-2) or ARM-B (L*er*; RLD, RLD-216-22) medium containing 5 μM DEX, cold treated, then germinated. The seedlings were then transferred twice (for two weeks each) onto a selective ARM-B medium containing spectinomycin and 5 μM DEX. During the 1st and 2nd selection cycles the medium contained 25 mg l^{-1} and 100 mg l^{-1} spectinomycin HCl, respectively. Spectinomycin resistant mutants were identified during the second selection cycle.

The spontaneous mutant L*er*-SPC2 was identified in L*er*-pKO216-2 root culture. Sterilized seed were sprinkled onto SCM plates (Zuo et al., 2002) supplemented with 1 ml vitamin stock solution, 2 % glucose, 1 mg l^{-1} IAA, 0.3 mg l^{-1} IPA and 5 μM Dex. One week after germination the hypocotyls with the roots were transferred into liquid ARM-B medium containing 5 μM DEX. After ~1.5 months of incubation on a gyratory shaker (90 rpm) in the light (16 hrs cycle) roots were transferred to a fresh ARMB medium

containing spectinomycin (25 mg l^{-1}) and 5 μM DEX. After two weeks the medium was replaced with a new batch of the same selective medium; during the 3rd selection cycle the spectinomycin concentration was increased to 100 mg l^{-1}. The spontaneous L*er*-SPC2 mutant was identified three months after starting the spectinomycin selection.

Sequencing of rrn16 *genes*

A 958 bp fragment of the plastid *rrn16* gene was PCR amplified using primers Lf1 5'-GGAGGAGTACGCCATGCTAATGTG-3' and Lf2 5'-CGCCTGTGTCCACCACGTACCGAC-3' with the following PCR program: 3 min at 94°C, followed by 30 cycles of 30 seconds at 94°C, 1 minute at 55°C and 1 min at 72°C; and 1 cycle of 11 minutes at 72°C. The PCR fragment was excised from an agarose gel and the DNA was isolated with the QIAquick Gel Extraction Kit (Qiagen, Valencia, CA). Direct sequencing of the PCR fragment was performed by Genewiz (South Plainfield, NJ) using primer Lf3 5'-GCAGCCTGCAATCCGAACTG-3'.

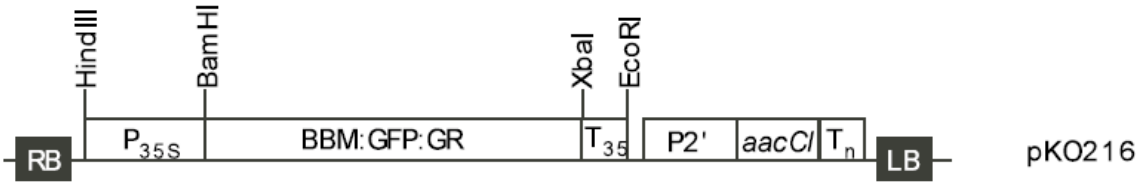


Figure 4.1. The map of the Agrobacterium binary vector pKO216 T-DNA region. Shown are the BBM:GFP:GR coding region expressed in a 35S promoter-terminator (P_{35S}/T_{35S}) cassette, the *aacC1* gentamycin resistance gene expressed in the P_2/T_{nos} cassette, and the T-DNA left (LB) and right (RB) border regions.

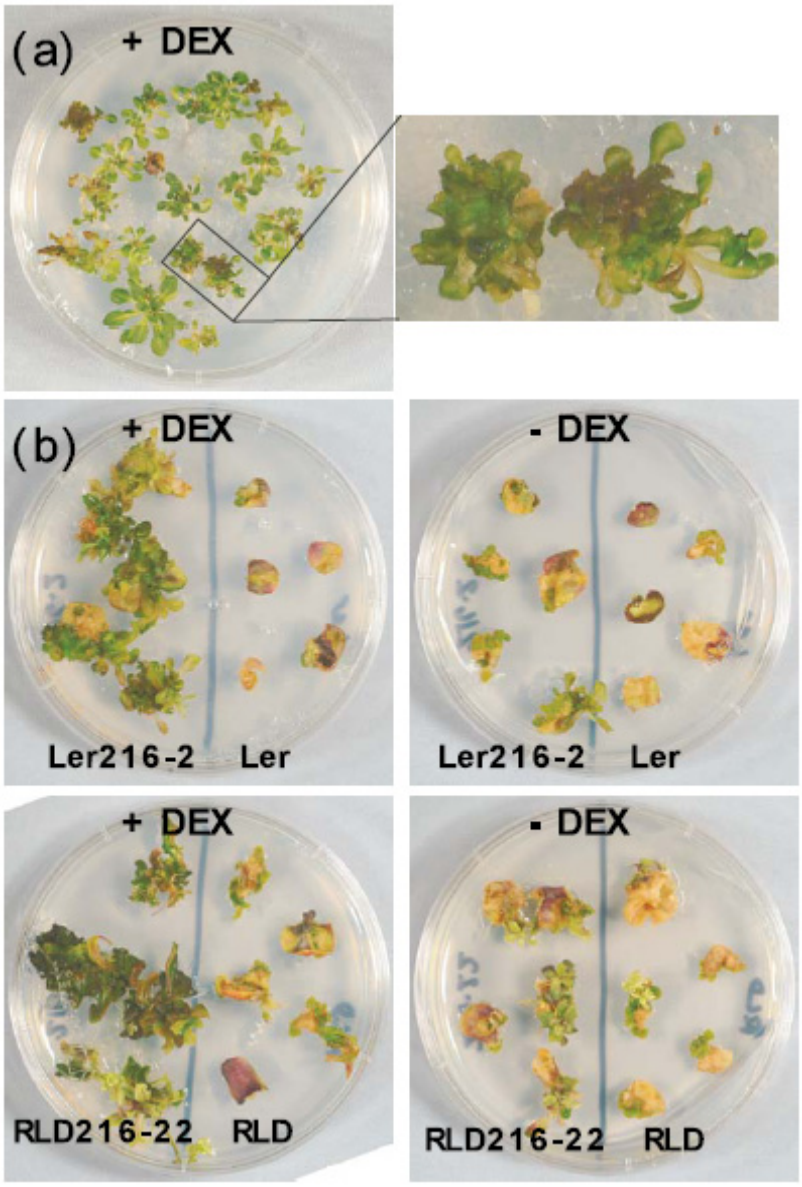


Figure 4.2. Dexamethasone-inducible plant regeneration in *A. thaliana* transformed with the BBM:GFP:GR gene in plasmid pKO216.
(a) Formation of embryogenic clusters from L*er*-pKO216-2 seedlings germinated on hormone-free ARM medium containing dexamethasone.
(b) DEX (5 μM) induces shoot regeneration from BBM:GFP:GR-transformed L*er*-216-2 and RLD-216-22 Arabidopsis leaves on ARM-B medium.

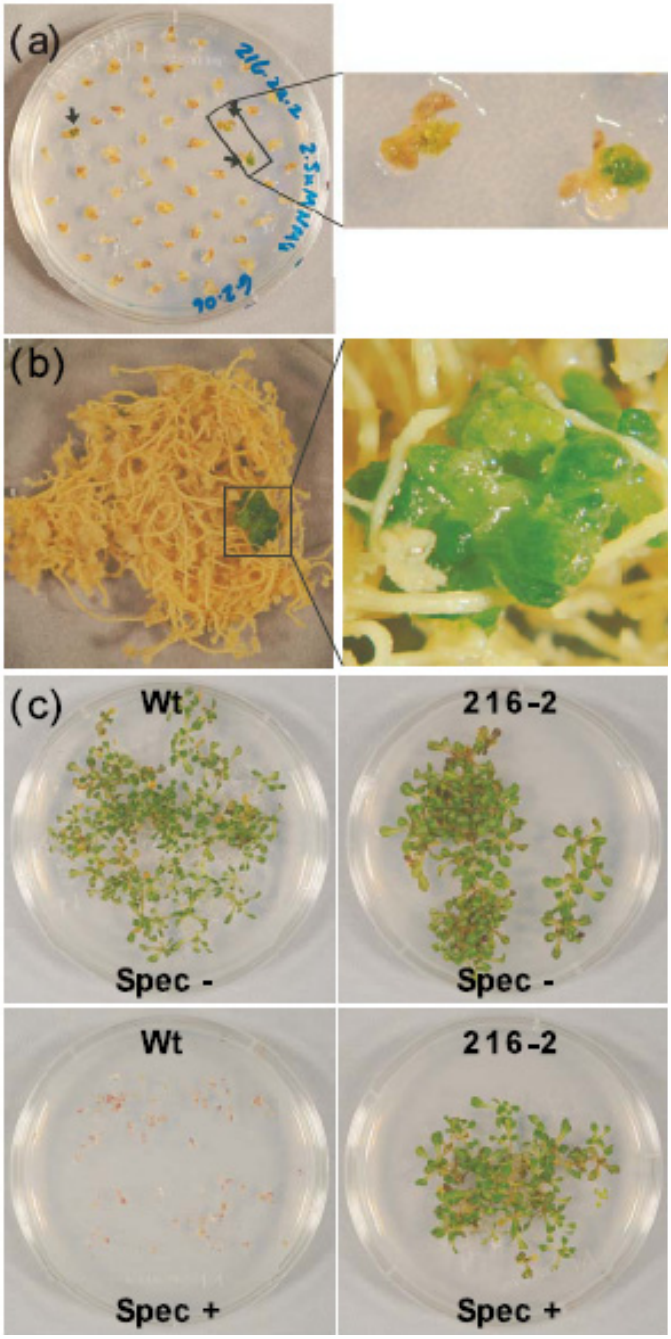


Figure 4.3. Isolation of spectinomycin resistant mutants in *Arabidopsis thaliana*.
(a) Identification of spectinomycin resistant mutants in mutagenized Arabidopsis seed culture. The seedlings were germinated on ARM-B medium containing 5 μM DEX and spectinomycin (25 mg l^{-1}) and selected on ARM-B medium containing 100 mg l^{-1} spectinomycin. Note mutants forming green embryogenic callus.
(b) The L*er*-SPC2 spontaneous spectinomycin resistant mutant in root culture.
(c) Inheritance of spectinomycin resistance in the L*er*-SPC2 seed progeny.

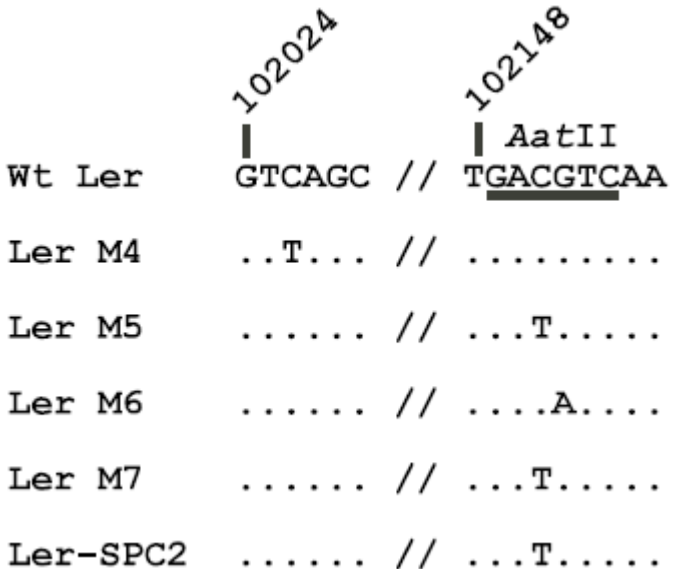


Figure 4.4. Mutations conferring spectinomycin resistance in the plastid *rrn16* gene encoding the 16S rRNA ribosomal subunit.

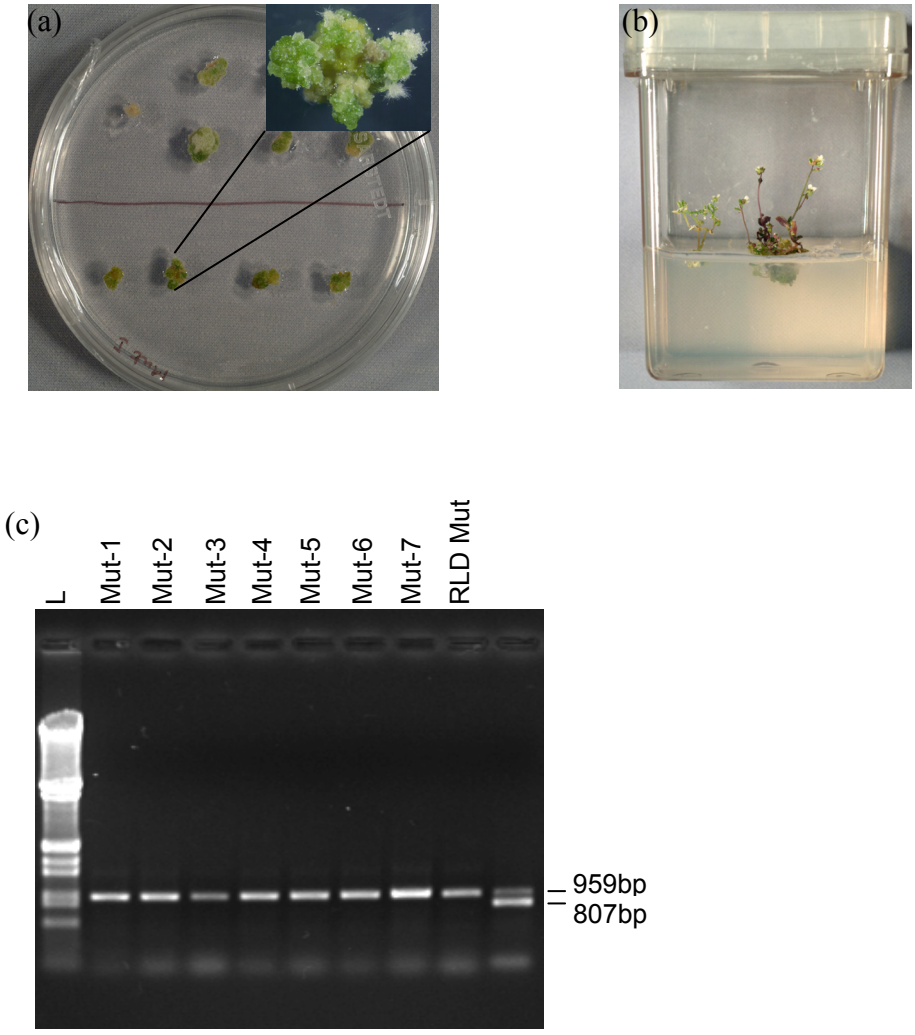


Figure 4.5. Selection of spectinomycin resistant mutants in PGA6-1 lines.
(a) Selection of spectinomycin resistant mutants (25 mg/l) in seedling callus after mutagenesis (2.5 mM NMU) in the presence of 17-β- Estradiol (5 mM).
(b) Flowering mutant plant regenerated from tissue shown in Figure 4a.
(c) *Aat*II digestion of PCR products confirms mutation in the *Aat*II site of the plastid *rrn16* gene.

Table 4.1. Selection of spectinomycin resistant mutants on ARM-B medium in mutagenized (2.5 mM NMU) Arabidopsis seedling culture

Line	**No. of seed**	**No. mutants**
RLD	200	11
RLD-216-22	200	7
L*er*	200	1
L*er*-216-2	410	5

Table 4.2. Selection of spectinomycin resistant mutants in mutagenized (10 mM NMU) Arabidopsis PGA6-1seedling culture

Ecotype	**No. of seed**	**Mutagen treatment**	**No. Spec^r mutants**
Ws	500	10 mM NMU (2 hrs)	7

References

Azhagiri, A. and Maliga, P. (2007) DNA markers define plastid haplotypes in *Arabidopsis thaliana. Curr Genet* **in press**.

Boutilier, K., Offringa, R., Sharma, V.K., Kieft, H., Ouellet, T., Zhang, L., Hattori, J., Liu, C.M., van Lammeren, A.A.M., Miki, B.L.A., Custers, J.B.M. and van Lookeren Campagne, M.M. (2002) Ectopic expression of BABY BOOM triggers a conversion of vegetative to embryonic growth. *Plant Cell* **14,** 1737-1749.

Clough, S.J. and Bent, A.F. (1998) Floral dip: a simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana. Plant J* **16,** 735-743.

Czako, M., Wilson, J. and Marton, L. (1993) Sustained root culture for generation and vegetative propagation of transgenic *Arabidopsis thaliana. Plant Cell Rep* **12,** 603-606.

Davis, S.J. and Vierstra, R.D. (1998) Soluble, highly fluorescent variants of green fluorescent protein (GFP) for use in higher plants. *Plant Mol Biol* **36,** 521-528.

Feldmann, K.A. and Marks, M.D. (1986) Rapid and efficient regeneration of plants from explants of *Arabidopsis thaliana. Plant Sci* **41,** 63-69.

Fromm, H., Edelman, M., Aviv, D. and Galun, E. (1987) The molecular basis for rDNA-dependent spectinomycin resistance in *Nicotiana* chloroplasts. *EMBO J* **6,** 3233-3237.

Hagemann, R. (1982) Induction of plastome mutationas by nitroso-urea-compounds. In *Methods in chloroplast molecular biology* (Edelman, M., Hallick, R.B. and Chua, N.H., eds). Amsterdam: Elsevier Biomedical Press, pp. 119-127.

Hajdukiewicz, P., Svab, Z. and Maliga, P. (1994) The small, versatile pPZP family of Agrobacterium binary vectors for plant transformation. *Plant Mol Biol* **25,** 989-994.

Johanson, U. and Hughes, D. (1995) A new mutation in 16S rRNA of *Escherichia coli* conferring spectinomycin resistance. *Nucleic Acids Res* **23,** 464-466.

Kavanagh, T.A., O'Driscoll, K.M., McCabe, P.F. and Dix, P.J. (1994) Mutations conferring lincomycin, spectinomycin, and streptomycin resistance in *Solanum nigrum* are located in three different chloroplast genes. *Mol Gen Genet* **242,** 675-680.

Lloyd, A.M., Schena, M., Walbot, V. and Davis, R.W. (1994) Epidermal cell fate determination in Arabidopsis: patterns defined by a steroid-inducible regulator. *Science* **266,** 436-439.

Lotan, T., Ohto, M., Yee, K.M., West, M.A., Lo, R., Kwong, R.W., Yamagishi, K., Fischer, R.L., Goldberg, R.B. and Harada, J.J. (1998) Arabidopsis LEAFY COTYLEDON1 is sufficient to induce embryo development in vegetative cells. *Cell* **93,** 1195-1205.

Luo, Y. and Koop, H.U. (1997) Somatic embryogenesis in cultured immature zygotic embryos and leaf protoplasts of *Arabidopsis thaliana* ecotypes. *Planta* **202,** 387-396.

Marton, L. and Browse, J. (1991) Facile transformation of *Arabidopsis. Plant Cell Rep* **10,** 235-239.

Murashige, T. and Skoog, F. (1962) A revised medium for the growth and bioassay with tobacco tissue culture. *Physiol Plant* **15,** 473-497.

Santos Mendoza, M., Dubreucq, B., Miquel, M., Caboche, M. and Lepiniec, L. (2005) LEAFY COTYLEDON 2 activation is sufficient to trigger the accumulation of oil and seed specific mRNAs in Arabidopsis leaves. *FEBS Lett* **579,** 4666-4670.

