Biological Forum – An International Journal **8**(1): 397-403(2016)

ISSN No. (Print): 0975-1130
ISSN No. (Online): 2249-3239

Cloning of Flanking Regions of Plastidal Vectors in Soybean (*Glycine max* L.): partial *AtpB* and *RbcL*

Amir Sam Pordel, Bahram Baghban Kohne-Rooz and Parviz Soleimani
Genetic Engineering Laboratory, Faculty of Agriculture,
University of Tabriz, Tabriz, Iran

(Corresponding author: Parviz Soleimani)
(Received 30 March, 2016, Accepted 27April, 2016)
(Published by Research Trend, Website: www.researchtrend.net)

ABSTRACT: For successful plastid transformation, a foreign DNA and endogenous regulatory sequences on the transgene cassette are flanked with plastid DNA fragments homologous to a specific site on plastome. To clone the intergene between *rbcL* and *AtpB* of *Glycine max* plastome along with portions of *rbcL* and *AtpB* genes as flanking sequences, forward and reversed primers were designed using Primer-Blast in NCBA. Restriction sites and over-hanged AA were added to 5' of either primer. A pre-test gradient temperature was used to identify the best annealing temperature for multiple primer sets during polymerase chain reaction, which was 62.7°C in this case. Flanking sequences as long as 195 Nucleotides from 3' late part of RbcL and 306 Nucleotides from 5' region of *AtpB* were amplified. Results showed that *RbcL* part of the fragment carries 10 gaps out of 195 nucleotide pairs, which accounts for 5%, but there is no gap in *AtpB* part.

Keywords: *AtpB,* flanking sequence, gradient temperature, Plastid transformation, *RbcL,*

INTRODUCTION

Chloroplasts, and plastids as a whole, are eukaryotic algae and plant specific cells containing genomes of circular, small size (200 to 220 kb) double stranded DNA. Chloroplast genomes (plastome) typically contain 110 to 120 unique genes distributed across the plastid polyploidy genetic system with up to 10,000 copies of the chloroplast genome in each plant cell, compared to their cyanobacterial ancestor with more than 1,500 genes (Verma & Daniell, 2007; Martin *et al,* 2002).

There are many features associated with plastids, which make them a promising subject of research for land plants transformation through plastid genetic engineering. In most angiosperm plant species, plastid genes inherited uniparentally in a maternal fashion, even though transgenic chloroplasts might be present in pollen, plastid engineered plants are genetically contained (Daniell, 2002). Additionally, although they accumulate at a level 169 times more than the nuclear transgenic plants, alien transcripts do not silence genes in chloroplast transgenic lines (Dhingra *et al*, 2004). Another advantage of plastid transformation is the ability to accumulate in the chloroplast any foreign proteins or their biosynthetic products. This is prominent when the plastid recombinant protein makes up 70% of leaf protein (Daniell *et al*, 2009; Ruhlman *et al,* 2010) with no post-transcriptional gene silencing being observed (DeCosa *et al*, 2001). Besides providing an important tool to investigate many aspects of gene regulation and expression (Day &Goldschmidt-Clermont, 2011), plastid transformation has also obtained strong interest for applications in biotechnological researches (Meyers *et al*, 2010), especially producing chloroplast-derived biopharmaceutical proteins (Daniell, 2006), and taking away the harmful effects of some foreign proteins or their biosynthetic products, through accumulating in the chloroplast rather than in the cytoplasm (Bogorad, 2000).

Chloroplast genetic engineering involves incorporating foreign gene into the plastid's DNA genome (ptDNA) using plastid vectors. Homologous recombination has been the cornerstone of all characterized recombination events in plastids. Efficient homologous recombination requires insertion of a foreign (heterologous) DNA flanked with homologous plastid DNA fragments for successful recombination event and endogenous regulatory sequences embedded on the transgene cassette to a specific site in the plastid genome (Vermaand & Daniell, 2007; Bansal & Saha, 2012), referred to as the targeted region. The right and left nucleotide sequences are taken from the host plastid genome, which efficiently direct the site-specific transgene insertion through homologous recombination. Successful targeted integration into the chloroplast genome by homologous recombination of flanking sequences eliminates the concerns of position effect, the common problem with nuclear transgenic lines (Daniell, 2005).

