Innovation, Sciences et Développement économique Canada
Office de la Propriété Intellectuelle du Canada

Innovation, Science and Economic Development Canada
Canadian Intellectual Property Office

CA 2781900 C 2023/02/14

(11)(21) 2 781 900

(12) BREVET CANADIEN
CANADIAN PATENT

(13) C

(86) Date de dépôt PCT/PCT Filing Date: 2009/11/25
(87) Date publication PCT/PCT Publication Date: 2010/06/03
(45) Date de délivrance/Issue Date: 2023/02/14
(85) Entrée phase nationale/National Entry: 2012/05/24
(86) N° demande PCT/PCT Application No.: GB 2009/002754
(87) N° publication PCT/PCT Publication No.: 2010/061186
(30) Priorité/Priority: 2008/11/25 (GB0821516.2)

(51) Cl.Int./Int.Cl. *C12N 15/82* (2006.01)

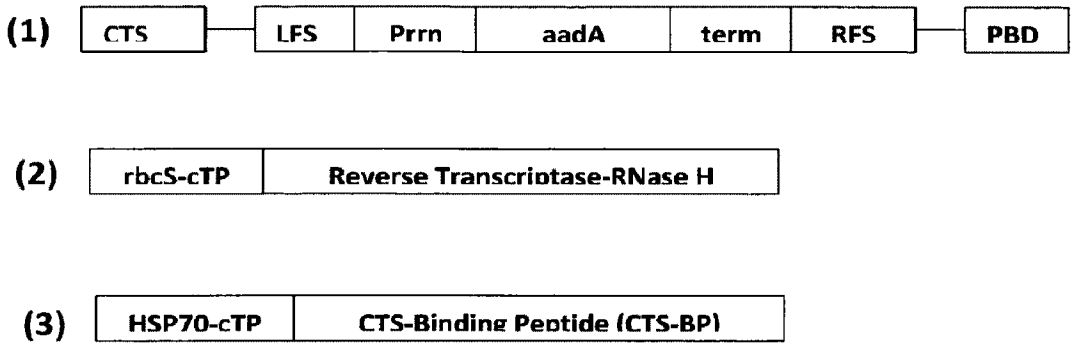
(72) Inventeurs/Inventors:
MALCUIT, ISABELLE, GB;
SOROKIN, ALEXANDER, GB

(73) Propriétaire/Owner:
ALGENTECH SAS, FR

(74) Agent: RIDOUT & MAYBEE LLP

(54) Titre : POLYNUCLEOTIDES POUR LA TRANSFORMATION D'UN PLASTE EN CELLULE VEGETALE ET POUR LA PRODUCTION D'UNE CELLULE OU D'UN VEGETAL CONTENANT LE PLASTE TRANSFORME, DE MEME QUE LA METHODE CONNEXE

(54) Title: POLYNUCLEOTIDES FOR TRANSFORMING A PLASTID IN A PLANT CELL, AND FOR PRODUCING A CELL OR A PLANT CONTAINING THE TRANSFORMED PLASTID, AND THE METHOD THEREFOR



(57) Abrégé/Abstract:
Method for heterologous protein production in plant cell plastids comprising introducing into plant cells nucleic acid components that encode heterologous proteins under the control of promoters operative in plastids, vectors, host cells, plants and uses thereof.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
3 June 2010 (03.06.2010)

PCT

(10) International Publication Number
WO 2010/061186 A3

(51) International Patent Classification:
C12N 15/82 (2006.01) *A01H 5/00* (2006.01)

(21) International Application Number:
PCT/GB2009/002754

(22) International Filing Date:
25 November 2009 (25.11.2009)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0821516.2 25 November 2008 (25.11.2008) GB

(71) Applicant *(for all designated States except US)*: **ALGENTECH SAS** [FR/FR]; Pépinière Genopole Enterprise, 4, rue Pierre Fontaine, F-91000 Evry Cedex 58 (FR).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **MALCUIT, Isabelle** [FR/GB]; 19 Catherine Place, London SW1E 6DX (GB). **SOROKIN, Alexander** [GB/GB]; 19 Catherine Place, London SW1E 6DX (GB).