Schmidt, R. and Willmitzer, L. (1988) High efficiency *Agrobacterium tumefaciens*-mediated transformation of *Arabidopsis thaliana* leaf and cotyledon explants. *Plant Cell Rep* **7,** 583-586.

Sikdar, S.R., Serino, G., Chaudhuri, S. and Maliga, P. (1998) Plastid transformation in *Arabidopsis thaliana. Plant Cell Rep* **18,** 20-24.

Srinivasan, C., Liu, Z., Heidmann, I., Supena, E.D., Fukuoka, H., Joosen, R., Lambalk, J., Angenent, G., Scorza, R., Custers, J.B. and Boutilier, K. (2007) Heterologous expression of the BABY BOOM AP2/ERF transcription factor enhances the regeneration capacity of tobacco (*Nicotiana tabacum* L.). *Planta* **225,** 341-351.

Stone, S.L., Kwong, L.W., Yee, K.M., Pelletier, J., Lepiniec, L., Fischer, R.L., Goldberg, R.B. and Harada, J.J. (2001) LEAFY COTYLEDON2 encodes a B3 domain transcription factor that induces embryo development. *Proc Natl Acad Sci U S A* **98,** 11806-11811.

Svab, Z. and Maliga, P. (1991) Mutation proximal to the tRNA binding region of the *Nicotiana* plastid 16S rRNA confers resistance to spectinomycin. *Mol Gen Genet* **228,** 316-319.

Timmermans, M.C.P., Maliga, P., Vieira, J. and Messing, J. (1990) The pFF plasmids: cassettes utilizing CaMV sequences for expression of foreign genes in plants. *J Biotechnol* **14,** 333-344.

Valvekens, D., Van Montagu, M. and Van Lusebettens, M. (1988) *Agrobacterium tumefaciens*-mediated transformation of *Arabidopsis thaliana* root explants by using kanamycin selection. *Proc Natl Acad Sci U.S.A.* **85,** 5536-5540.

Zuo, J., Niu, Q.W. and Chua, N.H. (2002) The *WUSCHEL* gene promotes vegetative-to-embryonic transition in *Arabidopsis. Plant J* **30,** 349-359.

Author Contributions to research described in Chapter 4

(A) Experiments on steroid inducible BBM regeneration system:

Kerry Lutz, Arun Azhagiri and Prof. Pal Maliga designed research; KL performed research; KL, AA and PM analysed data, KL and PM prepared the manuscript.

(B) Experiments on steroid inducible *WUS* regeneration system:

Arun Azhagiri, Kerry Lutz and Prof. Pal Maliga designed research; AA performed research; AA, KL and PM analysed data, AA and PM prepared the manuscript.

CHAPTER 5

Significance

My research has shown that maternal plastid inheritance is the norm in Arabidopsis; however, exceptional paternal plastids can be transmitted to the progeny. The frequency of transmission was determined to be $3.9x10^{-5}$. This number is inflated by antibiotic selection in tissue culture. Under field conditions such events are less likely to occur and will probably be 10-100 times lower (Ruf et al., 2007). The reasons for this are (i) selection pressure and (ii) self-fertilization. As mentioned in Chapter 2, the pollen transmission events probably involved only a very few ptDNA copies and these could be detected only because selection for spectinomycin resistance encoded in the paternal ptDNA was used. In the absence of selection it is highly likely that these clones would have been missed. This would be the same situation under field conditions. Any paternal plastid present in F1 seedlings would be part of a heteroplasmic population of maternal and paternal plastids. Sorting out of plastids during somatic cell division would lead to the loss of one of the two plastid types and it is likely that without selective pressure the paternal plastids would be lost as they would be present in very low numbers. A second factor that lowers frequency of paternal pollen transmission is that Arabidopsis, like most crop plants, is self-fertilizing and the probability of cross-fertilization, when transplastomic and non-transplastomic plants are grown close to each other, is very low. The reasons given above also go towards addressing two major public concerns regarding transgenic crops; (i) uncontrolled flow of transgenes *via* pollen from fields with transgenic crops to non-transgenic crops in adjacent fields and (ii) transgene flow from transgenic crops to related wild species. Pollen transfer can occur from plant to plant by direct physical contact of the flowers, by wind or insects. In the European Union, the threshold level that triggers labeling of foods as genetically modified (GM) is 0.9% and is based on the DNA content. Containment of transgene flow due to insects is particularly important in pollinated crops. This would be an issue in crops with modified nuclear genome but in plants with modified plastid genome this would not be of concern as plastids are maternally inherited (with very rare exceptions as the study reported in Chapter 1 shows). Recent field studies measuring the outcrossing rates of transgenic

maize into conventional or organic maize cultivations, have shown that in a distance of 25 m, an average of 0.01% to 0.3% of GM maize was found in the non-GM maize cultivation. In a distance of 250 m, the average outcrossing rate rested between 0% and 0.003%. Accordingly, all measured rates are far below the EU labeling threshold of 0.9% (Van de Wiel and Lotz, 2006; Weber et al., 2007). Compared to nuclear transformation, the level of transgene containment achieved through plastid transformation is ample enough that transgenic and non-transgenic crops can be grown in adjacent stands. If these limits are not sufficient, then plastid transformation can be combined with other containment measures, such as cytoplasmic male sterility and transgene mitigation strategies. This data extends similar reports in Nicotiana [10^{-4}-10^{-5} paternal plastid transmission frequencies (Ruf et al., 2007; Svab and Maliga, 2007)] to a new species and suggests that rare paternal plastid transfer to seed progeny may be an exception rather than the rule in plants considered to have a strictly maternal inheritance of plastids. The data reported in Chapter 2 can be used as a starting point for studies in Arabidopsis that will aid in understanding the mechanisms regulating inheritance of plastids through pollen. Any knowledge obtained from such studies would help in designing systems that would prevent or reduce transgene flow from transgenic crops to neighboring stands or to weedy relatives.

In addition, in the course of my research I developed new tools. They include identification of ecotype specific ptDNA markers and a plastid-encoded spectinomycin resistant Arabidopsis mutant using a steroid-inducible BBM plant regeneration system. The ptDNA markers, I identified, are also available in ecotypes commonly used in genetic studies. Study of rare paternal pollen transmission in Arabidopsis was possible because the ecotype-specific ptDNA markers enabled differentiation between parental plastids and the spectinomycin resistant mutant allowing fortracking and selecting paternal plastids in hybrid seed progeny.

The development of the phiC31 phage Integrase and the identification of *pseudo-lox* sites in the plastid genome provide alternate approaches to transforming the plastid genome. These technologies can be explored to transform plastids of plants with recalcitrant homologous recombination machinery. Proof of concept for the phiC31 Integrase was shown in tobacco (reported in Appendix I). The phiC31 Integrase, in conjunction with the BBM plant regeneration system, should facilitate the transformation of the Arabidopsis plastid genome. Further, the *pseudo-lox* sites identified in the plastid genome can be used as

targets for CRE-mediated insertion of transgenes into the plastid genome. Although the use of *pseudo-lox* sites as targets for insertion of genes into the plastid genome has not been attempted, the use of these sites for recombination with *loxP* sites has previously been reported (Sauer, 1992; Thyagarajan et al., 2000).

References

Ruf, S., Karcher, D. and Bock, R. (2007) Determining the transgene containment level provided by chloroplast transformation. *Proc Natl Acad Sci USA* **104,** 6998-7002

Sauer, B. (1992) Identification of cryptic *lox* sites in the yeast genome by selection for Cre-mediated chromosome translocations that confer multiple drug resistance. *J Mol Biol* **223,** 911-928.

Svab, Z. and Maliga, P. (2007) Exceptional transmission of plastids and mitochondria from the pollen parent and its impact on transgene containment. *Proc Natl Acad Sci U S A* **104,** 7003-7008.

Thyagarajan, B., Guimaraes, M.J., Groth, A.C. and Calos, M.P. (2000) Mammalian genomes contain active recombinase recognition sites. *Gene* **244,** 47-54.

Van de Wiel, C. C. M. and Lotz, L. A. P. (2006) Outcrossing and coexistence of genetically modified with (genetically) unmodified cops: a case study of the situation in the Netherlands. *Neth J Agr Sci* **54**, 17-35

Weber, W. E., Bringezu, T., Broer, I., Eder, J. and Holz, F. (2007) Coexistence between GM and non-GM maize crops - Tested in 2004 at the field scale level (Erprobunganbau 2004). *J Agron Crop Sci* **193,** 79-92.

Appendix 1

Developing the phiC31 phage site-specific Integrase to transform plastids

Thus far plastid transformation in higher plants has been based on incorporation of foreign DNA in the plastid genome by the plastid's homologous recombination machinery(Svab and Maliga, 1993). The report here is of an alternative approach that relies on integration of foreign DNA by the phiC31 phage site-specific integrase (INT) mediating recombination between bacterial *attB* and phage *attP* sequences(Groth et al., 2000). Plastid transformation by the new approach depends on the availability of a recipient line in which an *attB* site has been incorporated in the plastid genome by homologous recombination. Plastid transformation involves insertion of an *attP* vector into the *attB* site by INT and selection of transplastomic clones by selection for antibiotic resistance carried in the *attP* plastid vector. INT function was provided either by expression from a nuclear gene which encoded a plastid-targeted INT or by expressing INT transiently from a non-integrating plasmid in plastids. Transformation, in Nicotiana, was successful with both approaches using *attP* vectors with kanamycin-resistance or spectinomycin resistance as the selective marker. Transformation efficiency in some of the stable nuclear INT lines was as high as 17 independently transformed lines per bombarded sample. Since this system does not rely on the plastid's homologous recombination machinery we expect that INT-based vectors will make plastid transformation a routine in species in which homologous recombination rarely yields transplastomic clones.

Results

Construction of plants with the attB *plastid genome*

Minimal sequences required for *attB* and *attP* function in vitro were defined to be 34 bp and 39 bp, respectively (Groth et al., 2000). In mammalian cells, longer sequences were found to be more efficient (Groth et al., 2000). Therefore, we utilized 55-bp *attB* and 217-bp *attP* regions.

To create the *attB* recipient plastid genome, a suitable plastid transformation vector was designed. Plasmid pSAC114 contains plastid targeting sequences, an *attB* sequence and a linked spectinomycin resistance (*aadA*) gene flanked by *loxP* sites (floxed) (Figure A1.2a). Vector pSAC114 was introduced into tobacco leaf chloroplasts by the biolistic process to obtain the R1 recipient plastid genome. Plastid transformants containing the *attB* site were obtained by selection for spectinomycin resistance. The R1 recipient plastid genome structure was verified by DNA gel blot analysis (Figure A1.2b). Four pSAC114 transplastomic lines were purified to the homoplastomic state: Nt-pSAC114-2A, Nt-pSAC114-3B, Nt-pSAC114-5A and Nt-pSAC114-63A. The four lines are considered identical.

To obtain marker-free *lox/attB* plants (R2 recipient), a plastid-targeted *Cre* gene (pKO30) (Corneille et al., 2003) was introduced into the nucleus of the Nt-pSAC114-63A plant by *Agrobacterium* transformation. When all copies of the floxed *aadA* marker gene were removed (Figure A1.2a, c), the integrated nuclear *Cre* gene was segregated out in the seed progeny yielding the Nt-LA63 recipient line.

Plastid attP *vectors*

The pKO103 *attP* plastid vector was obtained by incorporating an *attP* site in an *E. coli* (pBSIIKS+) plasmid and replacing the ampicillin resistance gene with a *neo* gene that confers kanamycin resistance to both *E. coli* and plastids. Since the vector cannot replicate in plastids, kanamycin resistant plants will only be obtained if the vector integrates into the plastid genome. An alternative *attP* vector, pKO193, was obtained by replacing the ampicillin resistance gene with a spectinomycin resistance (*aadA*) gene as a

selective marker; a *loxP* site for excision of the marker gene; and a *bar* gene that confers herbicide resistance.

Stable nuclear INT for plastid transformation

Since construction of the Nt-LA63 recipient line took time, we decided to test INT-mediated transformation using R1 recipients (Figure A1.3). The R1 INT recipients were obtained by *Agrobacterium* transformation of the R1 recipient (Nt-pSAC114) with vector pKO117, which encodes a plastid-targeted *Int* and a gentamycin resistance (*aacC1*) gene (Figure A1.4). The INT lines with the R1 plastid genome are designated N1-N18 (Table A1.1).

To test the feasibility of INT-mediated plastid transformation, the pKO103 plastid vector (Figure A1.5b) was bombarded into the leaves of N1-N18 INT recipient lines. Two days after bombardment, the leaves (~7 cm long) were cut into ~1 cm^2 pieces (20 to 30 pieces per leaf) and selected for kanamycin resistance (50 mg/L). The number of kanamycin resistant clones per bombarded leaf was between 1 and 19. In 45 bombarded leaf cultures 237 independent kanamycin resistant lines were obtained. Of these, 113 were evaluated for integration of the pKO103 vector at the *attB* site (Table A1.1). The individual transformed lines were identified by the INT line and a serial number, e.g. N1-49. Letters after the serial number distinguish subclones derived from the same event.

INT-mediated insertion of the *attP* vector at the *attB* site was confirmed in 81 of the 113 clones (~72%; Table A1.1). Integration was verified by PCR amplification of border fragments obtained with primer pairs 01-02 and 03-04, one which annealed to the pKO103 vector part and one which annealed to the recipient plastid DNA (Figure A1.5c). Sequencing of PCR-amplified junction fragments revealed *attL* and *attR* junction sequences, the hallmarks of INT-mediated recombination (Figure A1.5e). DNA gel blot analysis in the 16 clones tested confirmed integration at the *attB* site as the *neo* probe recognized the predicted 6.1-kb fragment in the transplastomic clones (Figure 5d). DNA gel blots of 16 kanamycin resistant clones, in which no insertion at the *attB* site was found, were also probed with *neo*. No high-copy signal was detected on these blots therefore, no integration of *attP* vectors at alternative plastid target sites occurred (data not shown).

In one sample (N1-49A) a smaller than expected fragment (4.7 kb) was seen (Figure A1.5d). This smaller fragment is the result of a post-integration deletion event between 18-bp direct repeats in the pKO103 multiple cloning site (MCS) and restriction sites adjacent to the *attR* site (verified by sequencing PCR-amplified DNA).

Transformation mediated by INT transiently expressed in plastids

When a nuclear gene provides INT function, as described in the previous section, *Int* has to be removed when transformation is accomplished. To avoid this tedious process, we tested the feasibility of providing INT function transiently, from a non-integrating plasmid. This was accomplished by incorporating an *int* gene with plastid expression signals in an *E. coli* (pBSIIKS+) plasmid, which could be introduced together with the *attP* plastid vector (Figure A1.6). Two *int* helper plasmids, pKO107 and pKO111, were tested. Translation of INT from the pKO111 mRNA was dependent on the creation of an AUG translation initiation codon from an ACG codon by mRNA editing, whereas no editing was required for the translation of the pKO107 mRNA.

The transient INT system was tested by introducing two plasmids, the pKO193 *attP* plastid vector (Figure A1.7c) and an *int* helper plasmid (pKO107 or pKO111)(1:2 ratio), into the chloroplasts and selecting for spectinomycin resistance. Bombardment of five leaves of the Nt-LA63 recipient line with the mix containing the pKO107 helper plasmid yielded five spectinomycin resistant lines. Of these, three were confirmed as products of INT-mediated insertion (Figure A1.7e). Bombardment of five Nt-LA63 leaves with the mix containing the pKO111 plasmid yielded seven spectinomycin resistant lines of which one was the product of INT-mediated insertion (Figure A1.7e). The plastid *int* gene encoded in the helper plasmids did not incorporate in the plastid or nuclear genomes, as confirmed by PCR analysis (data not shown).

Discussion

We report here efficient, unidirectional integration of foreign DNA into the genome of a eukaryotic organelle, the tobacco plastid by the phiC31 phage site-specific integrase. Implementation of the system is dependent on the availability of a recipient line carrying an *attB* (or *attP*) site introduced into the plastid genome by homologous recombination. We have shown here that INT mediates plastid transformation when it is provided either from a nuclear-encoded plastid targeted gene or from a non-replicating helper plasmid.

Approaches to providing INT function

Incorporation of INT into the nucleus, although effective, requires that the gene be removed once recombination is complete. Transient expression of the recombinase allows for sufficient amounts of INT protein to be expressed without the need to remove the *int* gene once the reaction is complete. Plastid transformation with the transiently expressed INT was relatively efficient yielding 4 independently transformed transplastomic lines after bombardment of 10 leaves, five leaves each with the pKO107 (3 lines) and pKO111 (1 line) helper. In a preliminary experiment, we obtained 19 transplastomic clones in 8 bombarded leaf samples confirming that transient expression of INT reproducibly yields transplastomic clones. The preliminary experiment was carried out with the R1 recipient, the pKO103 *attP* transformation vector and the pKO111 INT helper plasmid. Thus, in the transient system, kanamycin (pKO103) and spectinomycin (pKO193) resistance markers were both suitable to recover transplastomic clones and both the edited (pKO111) and unedited (pKO107) *int* genes provided INT function.

We report here that we did not find integrated copies of the helper plasmids in the plastid or nuclear genomes. However, it is possible that the helper plasmid was transiently maintained during the initial stage of transformation resulting in maintenance of INT activity for a relatively long period of time. Possible mechanisms for transient maintenance of the helper plasmid are cointegrate formation by recombination *via* ptDNA sequences driving INT expression, or independent replication of the helper plasmid which has been shown to occur in *Chlamydomonas reinhardtii* chloroplasts (Boynton et al., 1988).

These transiently maintained helper plasmids, if present at all, were lost by the time we tested *int* by PCR in homoplastomic plants.

The attB *recipient lines*

Proof of concept for INT mediated plastid transformation was obtained with the R1 recipient plastid genome carrying a floxed *aadA* plastid marker gene. Removal of the floxed *aadA* by CRE yielded the marker free R2 recipient genome, which can now be transformed with an *attP* vector carrying either *aadA* or *neo* as the selective marker. The R2 recipient genome and the *attP* transformation vector each carry one suitably positioned *loxP* site so that the integrated marker gene can be excised *via* the directly oriented *loxP* sites (Figure A1.7d) when uniform transformation of all plastid genomes is obtained, as reported earlier (Corneille et al., 2001; Hajdukiewicz et al., 2001).

No functional attB *sites were found in the plastid genome*

We have previously identified functional *lox* sites in the plastid genome, which yielded recombination products with *loxP* sites (Corneille et al., 2003). Therefore, we tested if INT-mediated integration can occur into plastid sequences that function as *attB* sites. Candidates for integration at plastid *attB* sequences were kanamycin-resistant clones, in which no *attP* vector insertion was found at the *attB* site. Probing of total cellular DNA of these plants with *neo* did not yield a high-copy signal that would suggest integration of the vector in the plastid genome at alternative sites (data not shown). Thus, the kanamycin resistant clones were probably obtained by fortuitous integration of *neo* in the nucleus (Carrer et al., 1993).

Alternative att *vector designs*

INT-mediated transformation efficiency in plastids may be further improved by incorporating *attP* instead of *attB* in the plastid genome as a target site, because it was shown that the *attP* sequence could be a more efficient target (Thyagarajan, 2001; Belteki et al., 2003).