Both the aforementioned flanking sequences are necessary and generally about 1 kb in size, although they have no special properties except being homologous to the chosen target site (DeCosa*et al*, 2001). Accordingly, the first step in construction of chloroplast vectors is to identify regions to integrate transgenes and regulatory sequences that support optimal foreign gene expression.

So far, several sites within the plastid genome have been reported suitable for transgene integration (Maliga, 2004). Transgenes were first integrated into transcriptionally silent spacer regions. Recently, transcriptionally active spacer regions provide unique advantages over silent spacer ones, such as insertion of transgenes with no 5' or 3'untranslated regions (*UTRs*) or promoters (DeCosa*et al*, 2001). The genes of the rubisco large subunit (*RbcL*) and of ATP synthase polypeptides, the ATPase-□ (*AtpB*), have been studied in variety of dicot plants (Manen, 1994). They lie contiguously on the plastome in opposite directions, with their start codons facing each other at a distance of 600-800 nucleotides, depending on the plant you may study. Despite being highly variable in the plant kingdom, which makes it suitable for phylogenetic studies, this region contains the conserved promoting elements of each of the two genes (Zurawski& Clegg, 1987; Manen, 1994). In soybean, these genes are located on long single copy (LSC) region on the *cpDNA* (Saski*et al*, 2005), with both genes being transcribed in different direction.

In this paper, we endeavored to clone the *AtpB* - *RbcL* region in Soybean chloroplast genome with the goal to provide flanking sequences for plastid transformation cassettes.

MATERIALS AND METHOD

Genomic DNA extraction

Soybean (*Glycine max* L.) seeds were obtained from Gorgan Agricultural Research Center. The seeds planted into several pots, and after reaching 2-leaf stage, fresh leaves from plants that have been incubated in dark for 24 hours for starch degradation were used in DNA extraction. Total DNA of leaves was extracted with a modified CTAB method (Doyle & Doyle 1987; Huang, 2000). UV-Visible spectrophotometer was used to measure the absorbance of isolated genomic DNA at A260 and A280 nm. A sample run on 0.8% agarose gel was utilized to have a visible test of quantity and quality of extracted genomic DNA.

Bioinformatics and Primer design

Soybean chloroplast genome obtained from NCBI accession number NC_007942. Two primers were designed using Primer-Blast in NCBA, one forward and one reversed. Also, restriction sites and AA overhung were added to 5' of either primer. During primers design, GC contents, melting temperature, primer length, and primer-dimer formation were carefully considered.

These primers were employed in combination with templates of isolated genomic DNA for PCR amplification of sequences of interest. 5 µl of DNA extract was amplified by standard PCR methodology with the oligonucleotides 5'GAAGTAGTAGGATTGATTCTC as Forward, and 3'GACCATGTACCTGTTGACAT as Reverse primers, also 5 U Taq DNA polymerase (Fermentas, USA), 5 µl $MgCl_2$, 2 µl of each dNTP,10 µl PCR reaction Buffer were added to the mixture.

The PCR reactions were programmed on thermal cycles, one initial denaturizing cycle at 94°C for 5 min, initial annealing step at 50°C for 30 s, and elongation step at 72°C for 30 s, followed by 35 cycles at 94°C for 30s min, 50°C for 1 min, 72°C for 1:30 s, then followed by a final extension step at 72°C for 5 min (Joshi, 2010). Also was programed a Temperature gradient to determine the best temperature for amplification. Amplified PCR products were analyzed using electrophoresis in agarose gel (0.8%), stained with Redsafe®.

PCR products evaluation

The 100 µl of amplification mixture was loaded on a preparative 0.8 % agarose gel Then, the DNA was extracted from the gel using Extraction-kit (Bioneer co). The gel was let sit in 0.5 x TBE (Tris, Boric acid and EDTA) buffer for ~60 min and then stained in 1 µg/ml RedSafe® solution.

Ligation Reaction

The *pTG-19-T* (Vivantis) was added to PCR products with a 1 to 3. Then dT overhangs of the *pTG-19-T* was ligated to dA sequence of the PCR product.

Bacteria transformation

The DH5α strain of Escherichia coli was used as competent cell, and TSS competent cell protocol with heat shock was used to transfer the ligation product to the Bacteria.