An alternative approach to integrating an entire plastid transformation vector would be the cassette exchange system where only sequences flanked by *attP* (or *attB*) sites are integrated into the plastid genome *via* two *attB* (or *attP*) sites suitably positioned in the plastid genome (Thomason et al., 2001; Ow, 2002; Belteki et al., 2003). By this design only sequences between the *attP* (or *attB*) sites would be integrated in the plastid genome. However, even if the entire vector is incorporated in the plastid genome, remaining vector sequences may be removed along with the marker gene by CRE when transformation is accomplished *via* strategically positioned *loxP* sites. For this, the relative position of *ori* and *attP* sequences should be swapped in the *attP* vector shown in Figure A1.7c.

Application of the INT/attB/attP *system for plastid transformation*

We have shown here that the INT system is a useful alternative to homologous targeting for plastid transformation in tobacco. Plastid transformation efficiency in some of the stable nuclear INT lines was unusually high, up to 17 transplastomic clones per bombarded sample (Table A1.1). Homologous recombination-based plastid transformation efficiency numbers range from 0.5 to 5.0 per bombarded sample (Maliga, 2003, 2004). Since only a relatively small number of experiments were performed, it would be premature to draw the conclusion that INT-based plastid transformation is more efficient than homologous-recombination based transformation in tobacco. However, we expect that the INT-based system may enable reproducible plastid transformation in species, in which homologous recombination rarely yields transplastomic clones. Combination of the INT/*attB*/*attP* system for insertion of the transforming DNA and the CRE-*loxP* site-specific recombination system for removal of the marker gene (Corneille et al., 2001; Hajdukiewicz et al., 2001; Kuroda and Maliga, 2003) will facilitate applications of plastid transformation in basic science and biotechnology (Bock, 2001; Maliga, 2004).

Experimental procedures

Plastid vector for construction of attB *plants*

Plastid vector pSAC114 is a pPRV1 vector derivative (Zoubenko et al., 1994) in which an *aadA* gene, flanked by two directly oriented *lox* sites, and an *attB* sequence are targeted for insertion at the *trnV-rps12/7* intergenic region. Directly oriented *loxP* sites were available in plasmid pSAC61 (pBSIIKS+ backbone; Stratagene, La Jolla, CA) in a *Kpn*I-*Hin*dIII fragment (5'-ggtaccATAACTTCGTATAATGTATGCTATACGAAGTTATAGATCTATAACTTCGTATAATGTATGCTATACGAAGTTATaagctt-3'). Plasmid pSAC61 was obtained by cloning a *Kpn*I-*Sal*I adapter (5'-ggtaccATAACTTCGTATAATGTATGCTATACGAAGTTATagatctgtcgac-3') in a pBSIIKS+ vector (Stratagene, La Jolla, CA) to obtain plasmid pSAC60, then ligating a *Bgl*II-*Hin*dIII adapter (5'-agatctATAACTTCGTATAATGTATGCTATACGAAGTTATaagctt-3') in plasmid pSAC60. Nucleotides added to aid construction are shown in lower case. The *aadA* gene was cloned into the *Bgl*II site located between the *loxP* sequences in pSAC61. The *aadA* gene derives from plasmid pZS176, which contains an identical *aadA* gene in plasmid pZS197 (Svab and Maliga, 1993) other than it has an *Eco*RI site in the 5'-UTR. The *attB* sequence was cloned upstream of *aadA* as a *Pst*I-*Spe*I linker 5'-CTGCAGCCGCGGTGCGGGTGCCAGGGCGTGCCCTTGGGCTCCCCGGGCGCGTACTCCACTAGT-3'.

Agrobacterium *vector with nuclear* Int *gene*

Tobacco *attB* plants expressing a nuclear *Int* were obtained by transformation with *Agrobacterium* binary vector pKO117. Plasmid pKO117 is a pPZP222 vector derivative with gentamycin resistance (*aacC1*) as selective marker (Hajdukiewicz et al., 1994) and a plastid-targeted *Int* in the multiple cloning site. Plasmid pKO117 was obtained by replacing the *Cre* coding region in plasmid pKO27 (Corneille et al., 2001) with an *Int* coding region. The *Int* gene was then cloned into a pPZP222 vector. The *Streptomyces* ssp. phiC31 phage *Int* coding segment derives from plasmid pSET152 (Bierman et al., 1992)(GenBank Accession No.

AJ414670) obtained from Mark Buttner. To facilitate *Int* expression, an *Nco*I site was added as an *Nco*I-*Sun*I linker at the N-terminus to include the ATG translation initiation codon (5'-ccATGgctagcGACACGTACG-3') and an *Xba*I site was created downstream of the stop codon by PCR amplification (5'-TAGtctaga-3'). *Int* is expressed in a cassette consisting of the *Agrobacterium* P2' promoter and Tnos terminator, and is translationally fused with the pea Rubisco small subunit transit peptide (Timko et al., 1985) and twenty-two amino acids of the mature small subunit.

Helper plasmids for transient INT expression in plastids

Plasmid pKO107 is a pUC118 vector derivative obtained by replacing the *aadA* coding segment in plasmid pZS176 (above) with the *Int* coding segment (*Nco*I-*Xba*I fragment). Plasmid pKO111 was obtained by replacing the *neo* coding segment with an *Int* coding segment (*Nhe*I-*Xba*I fragments) in plasmid pSC4 (Chaudhuri and Maliga, 1996).

Plastid attP *vectors*

Plastid *attP* vector pKO103 is a pBSII KS+ (Stratagene, La Jolla, CA) derivative that contains an *attP* sequence, a *neo* marker gene, a *loxP* site (not shown in Figure A1.5) and a multiple cloning site (MCS). The *attP* sequence was PCR amplified from plasmid pSET152 (primers: 5'-caactcGAGCAATCGCCCTGGGTG-3' and 5'-caaaagcttCCCGGTCACAACCCCTTG-3') and cloned as an *Xho*I-*Hin*dIII fragment in the multiple cloning site of the pBSIIKS+ plasmid. The kanamycin resistance (*neo*) gene from plasmid pHK10 (blunt *Sac*I/*Hind*III fragment; *Xba*I site removed)(Kuroda and Maliga, 2001) was ligated in the *Hin*dIII site. The *attP*-*neo* fragment (*Xho*I-*Eco*RV blunt) was cloned into the *Kpn*I site of plasmid pSAC60. The ampicillin resistance gene and the f1 replication origin were removed by religating the *Ahd*I and *Sac*I digested plasmid.

Plastid *attP* vector pKO193 is pKO103 derivative, in which the *neo* gene was replaced with an *aadA* gene from plasmid pZS176 (see above) and the *bar* gene from plasmid pKO18 (*Eco*RI-*Hin*dIII fragment) (Lutz et al., 2001) was cloned into the multiple cloning site.

Plastid transformation

Transforming DNA was introduced into tobacco leaves on the surface of tungsten particles by the biolistic process (Svab and Maliga, 1993). Transplastomic clones transformed with plasmids pSAC114 or pKO193 were selected on RMOP medium containing 500mg/L spectinomycin HCl (Svab and Maliga, 1993). Clones transformed with plasmid pKO103 were selected on RMOP medium containing 50mg/L kanamycin sulfate (Carrer et al., 1993). Plastid transformation was confirmed by DNA gel blot analysis of total cellular DNA (Svab and Maliga, 1993). PCR amplification of junction sequences of pKO103-transformed clones was performed with primers O1 5'-CCGCCAGCGTTCATCCTGAGC-3' and O2 5'-TGACAGCCGGAACACGGCGGC-3' for the left (*attL*) junction or primers O3 5'-TGAAGAGCTTGGCGGCGAAT-3' and O4 5'-GAGATGTAACTCCAGTTCC-3' for the right (*attR*) junction.

Agrobacterium-*mediated transformation*

The *Int* gene was introduced into the nucleus of Nt-pSAC114 transformed plants by cocultivation of Nt-pSAC114 leaves with *Agrobacterium* containing the pKO117 binary vector (Hajdukiewicz et al., 1994). Nuclear gene transformants were selected by gentamycin resistance on RMOP shoot regeneration medium containing 100mg/L gentamycin. Transformation was confirmed by rooting shoots on selective plant maintenance (Murashige and Skoog salts, 3% sucrose, pH 5.6)(Murashige and Skoog, 1962) medium containing 100mg/L gentamycin.

Sequencing of attB/attP *junctions*

Junction sequences of pKO103-transformed clones were PCR amplified using primers O1 5'-CCGCCAGCGTTCATCCTGAGC-3' and O2 5'-TGACAGCCGGAACACGGCGGC-3' for the left junction or primers O3 5'-TGAAGAGCTTGGCGGCGAAT-3' and O4 5'-GAGATGTAACTCCAGTTCC-3' for the right junction. PCR products were sequenced using

oligonucleotides O1-O4 as primers with the T7 Sequenase PCR Product Sequencing Kit (USB, Cleveland, OH).

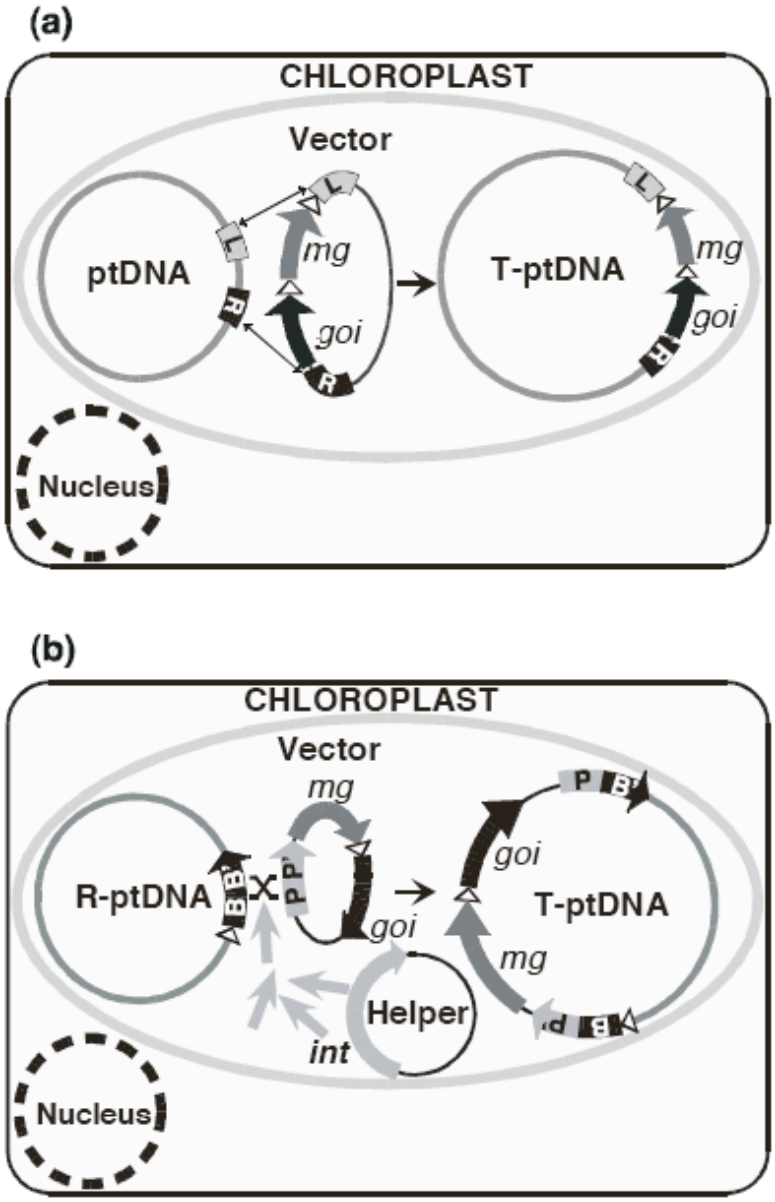


Figure A1.1. Approaches to plastid transformation.
(a) Plastid transformation based on the plastid's recombination machinery and homologous targeting. Note: wild-type plastid genome (ptDNA), left (L) and right (R) targeted regions; transformation vector with L and R targeting regions, marker gene (*mg*) flanked by *loxP* sites (floxed; open triangles) and gene of interest (*goi*); and the engineered plastid genome (T-ptDNA).
(b) Plastid transformation mediated by INT. Shown are: recipient plastid genome (R-ptDNA) with *attB* (BB') and *loxP* sites; plastid vector with *attP* site (PP'), marker gene (*mg*), *loxP* site and gene of interest (*goi*); the INT plasmid and engineered plastid genome (T-ptDNA) with *attL* (BP') and *attR* (PB') recombination junctions.

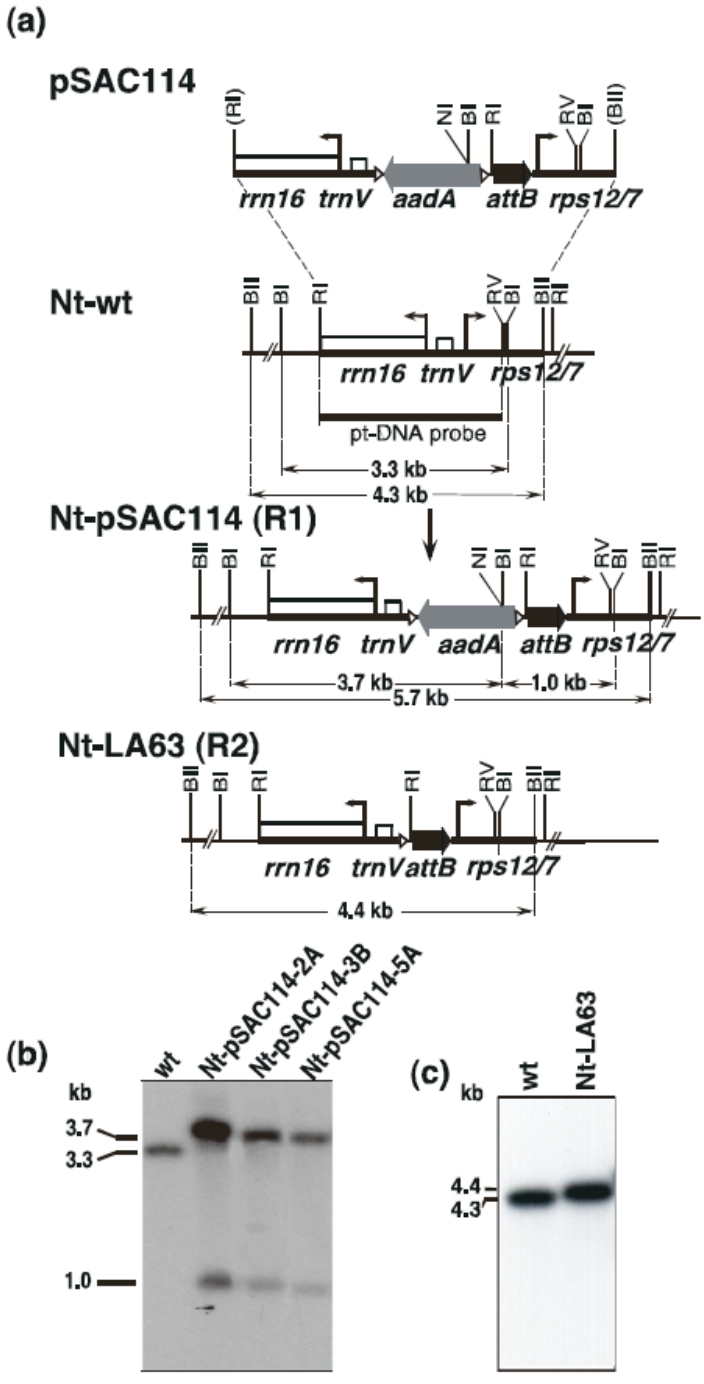


Figure A1.2. Verification of the *attB* plastid genome structure.
(a) Maps of the pSAC114 vector, the wild type (Nt-wt), the Nt-pSAC114 (R1) and Nt-LA63 (R2) plastid genomes. Positions of the plastid genes *rrn16*, *trnV*, *rps12/7*, the *aadA* gene, the *attB* sequence and the *Bam*HI (BI), *Bgl*II (BII), *Eco*RI (RI), *Eco*RV (RV) and *Nco*I (NI) restriction sites are marked.
(b) DNA gel blot analysis confirms ptDNA transformation. Total cellular DNA was digested with the *Bam*HI restriction endonuclease and probed with the wild-type *Eco*RI-*Eco*RV plastid DNA fragment (ptDNA probe, heavy line, Figure A1.2a). (c) DNA gel blot analysis to confirm *aadA* excision. Total cellular DNA was digested with the *Bgl*II restriction endonuclease and probed as Figure A1.2b.

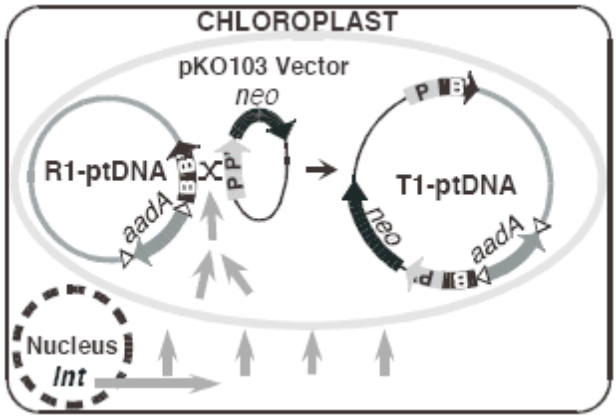


Figure A1.3. Plastid transformation mediated by INT expressed from a stably integrated nuclear gene. Shown are: recipient plastid genome (R1-ptDNA) with *attB* site (BB') and floxed *aadA* marker gene; plastid vector with *attP* site (PP') and a *neo* marker gene; and engineered plastid genome (T1-ptDNA) with *attL* (BP') and *attR* (PB') recombination junctions.

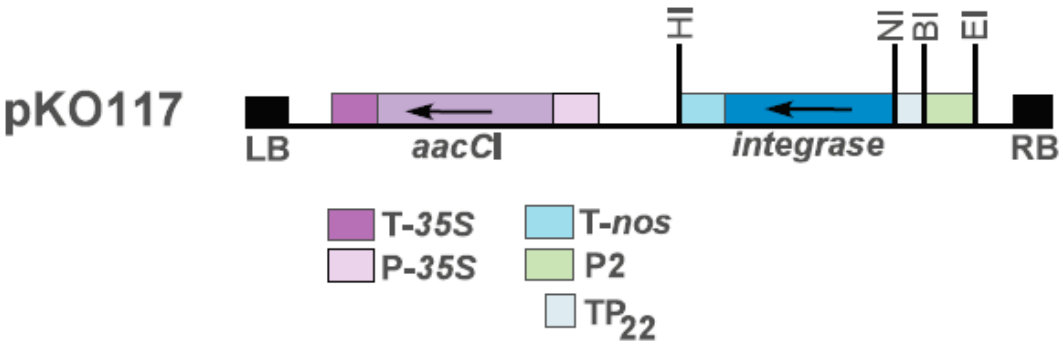


Figure A1.4. Schematic map of the pKO117 *Agrobacterium* binary vector T-DNA region with the plastid-targeted *Int* gene. Abbreviations: LB and RB, left and right border sequences; *aacC1*, gentamycin resistance gene; P2' and Tnos *Agrobacterium* P2' promoter and *nos* terminator; TP_{22}, Rubisco small subunit transit peptide, with 22 amino acids of the mature small subunit N-terminus.

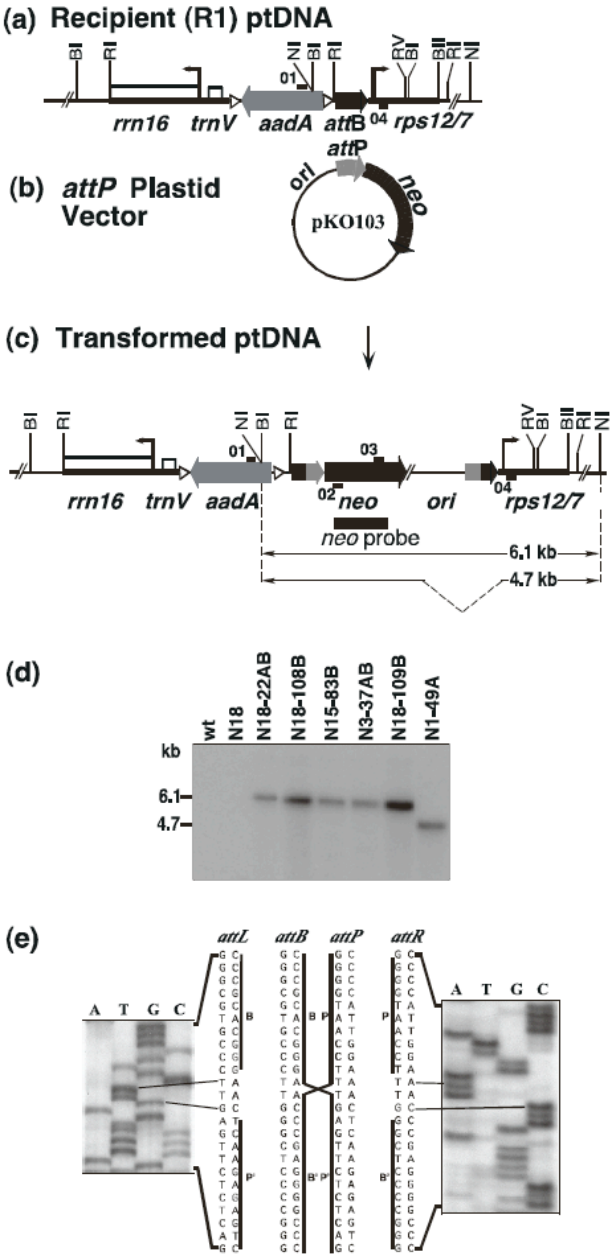


Figure A1.5. INT-mediated insertion of the pKO103 *attP* vector at the plastid *attB* site.
(a) Map of R1 plastid genome with an *attB* sequence (heavy black arrow) and floxed *aadA*. Position of plastid genes *rrn16*, *trnV*, *rps12/7*, the *Bam*HI (BI), *Bgl*II (BII), *Eco*RI (RI), *Eco*RV (RV) and *Nco*I (NI) restriction sites and 01 and 04 primers are marked.
(b) Map of the pKO103 *attP* plastid vector. Note *attP* site, *neo* marker gene and *Col*E1 replication origin (*ori*).
(c) Map of T1-ptDNA. Fragments detected by the *neo* probe and the positions of O1-O4 primers are shown.
(d) Probing of DNA gel blots with *neo* confirms integration of the pKO103 vector at the *attB* site. wt and N18 are DNA samples from wild-type and N18 recipient (R1) plants. Total cellular DNA was digested with the *Nco*I restriction endonuclease and probed with *neo* (heavy line, Figure A1.5c).
(e) The *attL* and *attR* junction sequences confirm INT-mediated insertion of the pKO103 vector at the *attB* site.

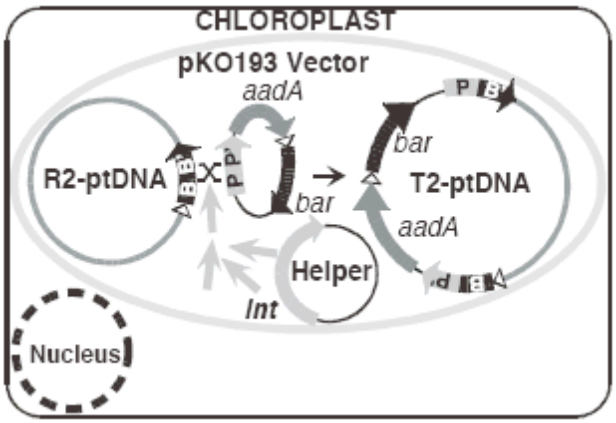


Figure A1.6. Plastid transformation mediated by INT transiently expressed in plastids. Note: recipient plastid genome (R2-ptDNA) with *attB* (BB') and *loxP* sites; plastid vector pKO193 with *attP* site (PP'), *aadA*, *loxP* site and *bar* gene; the INT plasmid and the engineered plastid genome (T2-ptDNA) with *attL* (BP') and *attR* (PB') recombination junctions.

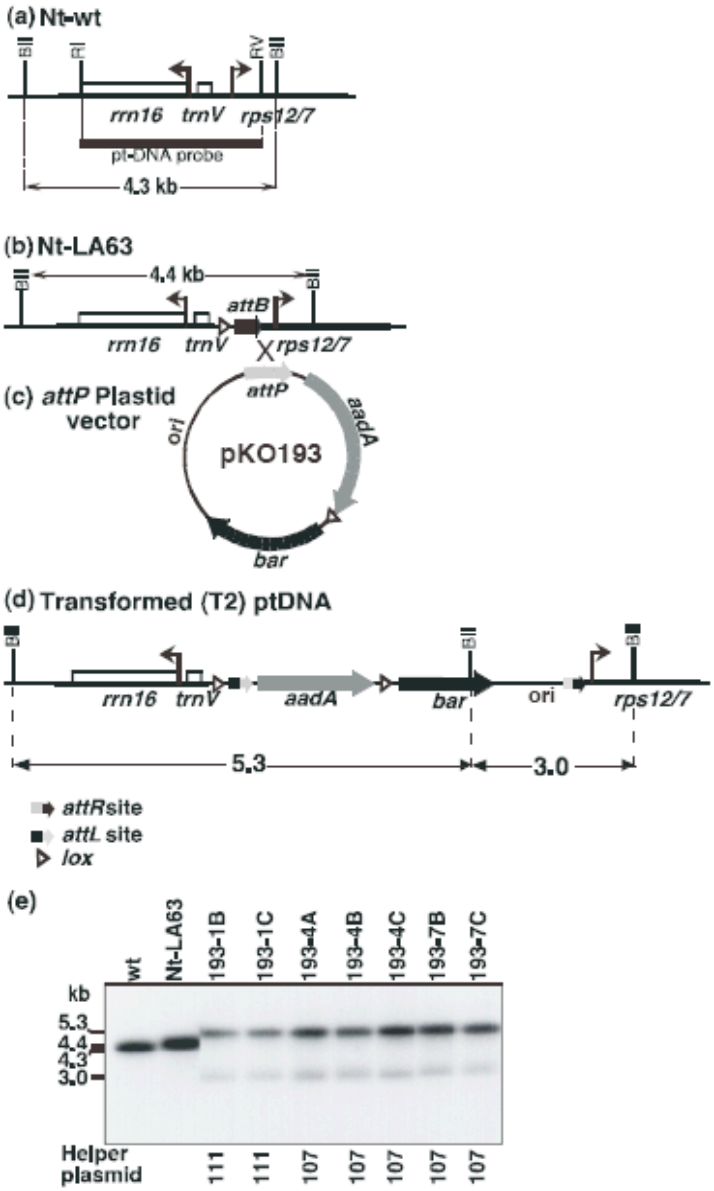


Figure A1.7. Transiently expressed INT mediates plastid transformation with pKO193.
(a) Map of the wild-type (Nt-wt) plastid genome. Positions of plastid genes *rrn16*, *trnV*, *rps12/7*, and the *Bgl*II (BII), *Eco*RI (RI), and *Eco*RV (RV) restriction sites are marked.
(b) Map of Nt-LA63 recipient plastid genome with a *loxP* site (open triangle) and an *attB* sequence (heavy black arrow). Positions of plastid genes *rrn16*, *trnV*, *rps12/7*, the *attB* sequence and the *Bgl*II (BII), restriction site are marked.
(c) Map of the pKO193 *attP* plastid vector. Marked are the *attP* site, the *aadA* marker gene, the *loxP* site, the *bar* gene and the *Col*E1 replication origin (*ori*).
(d) Map of the plastid genome transformed with the pKO193 vector. The 5.3 kb and 3.0 kb fragments detected by the *rrn16-rps12/7* targeting region probe are shown.
(e) DNA gel blot analysis confirms integration at the *attB* site. Wt and Nt-LA63 are DNA samples from wild-type plants and R2 recipient lines. Total cellular DNA was digested with the *Bgl*II restriction endonuclease and probed with the *Eco*RI-*Eco*RV *rrn16/rps12/7* targeting region (heavy line, Figure A1.7a). Position of hybridizing fragments is shown in Figure A1.7d.

Table A1.1 Integration of the pKO103 *attP* vector at the plastid *attB* site

INT Line	No. of Leaves	Kan^r	Kan^r /leaf	Insertion at *attB**	Efficiency**
N1	3	10	3.3	10/10	3.3
N2	3	14	4.7	6/10	2.8
N3	2	24	12.0	7/7	12.0
N4	1	5	5.0		
N5	1	2	2.0		
N6	4	13	3.3	8/10	2.6
N7	1	9	9.0		
N8	4	17	4.3	0/16	0.0
N9	1	11	11.0	5/10	5.5
N10	1	1	1.0		
N11	2	14	7.0	10/10	7.0
N12	2	13	6.5	10/10	6.5
N13	4	4	1.0		
N14	4	3	0.8		
N15	3	22	7.3	10/10	7.3
N16	4	25	6.3	6/10	3.8
N17	3	12	4.0		
N18	2	38	19.0	9/10	17.1

* Number of transplastomic clones/Number of clones tested.
** Efficiency = (Kan^r/leaf) x Insertion at *attB*. Values were calculated to adjust for differences in the fraction of transplastomic clones in INT lines.

References

Belteki, G., Gertsenstein, M., Ow, D.W. and Nagy, A. (2003) Site-specific cassette exchange and germline transmission with mouse ES cells expressing phiC31 integrase. *Nat Biotechnol* **21,** 321-324.

Bierman, M., Logan, R., O'Brian, K., Seno, E.T., Rao, R.N. and Schoner, B.E. (1992) Plasmid cloning vectors for the conjugal transfer of DNA from *Escherichia coli* to *Streptomyces* spp. *Gene* **116,** 43-49.

Bock, R. (2001) Transgenic plastids in basic research and plant biotechnology. *J Mol Biol* **312,** 425-438.

Boynton, J.E., Gillham, N.W., Harris, E.H., Hosler, J.P., Johnson, A.M., Jones, A.R., Randolph-Anderson, B.L., Robertson, D., Klein, T.M., Shark, K.B. and Sanford, J.C. (1988) Chloroplast transformation in *Chlamydomonas* with high velocity microprojectiles. *Science* **240,** 1534-1538.

Carrer, H., Hockenberry, T.N., Svab, Z. and Maliga, P. (1993) Kanamycin resistance as a selectable marker for plastid transformation in tobacco. *Mol Gen Genet* **241,** 49-56.

Chaudhuri, S. and Maliga, P. (1996) Sequences directing C to U editing of the plastid *psbL* mRNA are located within a 22 nucleotide segment spanning the editing site. *EMBO J* **15,** 5958-5964.

Corneille, S., Lutz, K., Svab, Z. and Maliga, P. (2001) Efficient elimination of selectable marker genes from the plastid genome by the CRE-*lox* site-specific recombination system. *Plant J* **72,** 171-178.

Corneille, S., Lutz, K.A., Azhagiri, A.K. and Maliga, P. (2003) Identification of functional *lox* sites in the plastid genome. *Plant J* **35,** 753-762.

Groth, A.C., Olivares, E.C., Thyagarajan, B. and Calos, M.P. (2000) A phage integrase directs efficient site-specific integration in human cells. *Proc Natl Acad Sci U S A* **97,** 5995-6000.

Hajdukiewicz, P., Svab, Z. and Maliga, P. (1994) The small, versatile pPZP family of Agrobacterium binary vectors for plant transformation. *Plant Mol Biol* **25,** 989-994.

Hajdukiewicz, P.T.J., Gilbertson, L. and Staub, J.M. (2001) Multiple pathways for Cre/*lox*-mediated recombination in plastids. *Plant J* **27,** 161-170.

Kuroda, H. and Maliga, P. (2001) Sequences downstream of the translation initiation codon are important determinants of translation efficiency in chloroplasts. *Plant Physiol* **125,** 430-436.

Kuroda, H. and Maliga, P. (2003) The plastid *clpP1* gene is essential for plant development. *Nature* **425,** 86-89.

Lutz, K.A., Knapp, J.E. and Maliga, P. (2001) Expression of *bar* in the plastid genome confers herbicide resistance. *Plant Physiol* **125,** 1585-1590.

Maliga, P. (2003) Progress towards commercialization of plastid transformation technology. *Trends Biotechnol* **21,** 20-28.

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

Murashige, T. and Skoog, F. (1962) A revised medium for the growth and bioassay with tobacco tissue culture. *Physiol Plant* **15,** 473-497.

Ow, D.W. (2002) Recombinase-directed plant transformation for the post-genomic era. *Plant Mol Biol* **48,** 183-200.

Svab, Z. and Maliga, P. (1993) High-frequency plastid transformation in tobacco by selection for a chimeric *aadA* gene. *Proc Natl Acad Sci U S A* **90,** 913-917.

Thomason, L.C., Calendar, R. and Ow, D.W. (2001) Gene insertion and replacement in *Schizosaccharomyces pombe* mediated by the *Streptomyces* bacteriophage phiC31 site-specific recombination system. *Mol Genet Genomics* **265,** 1031-1038.

Thyagarajan, B. (2001) Site-specific genomic integration in mammalian cells mediated by phage phiC31 integrase. *Mol Cell Biol* **21,** 3926-3934.

Timko, M.P., Kaush, A.P., Hand, J.M. and Cashmore, A.R. (1985) Structure and expression of nuclear genes encoding polypeptides of the photosynthetic apparatus. In *Molecular biology of the photosynthetic apparatus*. (Steinback, K.E., Bonitz, S., Arntzen, C.J. and Bogorad, L., eds). Cold Spring Harbor: Cold Spring Harbor Laboratory, pp. 381-396.

Zoubenko, O.V., Allison, L.A., Svab, Z. and Maliga, P. (1994) Efficient targeting of foreign genes into the tobacco plastid genome. *Nucleic Acids Res* **22,** 3819-3824.

Author Contributions to research described in Appendix 1

Kerry Lutz, Dr. Sylvie Corneille, Dr. Zora Svab and Prof. Pal Maliga designed research; construction of INT lines and transformation vectors were carried out by KL and SC; transformation with the INT-mediated system was carried out by AA and ZS; KL, SC, AA, ZS and PM analysed data, KL, AA and PM wrote the paper.

Appendix 2

CRE-mediated recombination between plastid *lox* DNA sequences and *loxP* sites in the plastid genome

The objective was to test whether or not CRE, the P1 phage site-specific recombinase, induces genome rearrangements in plastids. Testing was carried out in tobacco plants in which a DNA sequence, located between two inversely oriented *loxP* sites, underwent repeated cycles of inversions as a means of monitoring CRE activity. We report here that CRE mediates deletions between *loxP* sites and plastid DNA sequences (*p-lox*) in the 3'*rps12* gene leader (*lox-rps12*) or in the *psbA* promoter core (*lox-psbA*). We also observed deletions between two directly-oriented *lox-psbA* sites, but not between *lox-rps12* sites. Deletion *via* duplicated rRNA operon promoter (Prrn) sequences was also frequent in CRE-active plants. However, CRE-mediated recombination is probably not directly involved, as no recombination junction between *loxP* and Prrn could be observed. Tobacco plants carrying deleted genomes as a minor fraction of the plastid genome population were fertile and phenotypically normal suggesting that the absence of deleted genome segments was compensated by gene expression from wild-type copies. The deleted plastid genomes disappeared in the seed progeny lacking CRE. Observed plastid genome rearrangements are specific to engineered plastid genomes, which contain at least one *loxP* site or duplicated *psbA* promoter sequences. The wild-type plastid genome is expected to be stable, even if CRE is present in the plastid. In addition, these *p-lox* can be potential alternate targets for inserting transgenes.

Results

Experimental design

To test for CRE-induced genome rearrangements, plants were created in which a DNA sequence, located between two inversely oriented *loxP* sites, underwent repeated cycles of inversions. The DNA sequence undergoing inversions was the *neo* coding region, encoding neomycin phosphotransferase (NPTII) that confers kanamycin resistance. The *neo* coding region, in the transformation vector (pSAC38) is in an inverted orientation relative to the promoter (>*neo*<; ><symbolize *loxP* sites) (Figure A2.1a). Introduction of CRE into pSAC38-transformed plastids was to trigger inversion of the >*neo*< coding region enabling transcription of the *neo* sense strand by the Prrn promoter. The system was designed to allow monitoring of CRE activity by repeated cycles of >*neo*< inversions between two inversely oriented *loxP* sites, that maintains a plastid genome population with both *neo* orientations. Incorporation of the >*neo*< transgene and selectable marker *aadA* in the inverted repeats increases the plastid genome size from 155.9 kb to 160.7 kb. Transformation of wild type tobacco leaves with plasmid pSAC38 yielded two independently transformed lines: Nt-pSAC38-10C and Nt-pSAC38-9A. All experiments were carried out with the Nt-pSAC38-10C line, which will be referred to as line Nt-pSAC38.

Monitoring CRE activity by NPTII expression

The *cre* gene was introduced into the nuclear genome of Nt-pSAC38 plants to trigger inversion of the >*neo*< gene. To ensure plastid targeting of the CRE recombinase, the *cre* coding region was translationally fused with the Rubisco small subunit (SSU) transit peptide: one with 22 (*Cre3* gene) and one with 5 (*Cre4* gene) amino acids of the mature SSU N-terminus. The *cre* genes were expressed in P2' promoter and Tnos terminator cassettes. In the *Agrobacterium* binary vectors the nuclear *cre* gene is linked to the gentamycin resistance marker that allows for tracking of *cre*.

Cocultivation of Nt-pSAC38 leaves with *Agrobacterium* resulted in nine independently transformed lines. DNA gel blot analysis confirmed CRE activity in two clones (Cre3-5 and Cre4-6) shown by a mixed

population of expressed (4.6-kb and 1.0-kb fragments) and non-expressed (3.8-kb and 1.8-kb fragments) >*neo*< genes (Figure A2.1d). In two clones, Cre3-10 and Cre4-10, only two hybridizing fragments (3.8-kb, 1.8-kb) were present. This indicates that *cre*, although introduced (tested by PCR; data not shown), was not expressed. These clones have been excluded from further studies. In five clones, Cre3-4, Cre4-1, Cre4-2, Cre4-5 and Cre4-7, a smaller than wild-type fragment was found (3.0-kb instead of 3.3-kb). The plastid genome of these clones lacks the *trnV-aadA-neo* region (Figure A2.1c) which was deleted by recombination between the native rRNA operon promoter and its derivative (Prrn) driving expression of the >*neo*< gene, as confirmed by sequencing PCR amplification products. Deletion of plastid DNA sequences between Prrn repeats in CRE-active plants has been reported earlier (Corneille et al., 2001; Hajdukiewicz et al., 2001).