100 µl TSS was added to the pellet of bacteria, incubated on ice for 30 min, and then heat shocked for 90 s at 42°C. Then, 700 µl of LB was added, and the mixture was incubated for 60 min at 37°C for bacteria recovery. Then the compound was cultured on a medium containing LB Agar, X-gal, IPTG and Ampicillin (Selectable marker) for 16 hours. Some of the white colonies were then chosen for PCR cloning. These colonies were used as Template DNA in PCR cloning instead of Genomic DNA. Finally, they were run on 0.8% Agarose gel. Then the white colonies were incubated again for 16 hours.

Extraction and Digestion of Plasmid

The plasmids were extracted by Alkalinelysis method (Kotchony *et al*, 2003) with some modification, and they were excised by *Sac* I and *Hind* II Restriction Enzymes, until the fragment was obtained and sent for sequencing. The results of sequencing were studied and monitored using BLAST against previously identifies sequences.

RESULTS AND DISCUSSION

In this experiment, we successfully cloned the non-coding region between genes *RbcL* and *AtpB* of *Glycine max* chloroplast genome, along with poertions of *RbcL* and *AtpB* genes as flanking sequences. The two flanking sequences amplified here are 195 (from *RbcL*) and 306 (from *AtpB*) nucleotides in length (Fig. 6).

Genomic DNA of soybean was extracted by a modified CTAB method and UV-visible spectrophotometer was used to evaluate concentration of genomic DNA, which was 100 ng/µl. We performed PCR reaction to amplify RbcL - *AtpB* sequence from Soybean chloroplast genome. PCR Reaction was carried out in 3 main stages: Denaturation, Annealing, Extention, as described earlier. A temperature gradient was used to optimize the reaction conditions in a single run, identifying the best annealing temperature for multiple primer sets, performing reactions that require different annealing temperatures at the same time. After amplification, PCR products were then compared on Agarose gel. The differences between PCR temperatures revealed on Agarosegel (Fig. 1). As you can see, the best temperature was 62.7°C. So we continued the experiment using products of 62.7°C.

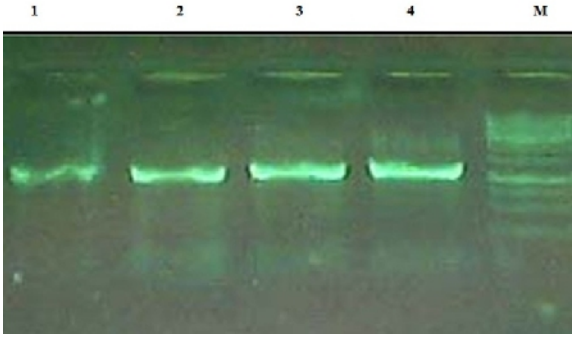

Fig. 1. *RbcL - AtpB* PCR products. Numbers one through four show temperature gradient. Number four proved the best temperature for amplification (62/7°C).

Fig. 2. Selection medium for white/blue colonies. Whites represent colonies with fragment and blues represent colonies without inserted fragment. Vector absorbed into DH5α strain under the influence of TSS and Head Shock.

In the next step, we attempted to ligate the DNA fragments chosen from previous step into a vector (pTG-19T, a linear vector), prior to being amplified in host bacteria. Ligation reaction performed using 3 to 1 ratio of PCR products to Vector. In ligation reaction, over-hanged Adenine nucleotides provided on each side of inserted sequence link to over-hanged Thymine nucleotides on the either side of the vector. After ligation, the linear vector changes into a circular vector, which makes its insertion into the host bacterium much easier.

The circular vector was then absorbed into DH5αstrain under the influence of transformation storage solution (TSS) material and Heat shock. The cells were incubated on a medium containing x-gal, IPTG, Ampicilin. X-gal consists of β-galactosidase that will integrate with α- galactosidase on vector to produce blue colors. But if their integration is interrupted by insertion of a foreign sequence, the α- galactosidase is damaged, hence white colors form instead of blue. Which means white colonies may have inserted fragment in their plasmid (Fig. 2).