The plastid-targeted *cre* was also introduced into the Nt-pSAC38 plants by pollination using the previously characterized strong activator lines Cre1-3 and Cre1-100 (Corneille et al., 2001). DNA gel blot analysis in the segregating seed progeny identified seedlings in which >*neo*< was present in both orientations (seedlings 2, 3, 27 and 29; Figure A2.2). An additional, minor 3.0-kb fragment is also visible in seedling No. 3. This fragment is the product of *trnV-aadA-neo* deletion.

Immunoblot analysis of leaf samples was carried out to test for NPTII accumulation in CRE-activated plastids. Levels of NPTII were about 1% of the total soluble cellular protein, about half of the amount found in Nt-pHK31 positive control plants (Maliga et al., 2002) (Figure A2.3). Thus, incorporation of a *lox* site in the 5'-UTR is compatible with significant levels of NPTII expression.

Testing for CRE-mediated recombination between loxP *and plastid sequences*

CRE-mediated recombination between *loxP* sites and pseudo-*loxP* sites have been observed in yeast (Sauer, 1992, 1996) and mammalian (Thyagarajan et al., 2000) genomes. Characteristic for CRE-mediated recombination is the junction sequence, which consists of a perfect 13-bp palindromic CRE *loxP* binding site, a recombination junction within the spacer, and the host genome sequence with some similarity to the 13-bp *loxP* palindrome (Sauer, 1992). Unpredicted deletions have also been observed in the plastid genome of plants expressing CRE (Corneille et al., 2001; Hajdukiewicz et al., 2001). To determine if deletions in

the plastid genome were mediated by CRE, we have sequenced PCR amplified recombination junctions. To facilitate communication of the results, we have individually labeled sequences involved in recombination: four transgenic *loxP* sites, *loxP1* and *loxP2* in inverted repeat A (IRA) and *loxP3* and *loxP4* in IRB; three *lox-psbA* sites, *lox-psbA1* in the *psbA* gene promoter, *lox-psbA2* and *lox-psbA3* driving *aadA* in IRA and IRB, respectively; and two *lox-rps12* sites located in the *rps12* leader (*lox-rps12-1,* IRA; *lox-rps12-2,* IRB) (Figure A2.4c).

A recombination junction between the *psbA2* promoter and the *loxP1* sequence was identified by sequencing the PCR fragment generated by primers O2 and O3 (Figure A2.4a). Note, that the sequence of IRA and IRB is identical, thus identical PCR products are generated from both repeats. In cases such as this we shall discuss only data for IRA. The recombination junction consists of the *psbA* promoter sequence (*lox-psbA2*), most of the *loxP1* spacer and 13-bp of the *loxP1* CRE binding site (Figure A2.4a). Encouraged by finding CRE-mediated recombination between *loxP1* and *lox-psbA2*, we also tested recombination between *loxP1* and the native *psbA* promoter (*lox-psbA1*). Sequencing of the PCR products (primers O4 and O5) confirmed CRE-mediated recombination between *loxP1* and the native *psbA* promoter (*lox-psbA1*) (Figure A2.4b) indicating a 17.3-kb deletion in the plastid genome. This genome segment encodes a total of 10 genes. Deletions were also tested between the directly oriented *lox-psbA2* site in IRA and the *lox-psbA1* site at its native location in the large unique region (Figure A2.4c). Sequencing of PCR-amplified products confirmed deletion of sequences between *lox-psbA2* and *lox-psbA1* sequences (primers: O3 and O5; 18.7-kb deletion yielding a 142-kb plastid genome) and *lox-psbA1* and *lox-psbA2* sequences (primers: O6 and O7; 142-kb deletion yielding an 18.7-kb plastid genome), the two deletions combined encompass the entire 160.7-kb transplastome.

We also inspected potential recombination junctions between the *loxP1* site and the previously described recombination 'hot-spot' in the 3'*rps12*leader sequence (Hajdukiewicz et al., 2001). PCR amplification using primers O4 and O8 yielded a fragment with a characteristic CRE-mediated junction sequence shown in Figure A2.5. Sequencing of PCR-amplified products from three different plants yielded an identical junction site. Therefore, the former "hot-spot' sequence was renamed *lox-rps12-1*.

Plastid deletion products, in the absence of direct selection, accumulate in readily detectable amounts only if non-essential genes have been excised. Fortunately, some of the deletion products, between

loxP1 and *lox-psbA2* or *lox-rps12-1* and *loxP1,* are in this category. Figure A2.6 shows a Southern blot verifying CRE-mediated deletion between the *loxP1* and *lox-psbA2* (Cre4-2 line) and *lox-rps12-1* and *loxP1* (Cre4-1) sites. Interestingly, these deletion products were enriched in plastids in which the t*rnV-aadA-neo* region has been deleted by homologous recombination via Prrn sequences.

We have observed deletion of a plastid genome segment by recombination between directly repeated Prrn promoters in Nt-pSAC38 plants containing an active CRE (Figure A2.1c). Deletion of plastid DNA sequences between directly repeated Prrn promoters has been reported earlier (Corneille et al., 2001; Hajdukiewicz et al., 2001). To determine if these deletions were mediated by CRE, potential recombination junctions were PCR-amplified between the *loxP* sites in combination with all Prrn promoter sequences (one native *rrn* promoter and one *neo* transgene Prrn promoter in each repeat). No evidence was found for CRE mediated recombination between *loxP* sites and Prrn sequences.

Testing recombination between loxP and plastid lox sequences identified by computer search

Computer searches were performed to identify sequences that may recombine with the *loxP* sites in the tobacco plastid genome using the findpatterns algorithms of the Wisconsin Software Package. Candidate sequences with similarity to the 34-bp *loxP* sequence were three sequences with a 22-bp match; five sequences with a 21-bp match and 23 sequences with a 20-bp match. The 31 sites were experimentally tested by PCR for recombination with an appropriately oriented *loxP* site. No recombination junctions were found between any of the plastid sequences and a *loxP* site (data not shown).

A search was also performed to identify the plastid sites with the best match to structurally important *loxP* nucleotides (query sequence: ATnACnnCnTATAnnnTAnnnTATAnGnnGTnAT) (Hoess et al., 1982; Hoess et al., 1986; Guo et al., 1997; Thyagarajan et al., 2000). Out of the seven best matches with the structurally important nucleotides two were already identified in the group of 31 sequences tested above. No recombination junctions were found between the remaining five best matches and the nearest properly oriented *loxP* site (data not shown).

Testing CRE-mediated deletion in the absence of perpetual DNA inversions

We also tested CRE-mediated deletions between *loxP* and plastid *lox* sites in the absence of perpetual >*neo*< inversions. The test plants carried a *loxP* site and a l*ox-psbA* site in each of the inverted repeats (Figure A2.7). The plants are seed progeny of >*codA*> deletion derivatives and segregated for the nuclear *cre* (Corneille et al., 2001). Out of the nine seedlings shown in Figure A2.7 eight contain a *cre*. PCR amplification products indicating recombination between *lox*P1 and *lox-psb*A1 and between *lox-psb*A1 and *lox-psb*A2 were obtained in each of the eight *cre*-containing seedlings, whereas the recombination products were absent in the seedling lacking *cre*.

Discussion

CRE activity yields large deletions in the plastid genome

We report here CRE-mediated recombination between *loxP* and plastid DNA sequences that yielded large deletions in the 160.7-kb transgenic tobacco plastid genome. The test system, a >*neo*< gene flanked by two inverted *loxP* sites, allowed continuous monitoring of CRE activity ensuring that *cre* is not silenced.
We found two plastid sequences, *lox-psbA* and *lox-rps12*, which are recognized by CRE as substrates since characteristic CRE-mediated recombination junctions between the plastid *lox* sites and *loxP* sites. Recombination was also observed between two directly oriented l*ox-psbA* sites duplicated in the transformed plastid genome. CRE-mediated recombination *via lox-psbA* is reported in this study. Hajdukiewicz et al. previously reported recombination between *loxP* and sequences in the *3'rps12* 5'-UTR (Hajdukiewicz et al., 2001). Recombination junctions in three independent clones were located throughout a 34-bp region. Therefore, Hajdukiewicz et al. called this region a 'recombination hot-spot'. In our study the *loxP* sequence is in an opposite orientation relative to the *loxP* site described by Hajdukiewicz et al. (Hajdukiewicz et al., 2001), and the same recombination junction was identified in three PCR-amplified *lox-rps12* recombination products (Figure A2.5). The recombination junctions mapped to one of the three sites identified earlier. The same *rps12* sequence yielding deletions with *loxP* in either orientation may be explained by its degeneracy relative to the *loxP* consensus (17 matches out of 34 nucleotides).
Figure A2.8 depicts CRE-mediated recombination products of the master (monomeric) plastid genome, which are inferred from sequencing recombination junctions. Shown in the center is the Nt-pSAC38 transformed plastid genome. Two circles generated by recombination between the *lox-psbA1* and *lox-psbA2* sites are shown in the upper right corner. We also show deletions due to recombination between *lox-psbA1*/*loxP1* (lower right), *loxP1*/*lox-psbA2* (or *loxP4*/*lox-psbA3*) (lower left) and *lox-rps12-1*/*loxP1* (or *lox-rps12-2*/*loxP4*) sites (top left). Deletions are likely to involve both copies (IRA, IRB) of the repeated region as the sequence of the repeats is base by base identical (Ohyama et al., 1986; Shinozaki et al., 1986), and transformation of the inverted repeat region involves both copies (Svab et al., 1990). Plastid genes are

known to exist as a multimeric series (Kolodner and Tewari, 1979; Deng et al., 1989; Lilly et al., 2001); deletion derivatives of monomeric and head-to-tail multimeric genome copies yield identical PCR products. We considered the possibility that the observed PCR products may have formed by template switching during the PCR reaction (Bradley and Hillis, 1997). We believe, this is not the case since: (i) deletion products involving *lox* sites occur only CRE-active plants; (ii) the site of recombination involving *loxP* sequences is in the spacer region and it is specific to CRE; (iii) products are also detectable by DNA gel blot analysis.

It should be noted that the observed plastid genome rearrangements are specific to engineered plastid genomes because the wild-type plastid genome does not contain *lox*P sites or duplicated *psbA* promoter sequences. Furthermore, no CRE-mediated recombination was found between the *lox-psbA1* and *lox-rps12-1* sites present in wild type plastid genomes.

Plastid lox *sites may not be identified by similarity to* loxP

There are many degenerate *loxP*-like sequences in the plastid genome, which are not substrates for CRE-mediated recombination. Interestingly, the two plastid sites that are CRE substrates were not among the 36 best matches identified by the computer search. The *lox-psbA* sequence (as well as the *lox-rps12* sequence) has only 17 matches out of the 34 *loxP* nucleotides. It is likely therefore that the overlap of *lox-psbA* with the conserved *psbA* promoter core is important for this site being a substrate for CRE, possibly through the RNA polymerase making accessible this site for interaction with the recombinase. The *lox-psbA* site is conserved in the plastids of monocots and dicots (Figure A2.9). Thus, CRE-mediated insertion at the *lox-psbA* site may facilitate plastid transformation in recalcitrant species.

Elevated homologous recombination in plastids containing CRE

Frequent deletion products in CRE-active plants were obtained by homologous recombination between two directly oriented Prrn sequences. This deletion product has been described in *Agrobacterium*-transformed plants (Corneille et al., 2001; Hajdukiewicz et al., 2001). We did not observe these deletion products in

young seedlings in our earlier study (Corneille et al., 2001). However, such deletion products were present in some of the older seedlings in the present study in which CRE was expressed from a *cre* gene obtained by an independent insertion event. Thus, formation of the deletions products is not specific to *Agrobacterium*-transformed cells and may also be observed in seedlings dependent on age and CRE expression levels. Deletions could be the result of CRE-mediated recombination *via* plastid *lox*-sites, or mediated by the plastid's homologous recombination machinery. To distinguish between the two possibilities, PCR amplification of deletion products was attempted between *loxP* sites and *Prrn* sequences. No such deletion product could be amplified. Therefore, we believe that deletions between *Prrn* sequences are mediated by the plastid's homologous recombination machinery. Interestingly, in the absence of CRE, such deletion products have never been seen, and the transformed plastid genomes are generally stable. Thus, enhanced deletions *via* homologous recombination is likely the result of CRE interacting with the plastid recombination machinery through its known DNA-binding and DNA-bending properties (Guo et al., 1997; Gopaul et al., 1998; Van Duyne, 2001).

Large deletions in the polyploid plastid genetic system have no long-term penalty

Expression of CRE in both plants and animals lead to deleterious, unpredicted mutations in the nuclear genome, causing phenotypic abnormalities, sterility, and loss of ability to regenerate plants (Schmidt et al., 2000; Loonstra et al., 2001; Coppoolse et al., 2003). In the present study we report CRE-mediated deletion of large plastid genome segments. Mutations in non-essential tobacco plastid genes, such as genes encoding photosynthetic function, can be readily obtained even if they cause a mutant phenotype (Kanevski and Maliga, 1994; Allison et al., 1996); reviewed in (Bock and Hippler, 2002). Unlike mutations induced by CRE in the nuclear genome, large deletions of the plastid genome did not lead to an obvious phenotype. Lack of a readily detectable phenotype can be explained by the deleted genome copies being only a minor fraction of the plastid genome population, and by compensation for the lacking function by genes expressed from wild-type genome copies. Deleted plastid genome copies lacking essential genes are rapidly eliminated in the absence of selection pressure (Svab and Maliga, 1993; Drescher et al., 2000; Shikanai et al., 2001), or a mechanism, such as CRE action, that continuously recreates the mutation.

Experimental procedures

Vector construction

The inverted *neo* gene was contained in a *Sac*I-*Hin*dIII fragment. The plastid rRNA operon (*rrn*) promoter derivative was contained in a *Sac*I-*Xba*I fragment obtained by PCR using oligonucleotides 5'-GGGGAGCTCGCTCCCCCGCCGTCGTTCAATG-3' and 5'-GGGAATTCATAACTTCGTATAGCATACATTATACGAAGTTATGCTCCCAGAAATATAGCCA-3' as primers and plasmid pZS176 (progenitor of plasmid pZS197) (Svab and Maliga, 1993) as a template. The promoter fragment Prrn*lox*I contained a *loxP* site at the 3' end adjacent to the *Xba*I site. The *neo* coding region was contained in an *Xba*I (at the 3'-end) and *Nco*I fragment. The ribosome binding site from plasmid pZS176 was contained in an *Nco*I-*Eco*RI fragment. The TrbcL*lox*I is the *rbcL* 3'-untranslated region contained in an *Eco*RI-*Hin*dIII fragment obtained by PCR using oligonucleotides 5'-GGGGAATTCATAACTTCGTATAGCATACATTATACGAAGTTATAGACATTAGC-3' and 5'-GGGAAGCTTGCTAGATTTTGTATTTCAAATCTTG-3' as primers and plasmid pMSK48 (Khan and Maliga, 1999) as template. TrbcL*lox*I contained a *loxP* site adjacent to the *Eco*RI site in indirect orientation relative to the *loxP* site in the *neo* 5'UTR. The chimeric *Prrnlox*I:*neo*:*TrbcLlox*I gene was introduced into the tobacco plastid transformation vector pPRV111B (Zoubenko et al., 1994) as a *Sac*I-*Hin*dIII fragment to obtain plasmid pSAC38.

Two plastid targeted nuclear *cre* genes were tested. The *cre* genes in *Agrobacterium* binary vectors pKO30 and pKO31 encode the CRE recombinase. Two chimeric *cre* genes were prepared, one with 5 (*Cre4* gene; plasmid pKO31), the other with 22 (*Cre3* gene; plasmid pKO30) amino acids of the mature pea Rubisco small subunit (SSU) translationally fused with CRE at the N terminus (Timko et al., 1985). Both *cre* genes were described in (Corneille et al., 2001). The plastid targeted nuclear *cre* genes were introduced as *Eco*RI-*Hin*dIII fragments into the pPZP222 *Agrobacterium* binary vector (Hajdukiewicz et al., 1994) to obtain plasmids pKO30 and pKO31 with twenty two and five amino acids of the mature Rubisco SSU respectively.

Plant transformation

Plastid transformation using the biolistic protocol, selection of transplastomic tobacco lines (RMOP medium, 500 mg/L spectinomycin dihydrochloride) and characterization of the transplastomic lines by DNA gel blot analysis was performed as described (Svab and Maliga, 1993). Transformation with *Agrobacterium* vectors pKO30 or pKO31 and regeneration of transformed tobacco plants has also been reported (Hajdukiewicz et al., 1994). Briefly, nuclear gene transformants were selected by gentamycin resistance on RMOP shoot regeneration medium containing 100 mg/L gentamycin and 500 mg/L carbenicillin. Gentamycin resistance of the shoots was confirmed by rooting on plant maintenance (RM) medium containing 100 mg/L gentamycin. Kanamycin resistance level of *cre* transformed plants was tested on RMOP plant regeneration medium containing 50 mg/L kanamycin. DNA gel blot analysis was carried out on *Bam*HI digested total leaf cellular DNA and probed with the targeting region (1.9-kb *Apa*I-*Eco*RV fragment containing the *rrn16* gene), the *aadA* and *neo* coding regions were isolated as *Nco*I-*Xba*I fragments (Svab and Maliga, 1993).

PCR amplification of plastid DNA and sequencing of recombination junctions

DNA template for PCR amplification was obtained from isolated chloroplasts (Guedeney et al., 1996). Plastid genome segments were identified using the following oligonucleotides: primer 01, 5'-CCGCCAGCGTTCATCCTGAGC-3'; primer 02, 5'-GAGATGTAACTCCAGTTCC-3'; primer 03, 5'-CGCTCGATGACGCCAACTACC-3'; primer 04, 5'-TGAAGAGCTTGGCGGCGAAT-3'; primer 05, 5'-GCAACCCACTAGCATATCGAA-3'; primer 06, 5'-GCATCAGAGCAGCCGATTGT-3'; primer 07, 5'-CAGAAGTTGCCGTCAATAAGG-3'; primer O8, 5'-TGAAAGAGGTTGACCTCCTTG-3'. PCR amplification was carried out with the Amplitaq polymerase (Applied Biosystems, Foster City, CA). The PCR program was: 1 cycle, 94 ^{o}C, 1 min; 30 cycles, 94 ^{o}C, 30 sec; 55 ^{o}C, 45 sec; 72 ^{o}C, 1.5 min; 1 cycle, 72 ^{o}C, 7 min. DNA template for sequencing was isolated from gels and purified on Qiaquick gel extraction columns (Qiagen, Valencia, CA). DNA template was sequenced with the USB (Cleveland, OH) T7 Sequenase v. 2.0 DNA sequencing kit according to the manufacturer's protocol.

Immunoblot analysis

Total soluble leaf protein was extracted from plants grown in sterile culture. About 200 mg of leaf was homogenized in 1 mL of buffer containing 50 mM Hepes/KOH (pH 7.5), 1 mM EDTA, 10 mM potassium acetate, 5 mM magnesium acetate, 10 mM dithiothreitol and 2 mM phenylmethanesulfonyl fluoride. Protein concentration was determined by the Bradford Protein Assay reagent kit (Bio-Rad, Hercules, CA). NPTII was detected by immunoblot analysis (Carrer et al., 1993) and quantified by densitometry (DensoSpot program of Alpha Imager 2000; Alpha Innotech, San Leandro, CA) using a dilution series of commercial NPTII as control (5Prime→3Prime, Inc., Boulder, CO). Positive controls were leaf extracts of Nt-pHK31 plants (Maliga et al., 2002).

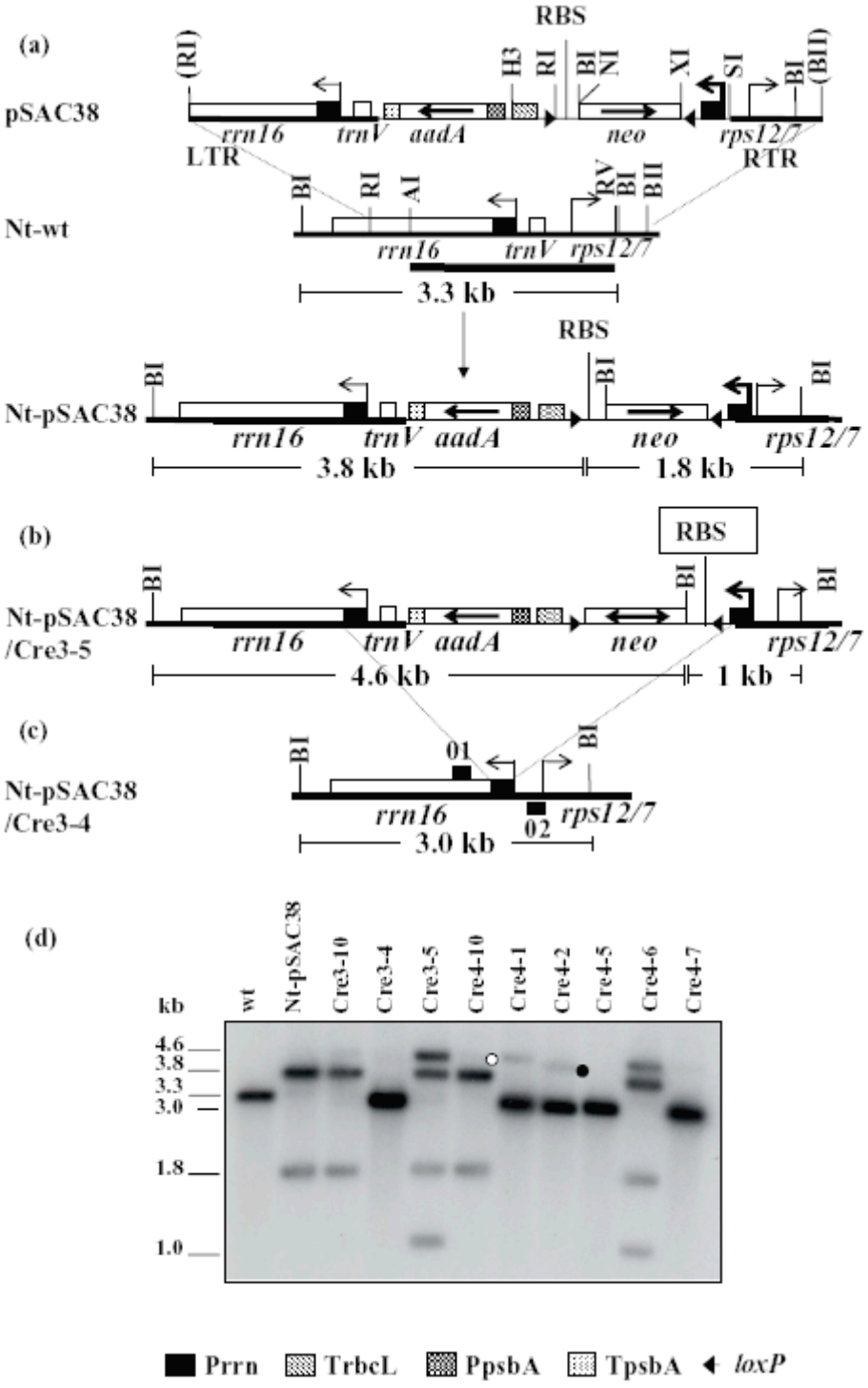


Figure A2.1. CRE-mediated plastid genome rearrangements in Nt-pSAC38 plants.
(a) Map of the plastid targeting region of transformation vector pSAC38; the cognate region of the wild type plastid genome (Nt-wt); and map of the plastid genome after transformation with plasmid pSAC38 (Nt-pSAC38). Positions of plastid genes *rrn16*, *trnV* and 3'*rps12/rps7*, *aadA* and >*neo*< transgenes, left targeting region (LTR) and right targeting region (RTR), relevant restriction sites (AI, *Apa*I; BI, *BamH*I;

BII, *Bgl*II; RI, *Eco*RI; RV, *Eco*RV; H3, *Hin*dIII; NI, *Nco*I; SI, *Sac*I; XI, *Xba*I) and *loxP* sites (filled triangles) are marked.
(b) Map of Nt-pSAC38 plastid genome with >*neo*< gene expressed after transformation with plastid targeted nuclear *cre* (line Nt-pSAC38/Cre3-5).
(c) Map of Nt-pSAC38 plastid genome with deletion of the *trnV-aadA-neo* segment by recombination *via* repeated Prrn sequences. Deletion was observed after *Agrobacterium*-mediated transformation with a nuclear-encoded, plastid targeted *cre.* This structure was present in lines Cre3-4, Cre4-1, Cre4-2 Cre4-5 and Cre4-7 shown in Figure A2.1d.
(d) DNA gel blot analysis confirms CRE activity by detecting mixed populations of expressed and non-expressed >*neo*< copies in lines Cre3-5 and Cre4-6. Shown are also wild-type control (wt), parental control without CRE activity (Nt-pSAC38) and an *Agrobacterium*-transformed plant in which *C*RE is not expressed (Cre3-10; Cre4-10). Map of deletions in lines Cre3-4, Cre4-1, Cre4-2 Cre4-5 and Cre4-7 is shown in Figure A2.1c. Total leaf cellular DNA was digested with the *Bam*HI restriction endonuclease and probed with the targeting region (P1, 1.9-kb *Apa*I-*Eco*RV fragment, heavy bar in Figure A2.1a). Fragment sizes detected by the P1 probe are marked in Figure A21a-c. Fragments marked with open and filled circles in the Cre4-1 and Cre4-2 lines are CRE-mediated deletion products (see Figure A2.6b).

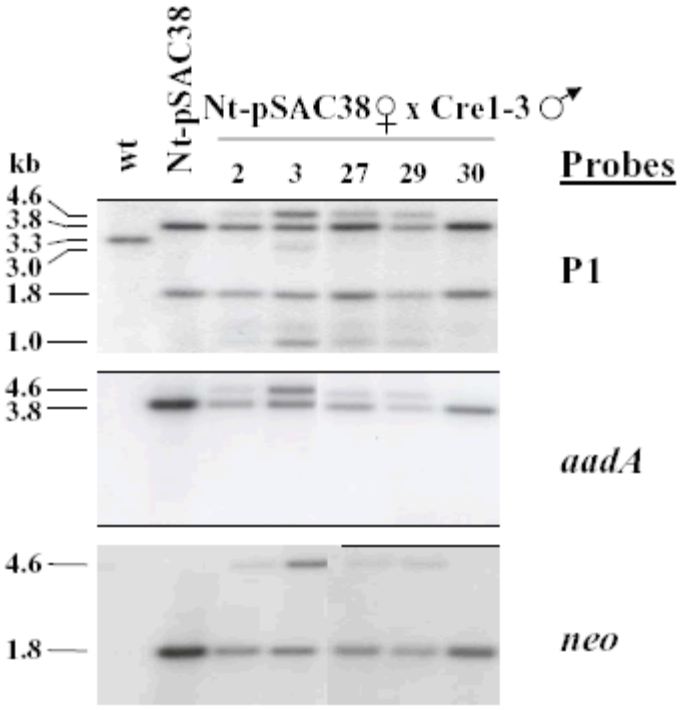


Figure A2.2. DNA gel blot analysis to test for CRE-mediated >*neo*< inversion in the seed progeny of Nt-pSAC38 plants (maternal parent) and Cre1-3 activator line (pollen parent). Total leaf cellular DNA was digested with the *BamH*I restriction endonuclease and hybridized with the plastid targeting region (P), *aadA* and *neo* coding region probes. For explanation of fragment sizes see Figure A2.1a-c.

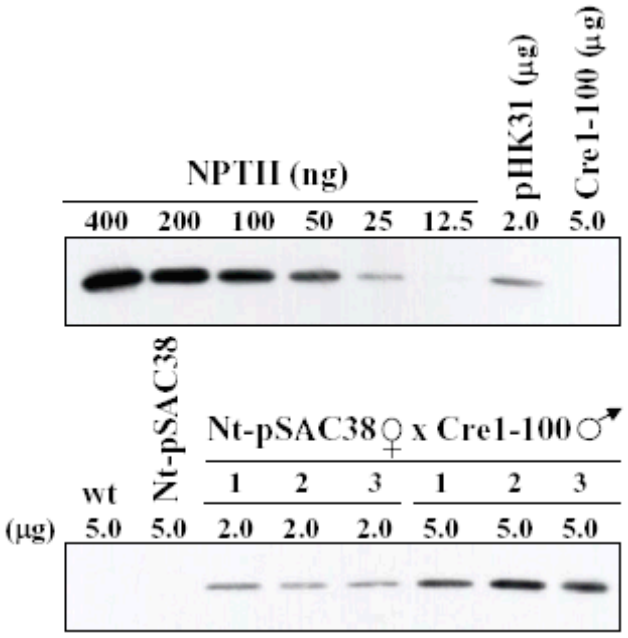


Figure A2.3. Immunoblot analysis for NPTII accumulation in leaves of CRE-activated Nt-pSAC38 plants. Quantification of NPTII is by comparison with commercial NPTII dilution series. Controls are non-transformed tobacco (wt), Nt-pSAC38, Nt-pHK31 and Cre1-100 plants.

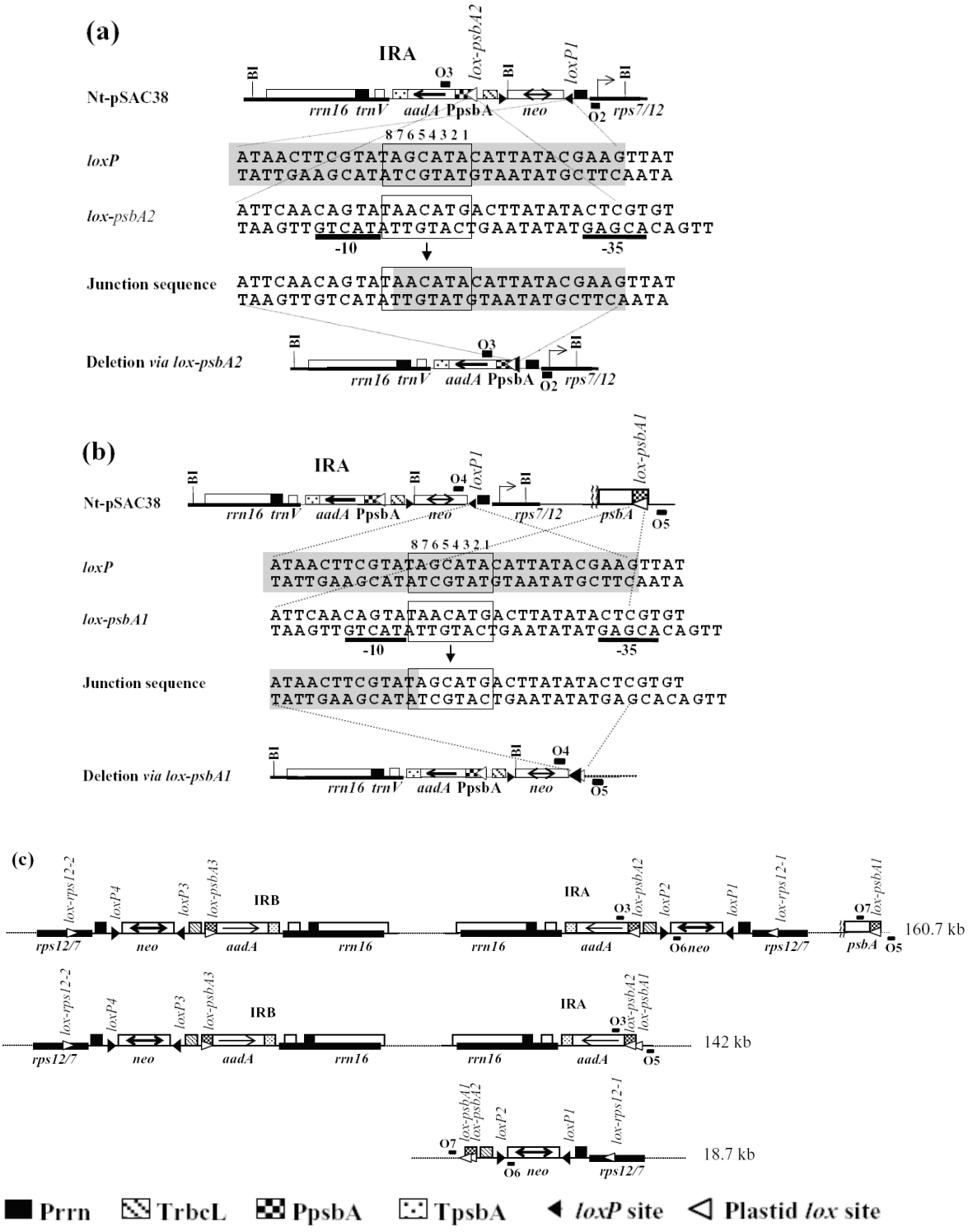


Figure A2.4. CRE-mediated recombination between *loxP*1 and *lox-psbA* and between two *lox-psbA* sequences. (a) Recombination between *loxP1* and *lox-psbA2* sequences.
(b) Recombination between *loxP1* and *lox-psbA1* sequences.

(c) Recombination between *lox-psbA1* and *lox-psbA2* sequences. *loxP* and *lox-psbA* sites are marked with filled and open triangles, respectively. Genes in the maps are labeled as in Figure A2.1. Primers used for PCR amplification are also marked.

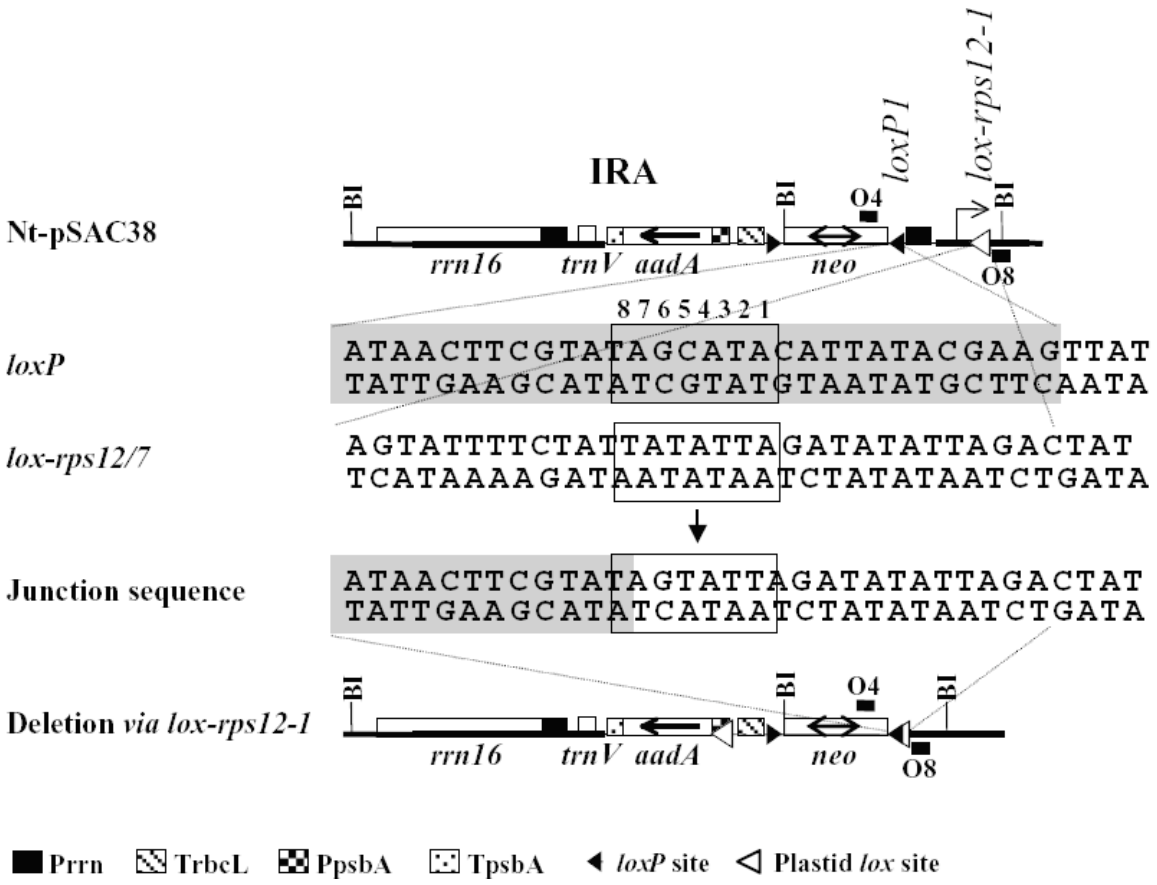


Figure A2.5. CRE-mediated recombination between *loxP*1 and *lox-rps12-1* sequences. *loxP* sites and *lox-rps12-1* sequences are marked with filled and open triangles, respectively. Genes in the maps are labeled as in Figure A2.1. Primers used for PCR amplification are also marked.