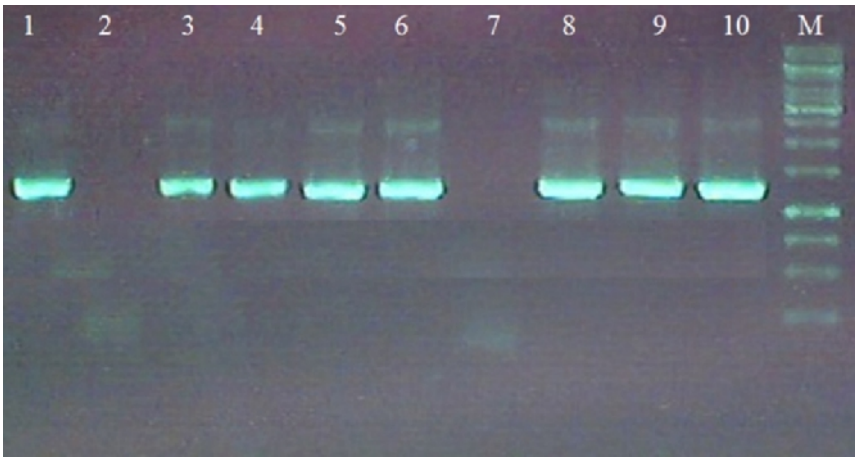

Fig. 3. Analysis of PCR Cloning of white colonies on gel. The white colonies number 2 and 7 did not form any sharp band, therefore their color was resulted from mutation, while the rest of white colonies expectedly carried the inserted fragment.

White colony formation could also be the result of mutation in β-galactosidase. To make sure that white colonies have resulted from the foreign fragment insertion and not mutation, further PCR analysis was carried out on white colonies (Fig. 3). One important property of expression vector is to have appropriate restriction sites at either side flanking sequences. The function of these restriction site is to ease insertion or excision of cassette when building the expression vector. When we designed the forward and reverse primers, *Sac I* and *Hind II* restriction sites as well as two-nucleotide extension AA were added to 5' end of the primers (Fig. 4).

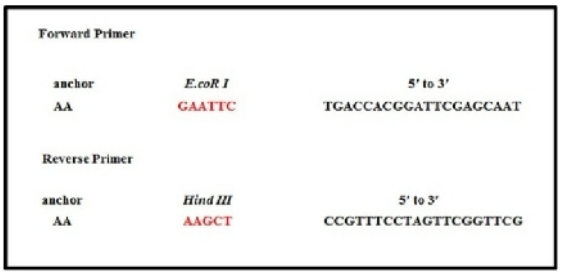

Fig. 4. Primers Forward and Reverse composition along with restriction site and Anchor AA are portrayed. Primer design was performed in Primer-Blast (NCBI). In the ligation process, AA overhung on the PCR amplified segment will link to TT overhung on the pTG19-T.

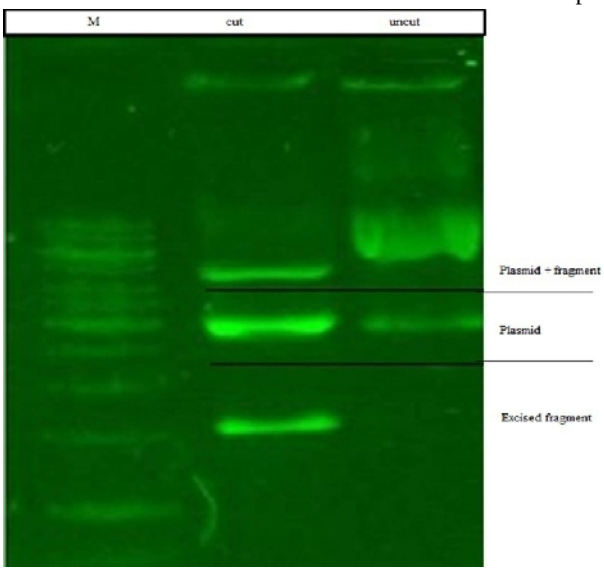

Fig. 5. Gel agarose scan of digested plasmids. First *RbcL - AtpB* segment extracted from plasmid by Alkalinelysis. Then was digested by *Sac I* and *Hind III* restriction enzymes. Cut: restriction enzymes Digested plasmid (4100 bp), Uncut: non- Digested plasmid.

Then let the PCR amplification happen using these primers. Now we had a DNA segment containing flanking sequences with two restriction sites at either side. To test the efficiency of these sites, first, white colonies were incubated for 16 hours, then plasmids were extracted by Alkaline lysis mini-prep method consisted of Solution I (re-suspends the bacteria), Solution II (dissolves the cell wall) known as Denaturation and central part of the procedure, and Solution III (neutralizes alkalinity of the mixture) (Bimboin and Doly, 1979). After extraction, the plasmids were digested with *Sac I* and *Hind III* Irestriction enzymes and the products were observed on an agarose gel (Fig. 5). The results in Fig. 5.gives evidence about the affectivity of enzymes as well as the precision of the process. So, we sent the obtained fragments for sequencing. The results of sequencing, illustrated in Fig. 6, yielded a sequence with 1276 bp length, with flanking sequences from *RbcL* and *AtpB* on the left and right, respectively. The two flanking sequences amplified here have 195 (from *RbcL*) and 306 (from *AtpB*) nucleotides length (Fig. 6).

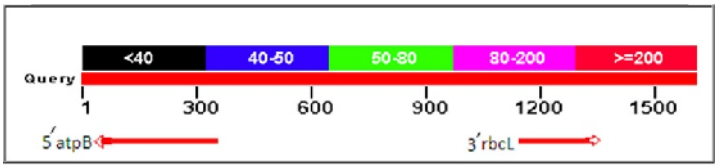

Fig. 6. An illustration pertaining to the supplementary BLAST of the cloned segment against natural full-length *RbcL* (A) and *AtpB* (B) genes showed the exact amount of coverage and similarity to each gene. The result of the blast for the cloned segment against *RbcL* gene showed that the forward primer has helped amplify 3' late part of *RbcL* (195 Nu.) and 5' region of *AtpB* (306 Nu.). The amount of identities to 3' *RbcL* and to *AtpB* is 87% and 100%, respectively, with no gaps for the latter. Also in 3' late part of *RbcL*, the fragment carries 10 gaps out of 195 nucleotide pairs, which accounts for 5%.

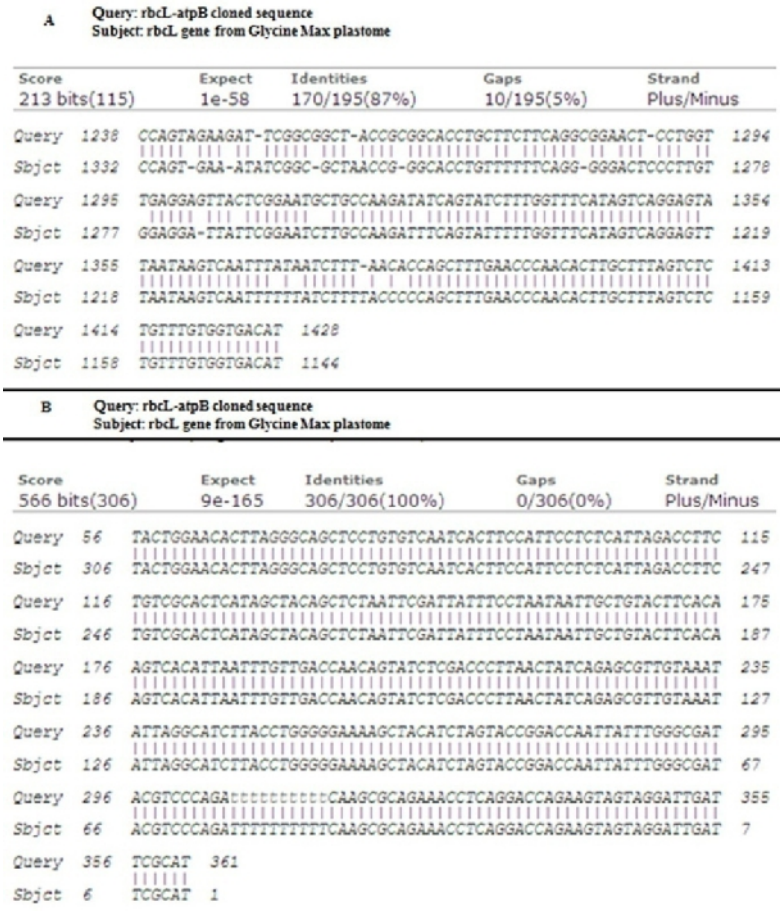

Fig. 7. Simultaneous comparison of *RbcL* - *AtpB* fragment to *RbcL* and *AtpB* gene of Soybean plastome using BLAST. It shows that 3' and 5' regions of the cloned fragment matches the 3' and 5' regions of *RbcL* and *AtpB* genes, respectively. The arrows show the transcription direction of respective genes.