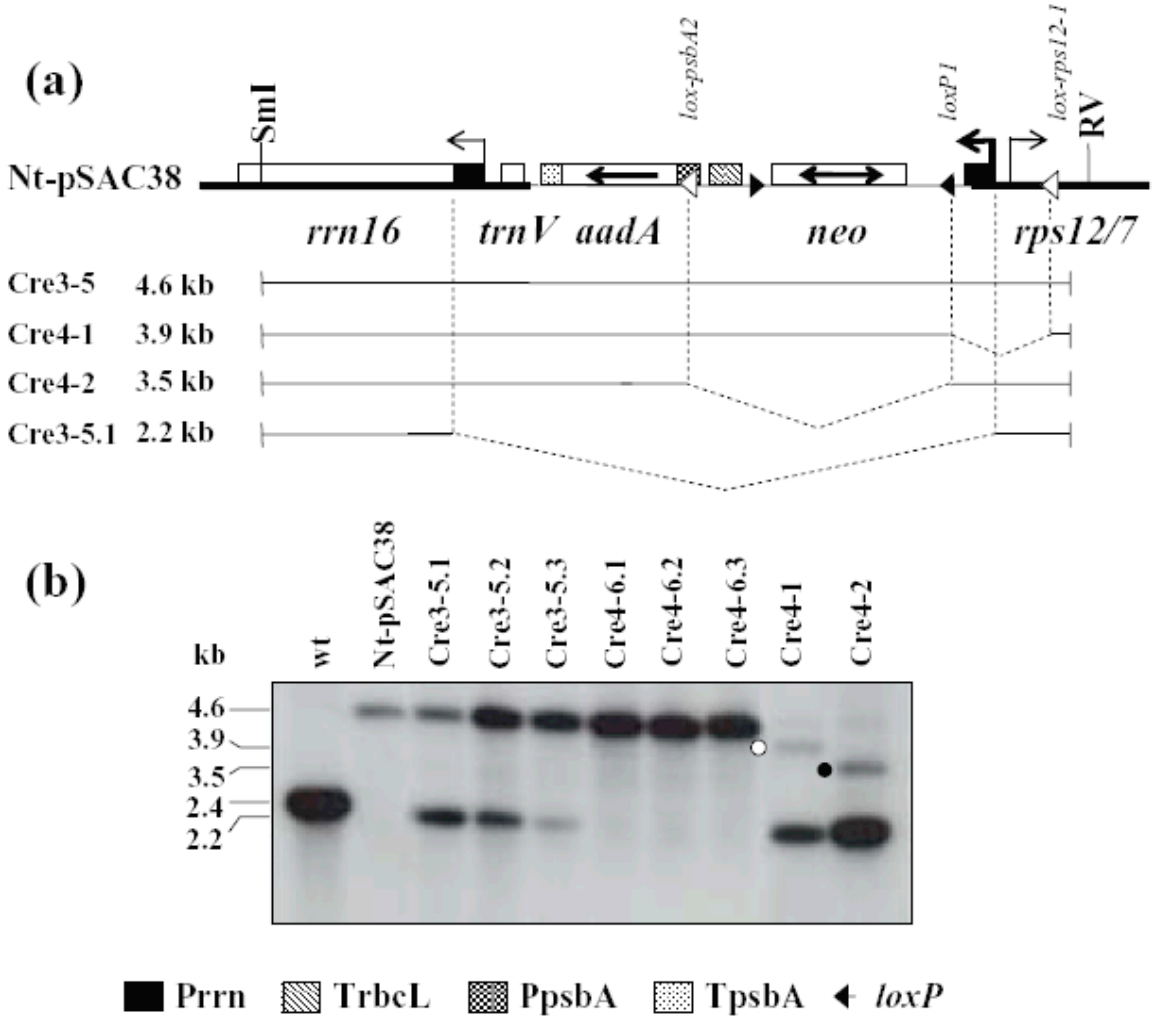


Figure A2.6. Southern analysis confirms CRE-mediated deletion between the *loxP1* and *lox-psbA2* (Cre4-2 line) and *lox-rps12-1* and *loxP1* (Cre4-1) sites. We also show examples for deletion of the *trnV-aadA-neo* region in the Cre3-5 seed progeny (Cre3-5.1, Cre3-5.2, Cre3-5.3), and for uniform transplastomes in the Cre4-6 seed progeny (Cre4-6.1, Cre4-6.2, Cre4-6.3).
(a) Map of Nt-pSAC38 plastid genome and CRE-mediated deletion derivatives. For explanation see caption to Figure A2.1.
(b) Southern blot obtained after digestion of total cellular DNA with the *Sma*I (SmI) and *Eco*RV (RV) restriction endonucleases and probing with the plastid targeting sequence (*Apa*I-*Eco*RV fragment, Figure A2.1a). The *Sma*I-*Eco*RV fragment includes the entire transgenic region. The 3.5-kb *loxP1* and *lox-psbA2* fragment in the Cre4-2 line (filled circle) and the 3.9-kb *lox-rps12-1* and *loxP1* deletion product in the Cre4-1 line (open circle) are marked. Same deletion products after digestion with a different restriction enzyme (*Bam*HI) are marked in Figure A2.1d. The 2.2-kb fragment is the product of deletion between the Prrn promoter driving the *neo* gene and the native *rrn* promoter sequences.

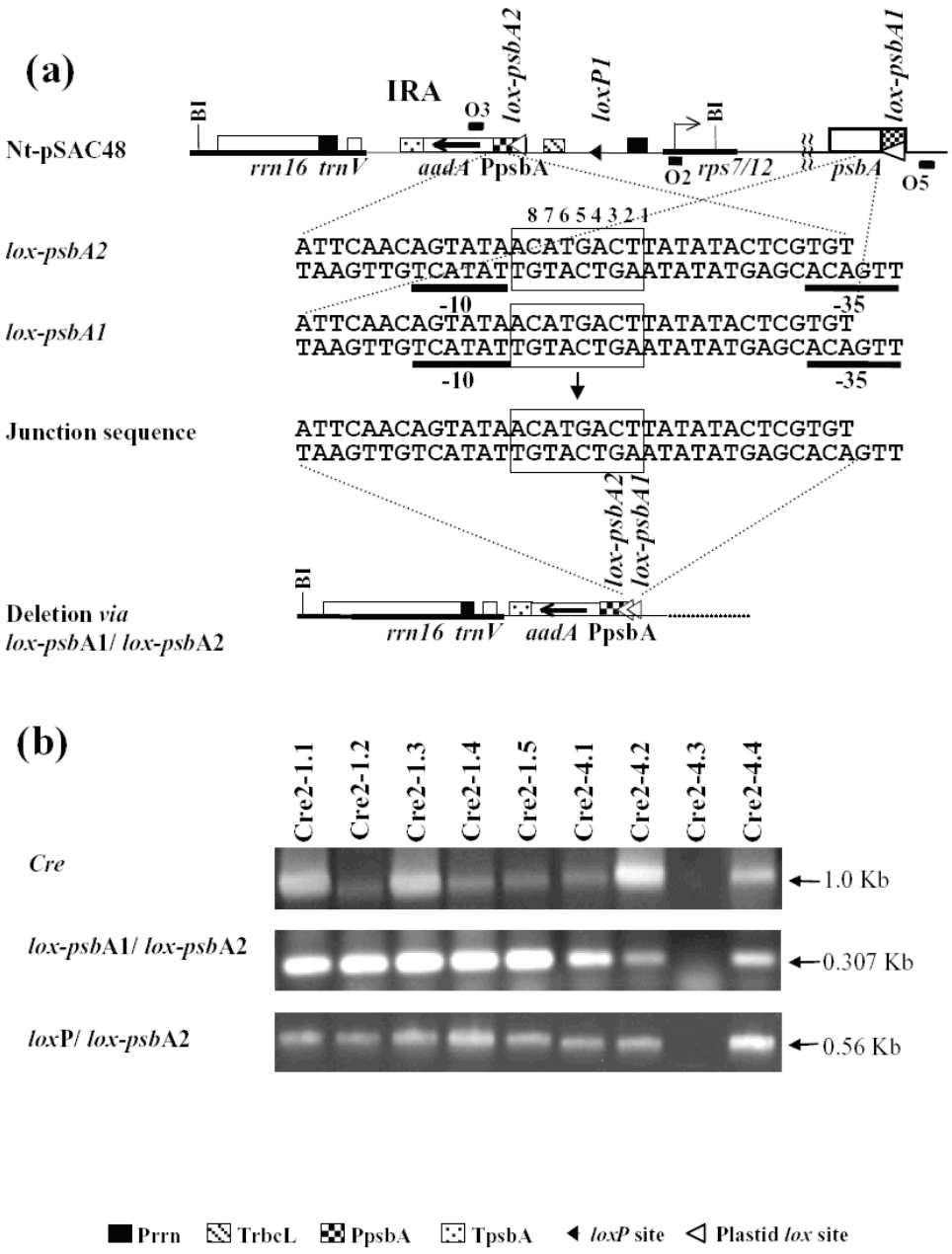


Figure A2.7. CRE-mediated recombination in Nt-pSAC48 >*codA*> deletion derivatives.
(a) Plastid map showing recombination between *lox-psbA1* and *lox-psbA2* sequences and junction sequence. Genes in the maps are labeled as in Figure A2.1. Primers used for PCR amplification are also marked.
(b) *cre*, *lox-psbA1*/*lox-psbA2* and *loxP1*/*lox-psbA2* PCR amplification products.

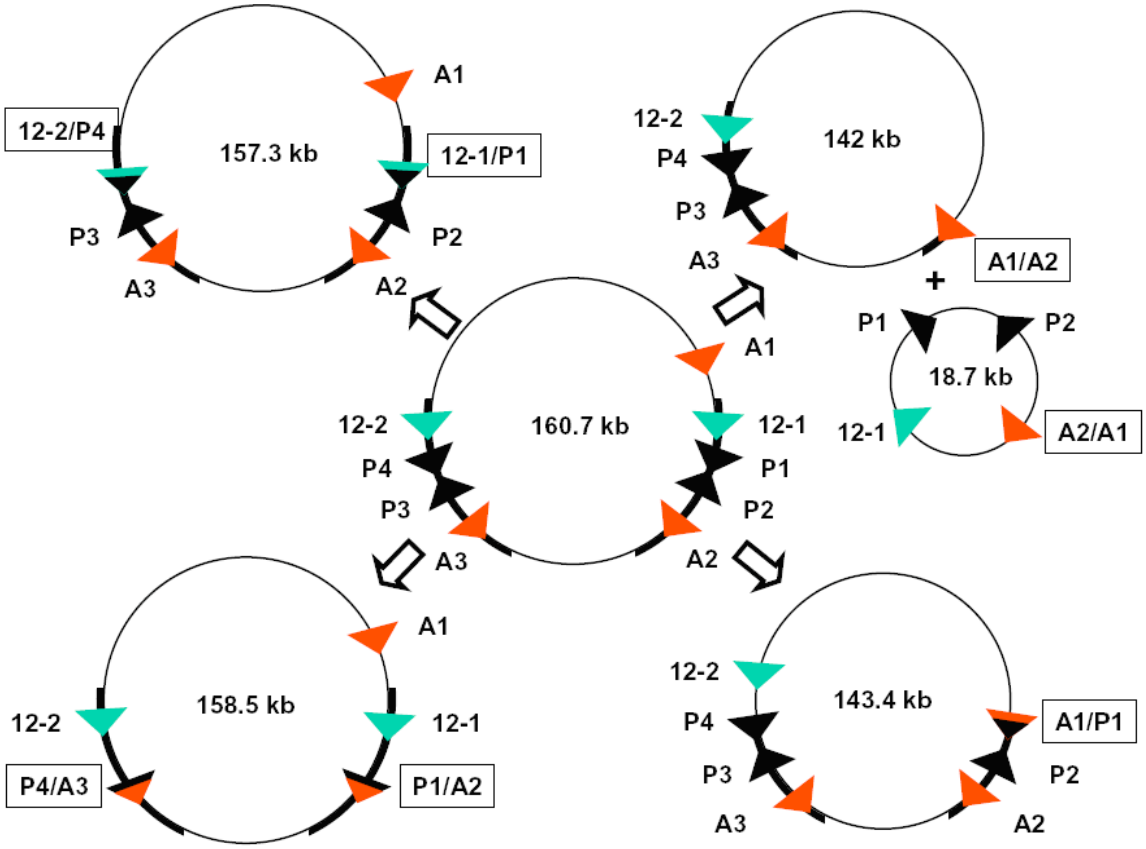


Figure A2.8. CRE-mediated deletion derivatives (size inside circle) of the transformed plastid genome (160.7 kb). Recombination events yielding deletion derivatives are boxed. Abbreviations: A1, *lox-psbA1;* A2, *lox-psbA2;* A3, *lox-psbA3;* P1 through P4, *loxP1* to *loxP4*; 12-1, *lox-rps12-1*; 12-2, *lox-rps12-2*.

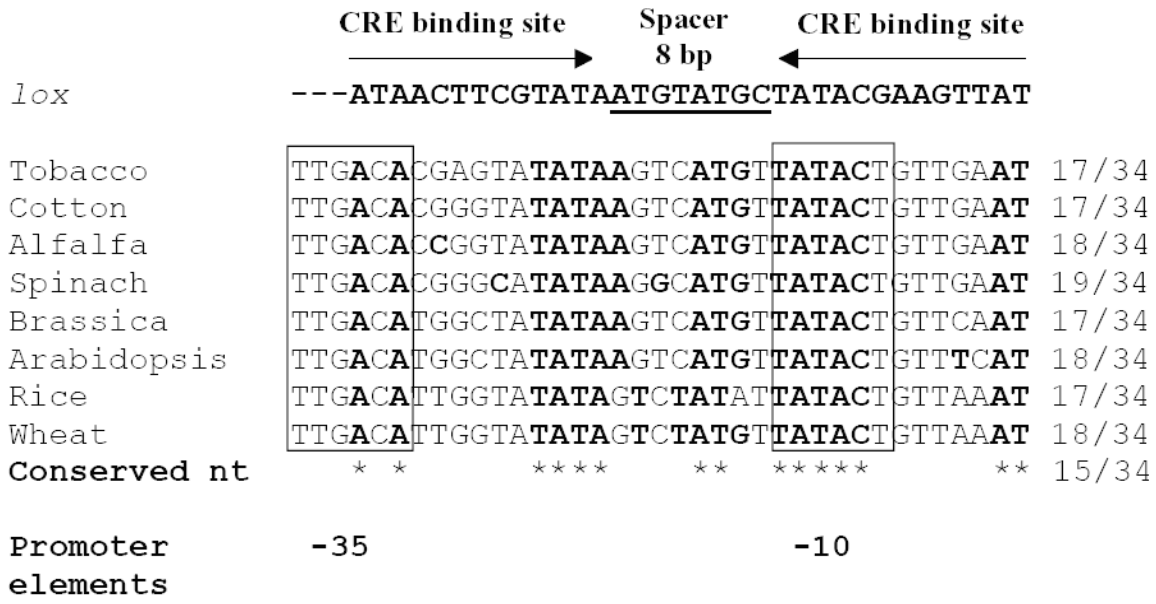


Figure A2.9. Conservation of *lox-psbA* sequence in the *psbA* core promoter. The conserved –35 and –10 core promoter elements are boxed. On top is shown the *loxP* sequence. Nucleotides that match the *loxP* sequence are in bold. Number of matching *loxP* nucleotides out of the 34 is listed on the right. Asterisks at bottom mark *loxP* nucleotides present in all species.

References

Allison, L.A., Simon, L.D. and Maliga, P. (1996) Deletion of *rpoB* reveals a second distinct transcription system in plastids of higher plants. *EMBO J* **15,** 2802-2809.

Bock, R. and Hippler, M. (2002) Extranuclear inheritance: functional genomics in chloroplasts. *Prog Bot* **63,** 106-131.

Bradley, R.D. and Hillis, D.M. (1997) Recombinant DNA sequences generated by PCR amplification. *Mol Biol Evol* **14,** 592-593.

Carrer, H., Hockenberry, T.N., Svab, Z. and Maliga, P. (1993) Kanamycin resistance as a selectable marker for plastid transformation in tobacco. *Mol Gen Genet* **241,** 49-56.

Coppoolse, E.R., de Voormen, M.J., Roelofs, D., Smit, J., van Gennip, F., Hersmus, B.J.M., Nijkamp, H.J.J. and van Haaren, M.J.J. (2003) Cre recombinase expression can result in phenotypic aberration in plants. *Plant Mol Biol* **51,** 263-279.

Corneille, S., Lutz, K., Svab, Z. and Maliga, P. (2001) Efficient elimination of selectable marker genes from the plastid genome by the CRE-*lox* site-specific recombination system. *Plant J* **72,** 171-178.

Deng, X.W., Wing, R.A. and Gruissem, W. (1989) The chloroplast genome exists in multimeric forms. *Proc Natl Acad Sci U S A* **86,** 4156-4160.

Drescher, A., Ruf, S., Calsa, T., Carrer, H. and Bock, R. (2000) The two largest chloroplast genome-encoded open reading frames of higher plants are essential genes. *Plant J* **22,** 97-104.

Gopaul, D.N., Guo, F. and Van Duyne, G.D. (1998) Structure of the Holliday junction intermediate in Cre-*loxP* site-specific recombination. *EMBO J* **17,** 4175-4187.

Guedeney, G., Corneille, S., Cuine, S. and Peltier, G. (1996) Evidence for an association of *ndhB*, *ndhJ* gene products and ferredoxin-NADP-reductase as components of a chloroplastic NAD(P)H dehydrogenase complex. *FEBS Letters* **378,** 277-280.

Guo, F., Gopaul, D.N. and Van Duyne, G.D. (1997) Structure of Cre recombinase complexed with DNA in a site-specific recombination synapse. *Nature* **389,** 40-46.

Hajdukiewicz, P., Svab, Z. and Maliga, P. (1994) The small, versatile pPZP family of Agrobacterium binary vectors for plant transformation. *Plant Mol Biol* **25,** 989-994.

Hajdukiewicz, P.T.J., Gilbertson, L. and Staub, J.M. (2001) Multiple pathways for Cre/*lox*-mediated recombination in plastids. *Plant J* **27,** 161-170.

Hoess, R.H., Ziese, M. and Sternberg, N. (1982) P1 site-specific recombination: nucleotide sequence of the recombining sites. *Proc Natl Acad Sci U S A* **79,** 3398-3402.

Hoess, R.H., Wierzbicki, A. and Abremski, K. (1986) The role of the *loxP* spacer region in P1 site-specific recombination. *Nucleic Acids Res* **14,** 2287-2300.

Kanevski, I. and Maliga, P. (1994) Relocation of the plastid *rbcL* gene to the nucleus yields functional ribulose-1,5-bisphosphate carboxylase in tobacco chloroplasts. *Proc Natl Acad Sci U S A* **91,** 1969-1973.

Khan, M.S. and Maliga, P. (1999) Fluorescent antibiotic resistance marker to track plastid transformation in higher plants. *Nat Biotechnol* **17,** 910-915.

Kolodner, R.D. and Tewari, K.K. (1979) Inverted repeats in chloroplast DNA from higher plants. *Proc Natl Acad Sci U S A* **76,** 41-45.

Lilly, J.W., Havey, M.J., Jackson, S.A. and Jiang, J. (2001) Cytogenomic analyses reveal the structural plasticity of the chloroplast genome in higher plants. *The Plant Cell* **13,** 245-254.

Loonstra, A., Voojis, M., Beverloo, H.B., Allak, B.A., Van Drunen, E., Kanaar, R., Berns, A. and Jonkers, J. (2001) Growth inhibition and DNA damage induced by Cre recombinase in mammalian cells. *Proc Natl Acad Sci U S A* **98,** 9209-9214.

Maliga, P., Kuroda, H., Corneille, S., Lutz, K. and Svab, Z. (2002) Chloroplasts for the production of recombinant proteins. PS2001 Proceedings: 12th International Congress on Photosynthesis. CSIRO Publishing: Melbourne, Australia, 2001. ISBN 0643067116, www.publish.csiro.au/ps2001.

Ohyama, K., Fukuzawa, H., Kohchi, T., Shirai, H., Sano, T., Sano, S., Umesono, K., Shiki, Y., Takeuchi, M., Chang, Z., Aota, S., Inokuchi, H. and Ozeki, H. (1986) Chloroplast gene organization deduced from the complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA. *Nature* **322,** 572-574.