The supplementary BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi) of the cloned segment showed the exact length of covered regions of *RbcL* and *AtpB* genes along with the amount of the similarity to each gene (Figs. 6 and 7). The Fig. 6 indicates that the forward primer has helped amplify 3' late part of *RbcL* (195 Nu.) and 5' region of *AtpB* (306 Nu.). The amount of identities to 3' *RbcL* is 87% and to *AtpB* is 100%, with no gaps for the latter. Also in 3' late part of *RbcL*, the fragment carries 10 gaps out of 195 nucleotide pairs, which accounts for 5%. An additional research must be done so as to comprehend the reason for gaps, whether it comes from PCR sensitivity or true variation of the *Glycine max* subspecies. The simultaneous BLAST of *RbcL* - *AtpB* fragment against *AtpB* and *RbcL* genes also confirms the process (Fig. 7).

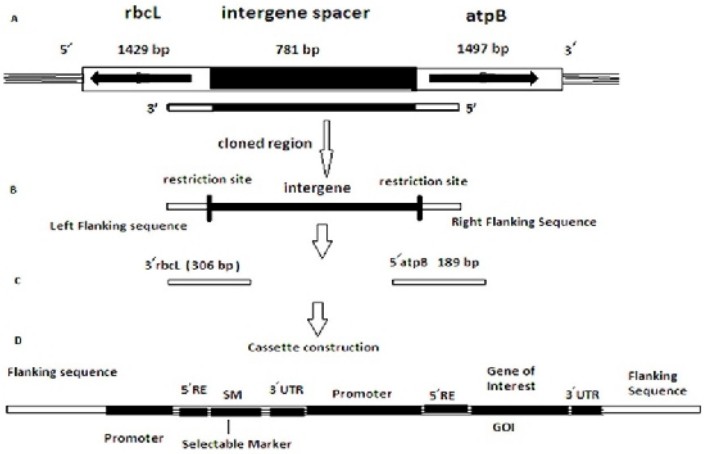

Fig. 8. Depiction of the process from fragment extraction to cassette construction; Using meticulously designed primers, 3' *RbcL* and 5' *AtpB* regions were cloned with PCR technique; A) the location of both *RbcL* and *AtpB* genes on Soybean plastid genome (for simplicity some parts of the process were not shown), B) the cloned and extracted fragment bearing left and right flanking sequences, C) Shows free flanking sequences ready to be utilized in transformation process and D) A schema of a typical cassette.

Fig. 8 summarizes the process from the extraction of the fragment to construction the expression cassette. Successful chloroplast transformation requires homologous recombination, whereby the plasmid genome takes on the transgene. Accordingly, vectors designed for chloroplast transformation carry homologous flanking sequences on either side of the transgene cassette to facilitate homologous recombination. Flanking sequences left and right are both necessary for successful homologous recombination, though possessing no physiological property other than being homologous to the chosen target and helping integration the transgene into the targeting site (Verma & Daniell, 2007).

REFERENCE

Bansal, K.C. & Saha, D. (2012). Chloroplast Genomics and Genetic Engineering for Crop Improvement. *Agriculture Resources*.**1**: 53-66.

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

Daniell, H., Kumar, Sh. & Dufourmantel, N. (2005). Break through in chloroplast genetic engineering of agronomically important crops. *Trends in Biotechnology*.**23**: 238-245.

Daniell, H., Lee, S.B., Grevich, J., Saski, C., Quesada-Vargas, T., Guda, C., Tomkins, J. & Jansen, R.K. (2006).Complete chloroplast genome sequences of *Solanum bulbocastanum*, *Solanum lycopersicum* and comparative analyses with other Solanaceae genomes. *Theoretical and Applied Genetics*.**112**: 1503-1518.

Daniell, H., Singh, N.D., Mason, H. & Streatfield, S.J. (2009).Plant-made vaccine antigens and biopharmaceuticals. *Trends in Plant Science*.**14**: 669-679.

Daniell, H. (2002). Molecular strategies for gene containment in transgenic crops. *Nature Biotechnology*.**20**: 581-586.

Day, A. & Goldschmidt-Clermont, M. (2011).The chloroplast transformation toolbox: selectable markers and marker removal. *Plant Biotechnology Journal*.**9**: 540-553.

DeCosa, B., Moar, W., Lee, S.B., Miller, M. &Daniell, H. (2001). Over expression of the Bt cry2Aa2 operon in chloroplasts leads to formation of insecticidal crystals. *Nature Biotechnology*. **19**: 71-74.

Dhingr, A., Portis, A.R. & Daniell, H. (2004).Enhanced translation of a chloroplast expressed RbcS gene restores SSU levels and photosynthesis in nuclear antisense *RbcS* plants. *Proceedings of the Natural Academy of Sciences*.**101**: 6315-6320.

Doyle, J.J. & Doyle, J.L. (1987).A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemistry Bulletin*.**19**: 11-15.

Huang, J., Ge, X. & Sun, M. (2000). Modified CTAB protocol using a silica matrix for isolation of plant genomic DNA. *Bio Techniques*. **28**: 432-434.

Kotchoni, O.S., Shonukan, O.O. & Gachomo, W.E. (2003). *Bacillus pumillus*, BPCR16, a promising candidate for cellulase production under conditions of catabolic repression. *African Journal of Biotechnology*. **2**: 140-146.

Maliga, P. (2004). Plastid transformation in higher plants. *Annual Reviews in Plant Biology*. **55**: 289-313.

Manen, J.F., Savolainen, V. & Simon, P. (1994). The *atpB* and *rbcL* Promoters in Plastid DNAs of a Wide Dicot Range. *Journal of Molecular Evolution*. **38**: 577-582.

Manen, J.F., Natali, A. & Ehrendorfer, F. (1994). Phylogeny of Rubiaceae Rubieae inferred from the sequence of a cpDNA intergene region. *Plant Systematics and Evolution*.**190**: 195-211.

Martin, W.R.T., Richly, E., Hansen, A., Cornelson, S., Lins, T., Leister, D., Stoebe, B., Hasegawa, M. & Penny, D. (2002). Evolutionary analysis of Arabidopsis, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. *Proceedings of the Natural Academy of Sciences*.**99**:12246-12251.

Meyers, B., Zaltsman, A., Lacroix, B., Kozlovsky, S.V. & Krichevsky, A. (2010). Nuclear and plastid genetic engineering of plants: comparison of opportunities and challenges. *Biotechnology Advances*. **28**:747-756.

Ruhlman, T., Verma, D., Samson, N. & Daniell, H. (2010). The role of heterologous chloroplast sequence elements in transgene integration and expression. *Plant Physiology*. **152**: 2088-2104.

Saski, C., Lee, S.B., Daniell, H., Wood, T.C., Tomkins, J., Kim, H.G. & Jansen, R.K. (2005). Complete chloroplast genome sequence of *Glycine max* and comparative analyses with other legume genomes. *Plant Molecular Biology*. **59**: 309-322.

Saski, C., Lee, S.B., Fjellheim, S., Guda, C., Jansen, R.K., Luo, H., Tomkins, J., Rognli, O.A., Daniell, H. & Clarke, J.L. (2007). Complete chloroplast genome sequences of *Hordeum vulgare*, *Sorghum bicolor* and *Agrostis stolonifera*, and comparative analyses with other grass genomes. *Theoretical and Applied Genetics*.**115**: 571-590.

Verma, D. & Daniell, H. (2007). Chloroplast Vector Systems for Biotechnology Applications. *Plant Physiology*. **145**: 1129-1143.

Zurawski, G. & Clegg, M.T. (1987). Evolution of higher plant chloroplast DNA-encoded genes: implication for structure-function and phylogenetic studies. *Annual Reviews in Plant Physiology*.**38**: 391-418.