Sauer, B. (1992) Identification of cryptic *lox* sites in the yeast genome by selection for Cre-mediated chromosome translocations that confer multiple drug resistance. *J Mol Biol* **223,** 911-928.

Sauer, B. (1996) Multiplex *CRE/lox* recombination permits selective site-specific DNA targeting to both a natural and an engineered site in the yeast genome. *Nucleic Acids Res* **24,** 4608-4613.

Schmidt, E.E., Taylor, D.S., Prigge, J.R., Barnett, S. and Capecchi, M.R. (2000) Ilegitimate Cre-dependent chromosome rearrangements in transgenic mouse spermatids. *Proc Natl Acad Sci U S A* **97,** 13702-13707.

Shikanai, T., Shimizu, K., Ueda, K., Nishimura, Y., Kuroiwa, T. and Hashimoto, T. (2001) The chloroplast *clpP* gene, encoding a proteolytic subunit of ATP-dependent protease, is indispensable for chloroplast development in tobacco. *Plant Cell Physiology* **42,** 264-273.

Shinozaki, K., Ohme, M., Tanaka, M., Wakasugi, T., Hayashida, N., Matsubayashi, T., Zaita, N., Chungwongse, J., Obokata, J., Yamaguchi-Shinozaki, K., Ohto, C., Torazawa, K., Meng, B.Y., Sugita, M., Deno, H., Kamogashira, T., Yamada, K., Kasuda, J., Takaiwa, F., Kato, A., Tohdoh, N., Shimada, H. and Sugiura, M. (1986) The complete nucleotide sequence of the tobacco chloroplast genome. *Plant Molecular Biology Reporter* **4,** 110-147.

Svab, Z. and Maliga, P. (1993) High-frequency plastid transformation in tobacco by selection for a chimeric *aadA* gene. *Proc Natl Acad Sci U S A* **90,** 913-917.

Svab, Z., Hajdukiewicz, P. and Maliga, P. (1990) Stable transformation of plastids in higher plants. *Proc Natl Acad Sci U S A* **87,** 8526-8530.

Thyagarajan, B., Guimaraes, M.J., Groth, A.C. and Calos, M.P. (2000) Mammalian genomes contain active recombinase recognition sites. *Gene* **244,** 47-54.

Timko, M.P., Kaush, A.P., Hand, J.M. and Cashmore, A.R. (1985) Structure and expression of nuclear genes encoding polypeptides of the photosynthetic apparatus. In *Molecular biology of the photosynthetic apparatus*. (Steinback, K.E., Bonitz, S., Arntzen, C.J. and Bogorad, L., eds). Cold Spring Harbor: Cold Spring Harbor Laboratory, pp. 381-396.

Van Duyne, G.D. (2001) A structural view of Cre-*loxP* site-specific recombination. *Annu. Rev. Biophys. Biomol. Struct.* **30,** 87-104.

Zoubenko, O.V., Allison, L.A., Svab, Z. and Maliga, P. (1994) Efficient targeting of foreign genes into the tobacco plastid genome. *Nucleic Acids Res* **22,** 3819-3824.

Author Contributions to research described in Appendix 2

Kerry Lutz, Dr. Sylvie Corneille, and Prof. Pal Maliga designed research; KL, SC and AA carried out research. AA tested recombination between loxP site and pseudo-loxP sites identified by computer search and CRE-mediated deletion in the absence of perpetual DNA inversions. Figures A2.6 and A2.7 were prepared by AA; KL, SC, AA and PM analysed data, KL, SC, AA and PM wrote the paper.

Appendix 3

Plastid transformation vectors

Plastid transformation requires efficient vector systems for delivery of foreign DNA into the plastids and is based on homologous recombination between the transforming DNA and the plastid genome. These vector systems are *E. coli* plasmid derivatives with plastid DNA (ptDNA) sequences (1-2 kb), flanking a selectable marker gene and multiple cloning site, which direct integration of the transforming DNA at specific sites in the plastid genome. Both the marker gene and the gene of interest have their own 5' and 3' regulatory regions that are derived from plastid sequences or heterologous systems. The 5'-regulatory sequence includes the promoter and translational control region and most commonly used 5'-UTRs are complete or truncated versions of plastid genes, such as *psbA, rbcL, atpB* and the rRNA operon. The 3'-regulatory region encodes the mRNA 3'-UTR and this region is important for mRNA stability and functions as an efficient transcription terminator. Most commonly used 3'-regulatory regions are derived from the plastid *psbA*, *rbcL* and *rps16* genes (Maliga 2004). Most recent vectors are set up for marker gene excision with the CRE-*loxP* (Corneille et al, 2001) and phage phiC31 (Kittiwongwattana et al, 2007) marker gene elimination systems.

As a part of my thesis research I constructed a series of vectors to transform both the *A. thaliana* and *N. tabacum* plastid genomes (Table A3.1). These vectors are characterized by the presence of *loxP* sequences flanking (f*lox*ed) the marker gene. These *loxP* sequences enable the elimination of the marker gene once homoplastomic plants are obtained, thus facilitating the recycling of the marker gene for subsequent transformation of the same plant. In the Arabidopsis transformation vectors, in addition to the f*lox*ed marker gene, an *attB* sequence has been incorporated in the vectors to facilitate site-specific insertion of transforming DNA. The use of the *attB* sequence as the insertion target for plastid transformation is discussed in Appendix 1.

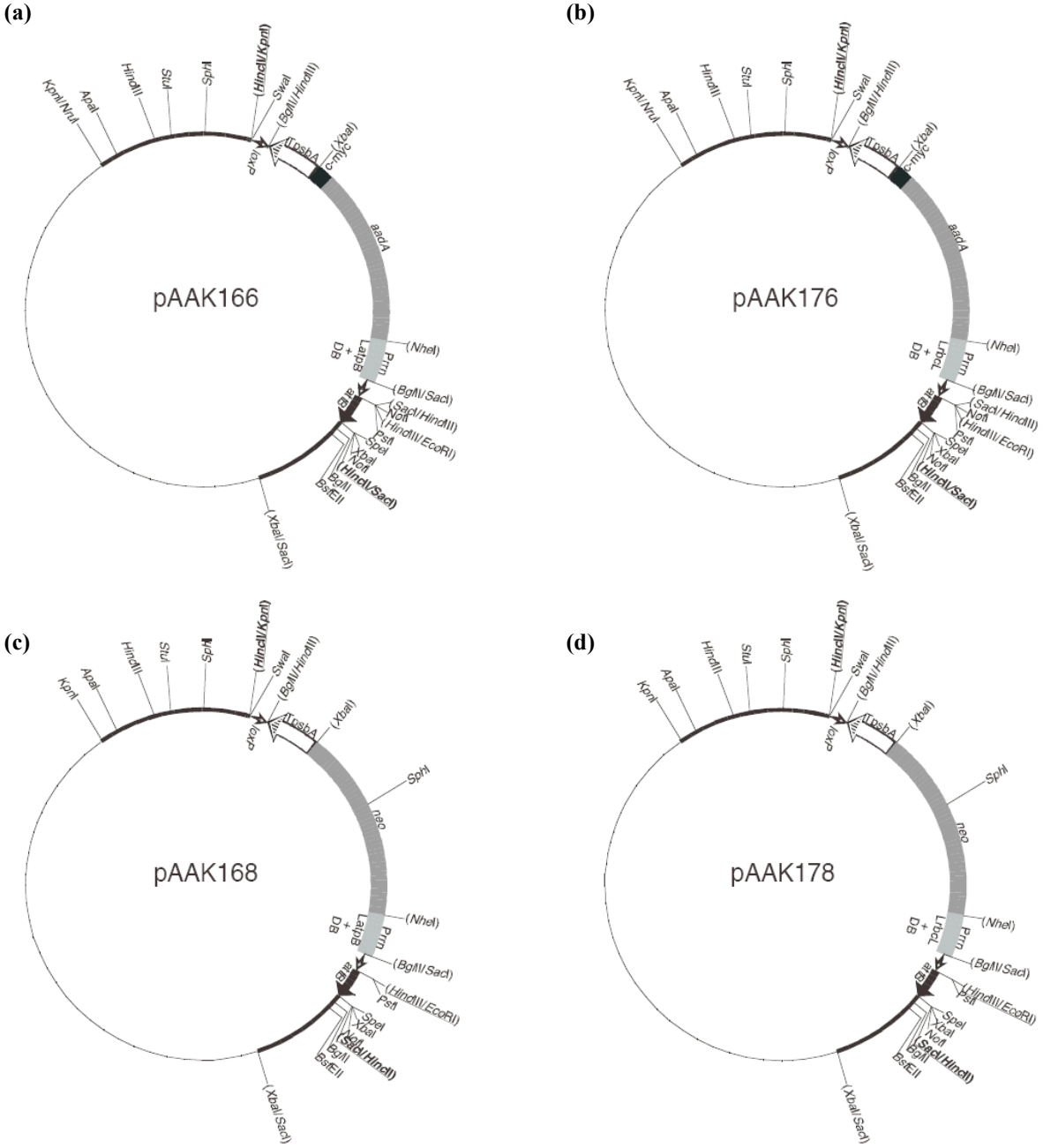
Vector Details

Arabidopsis plastid loxP/attB *vector*

The Arabidopsis vectors are pBSIIKS+ (Stratagene, La Jolla, CA, USA) derivatives that contain an *attB* sequence, a f*lox*ed *neo* or *aadA* marker gene and an MCS. These vectors target insertion to either the *trnV-rps12/7* or *trnI-trnA* intergenic regions in the Arabidopsis plastid genome. The vectors differ in the 5' regulatory sequences that drive expression of the marker genes. The 3' regulatory sequences (TpsbA) are the same for all the vectors. Vectors pAAK166 and pAAK176 contain an *attB* sequence upstream of a f*lox*ed *aadA* gene and they differ in the promoter/leader sequence that drives expression of the *aadA* gene (Figure A3.a & 3.b). These vectors target insertion of the transforming DNA to the *trnV-rps12/7* intergenic region. Vectors pAAK168 and pAAK178 contain an *attB* sequence upstream of a f*lox*ed *neo* gene and they differ in the promoter/leader sequence that drives expression of the *neo* gene (Figure A3.c & 3.d) and are targeted to the *trnV-rps12/7* intergenic region. Vectors pAAK262 and pAAK272 target insertion of the transforming DNA to the *trnI-trnA* intergenic region and contain an *attB* sequence upstream of a f*lox*ed *aadA* gene. They differ in the promoter/leader sequence that drives expression of the *aadA* gene (Figure A3.e & 3.f).

Tobacco plastid loxP *vector*

Vector pAAK205 (Figure A3.g) is a pBSIIKS+ (Stratagene, La Jolla, CA, USA) derivative that contains a f*lox*ed *neo* marker gene and an MCS that targets insertion to the *trnV-rps12/7* intergenic region in the tobacco plastid genome. The *neo* gene is driven by the Prrn promoter that is fused to the 5' translational control region of the plastid *atpB* gene. The TrbcL (3' regulatory sequence of the tobacco plastid *rbcL* gene) serves as the terminator of transcription.


(a)
pAAK166
(b)
pAAK176
(c)
pAAK168
(d)
pAAK178

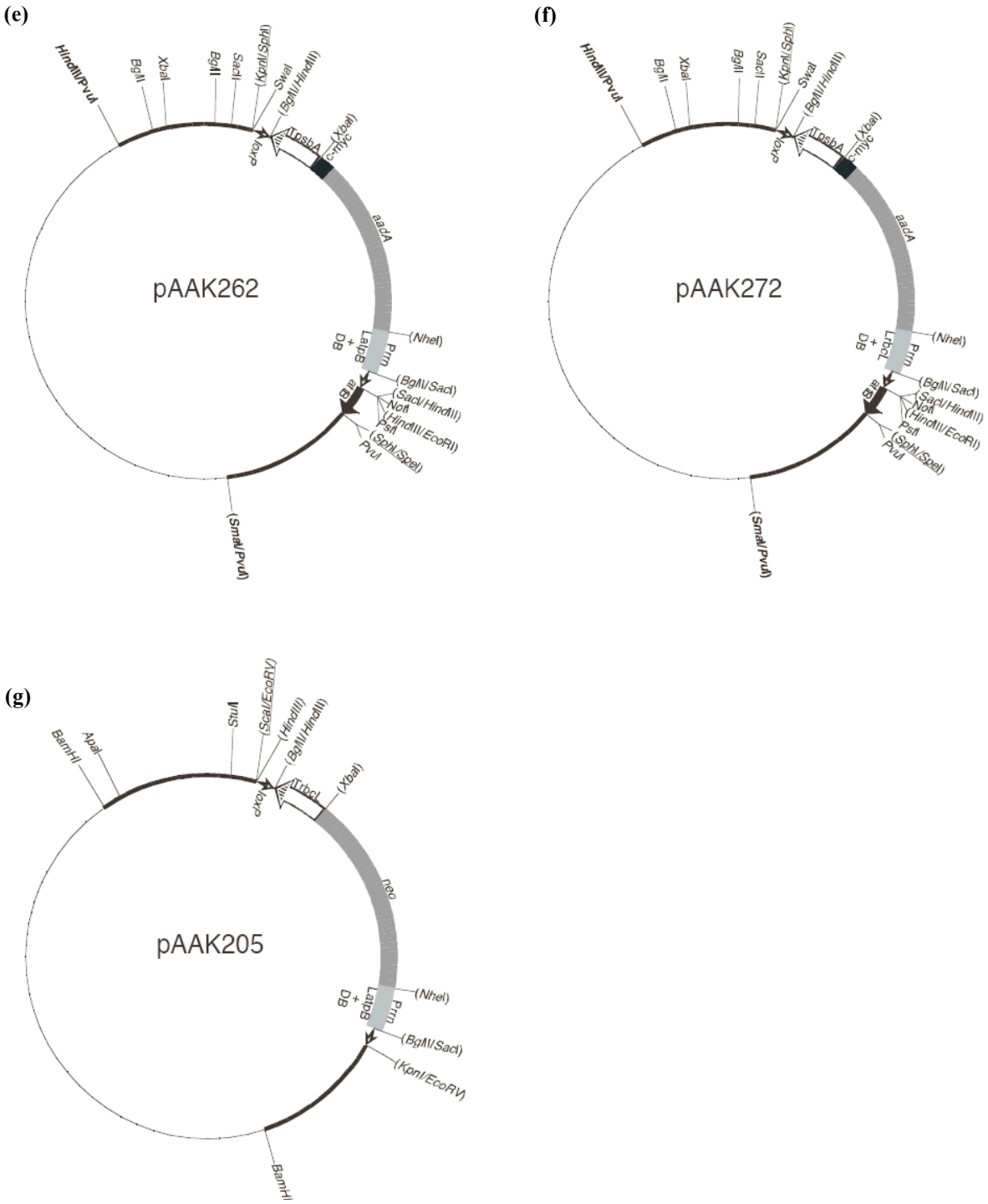


Figure A3.1. Maps of the *A. thaliana* and *N. tabacum* plastid vectors. Maps of *A.thaliana* plastid vectors are figures A3.a, A3.b, A3.c, A3.d, A3.e and A3.f and *N.tabacum* plastid vector map is figure A3.g. Arabidopsis plastid vectors pAAK166 and pAAK176 target insertion to the *trnV-rps12/7* intergenic region; vectors pAAK168 and pAAK178 target insertion to the *trnV-rps12/7* intergenic region; Arabidopsis plastid vectors pAAK166 and pAAK176 target insertion to the *trnI-trnA* intergenic region. Vector pAAK205 targets insertion to the *trnV-rps12/7* intergenic region of the tobacco plastid genome. Shown are promoter

Prrn/LatpB +DB (plastid rRNA operon promoter region fused to the plastid *atpB* 5' translation control region), Prrn/LrbcL +DB (plastid rRNA operon promoter region fused to plastid *rbcL* 5' translation control region), TpsbA (plastid *psbA* 3' regulatory region), TrbcL (plastid *rbcL* 3' regulatory region) *loxP* sites (black arrow heads), *aatB* sites (black block arrows), *aadA* (spectinomycin resistance selectable marker gene) *neo* (kanamycin resistance selectable marker gene) and restriction sites.

Table A3.1. *A. thaliana* and *N. tabacum* plastid transformation vectors

Vector	Marker Gene	5' regulatory sequences	3' regulatory sequences	*loxP*/*attB* sites	Insertion Site	Flanking sequences origin
pAAK 166	*aadA*	P*rrn*/L*atpB*	T*psbA*	*loxP/attB*	*trnV-rps12/7*	Arabidopsis
pAAK 176	*aadA*	P*rrn*/L*rbcL*	T*psbA*	*loxP/attB*	*trnV-rps12/7*	Arabidopsis
pAAK 168	*neo*	P*rrn*/L*atpB*	T*psbA*	*loxP/attB*	*trnV-rps12/7*	Arabidopsis
pAAK 178	*neo*	P*rrn*/L*rbcL*	T*psbA*	*loxP/attB*	*trnV-rps12/7*	Arabidopsis
pAAK 262	*aadA*	P*rrn*/L*atpB*	T*psbA*	*loxP/attB*	*trnI-trnA*	Arabidopsis
pAAK 272	*aadA*	P*rrn*/L*rbcL*	T*psbA*	*loxP/attB*	*trnI-trnA*	Arabidopsis
pAAK 205	*neo*	P*rrn*/L*atpB*	T*rbcL*	*loxP*	*trnV-rps12/7*	Tobacco

References

Corneille, S., Lutz, K., Svab, Z. and Maliga, P. (2001) Efficient elimination of selectable marker genes from the plastid genome by the CRE-*lox* site-specific recombination system. *Plant J* **72,** 171-178.

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

Kittiwongwattana, C., Lutz, K., Clark, M., Maliga, P. (2007) Plastid marker gene excision by the phiC31 phage site-specific recombinase. *Plant Mol Biol* **64,** 137-143.

Author Contributions to research described in Appendix 3

Arun Azhagiri, Kerry Lutz, and Prof. Pal Maliga designed research; AA carried out research; AA and PM analysed data, AA and PM wrote the manuscript.

CURRICULUM VITAE

Arun K. Azhagiri

Education:

1989-1992	B.Sc. Microbiology, St. Joseph's College of Arts and Science, Bangalore University, Bangalore, India.
1992-1994	M.Sc. Industrial Microbiology, South Gujarat University, Surat, India
1994-1999	Junior Research Fellow, National Centre for Biological Sciences, Bangalore, India.
1999-2007	Ph.D. Plant Biology, Rutgers, The State University of New Jersey, New Brunswick, NJ.

Publications:

Corneille, S., Lutz, K. A., Azhagiri, A. K., and Maliga, P. (2003) Identification of Functional *lox* Sites in the Chloroplast Genome. The Plant Journal 35: 6, 753-762.

Lutz, K. A., Corneille, S., Azhagiri, A. K., Svab, Z., and Maliga, P. (2004) A Novel Approach to Chloroplast Transformation Utilizes the phiC31 Phage Integrase. The Plant Journal 37: 6, 906-913.

Arun K. Azhagiri, Pal Maliga. (2007) DNA markers define plastid haplotypes in Arabidopsis thaliana. Current Genetics 51:4, 269