(74) Agent: **DR WALTHER WOLFF & CO**; 19 Catherine Place, London SW1E 6DX (GB).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

(88) Date of publication of the international search report:
22 July 2010

(54) Title: METHOD FOR THE TRANSFORMATION OF PLANT CELL PLASTIDS

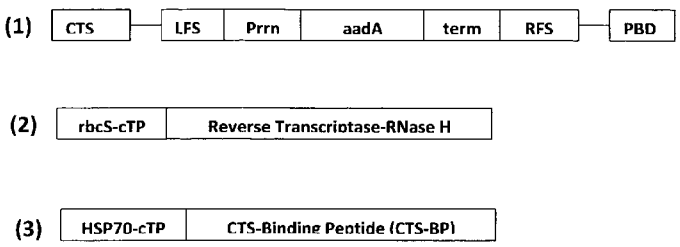


Figure 1

(57) Abstract: Method for heterologous protein production in plant cell plastids comprising introducing into plant cells nucleic acid components that encode heterologous proteins under the control of promoters operative in plastids, vectors, host cells, plants and uses thereof.

WO 2010/061186 A3

Polynucleotides for Transforming a Plastid in a Plant Cell, and for Producing a Cell or a Plant Containing the Transformed Plastid, and the Method Therefor

The present invention relates to a method for producing heterologous or exogenous RNA species in plant cell material such as genetically transformed plant cells in culture, plant tissue and plants derived from genetically transformed plant cells. In particular, the method relates to a more efficient method for producing\RNA species and/or heterologous or exogenous proteins in plastids comprised 'in plant cell material, the genetic material required therefor, such as DNA and RNA, vectors, host cells, methods of introduction of genetic material into plant cells, plant cells comprising genetically modified plastids, and uses thereof.

A disadvantage of prior art plant plastid transformation methods is that the transformation efficiency in terms of numbers of transformed plastids per cell tends to be low. A further disadvantage of prior art methods is that the delivery of genetic information into the plastid tends to be erratic in the sense that the delivery mechanisms employed rely on chance for the successful delivery of genetic information, such as RNA, into the plastid genome. Prior art methods do not rely on efficient endogenous cellular processes for the transfer of RNA into the plastid genome, subsequent reverse transcription and recombination of it within the plastid genome, and where desired, followed by expression of protein of interest therefrom. As such, prior art processes for genetically modifying plastids appear inefficient. These and other disadvantages of prior art plastid transformation technology will become apparent from the foregoing description.

The present inventors have found that by using or adapting endogenous cellular processes for the transfer of polynucleotide sequences, such as RNAs, from the cytoplasm to the plastid in the plant cell, polynucleotide sequences derived from nuclear transformation of the nucleus of a plant cell can be efficiently transferred or targeted to the plastid genome within a plant cell that is so transformed, and expressed more efficiently in the

plastid as described herein. Furthermore, it is apparent that once the plastid is transformed with sequences of the invention, it is not necessary for the nuclear encoded trangenes that are required for the initial transformation of the plastid to remain in the nuclear genome. As a consequence, the nuclear encoded transgenes can be removed through deliberate or natural segregation in subsequent generations of plants. For the purposes of the present invention the terms "plastid" and "plastids" and "plastid population" are used interchangeably, as are the terms "plant cell" and "plant cells", unless context demands otherwise. By employing or adapting endogenous cellular processes for the transfer of RNA derived from polynucleotide sequences introduced to the nucleus to the plastid genome, as described herein, the method of the invention is considered to be unique over prior art methods for the generation of plant cells or plants possessing genetically modified plastids. The plastid population of the plant cell is constantly bombarded by RNA that is derived from the nucleus of the cell, which is carried over the plastid membrane and into the plastid where it is reverse transcribed, integrated into the genome and then transcribed, resulting in the generation of RNA from which proteins of interest may be expressed.

There exists a need for a more efficient plastid transformation method for the production of RNAs, and where required, proteins of interest in the plastids of transformed plant cells and plant tissue derived therefrom.

The basis for the present invention, which does not appear to have been realised in the prior art, is the supply of a plant plastid transformation unit comprising nucleic acid sequences that encode: i) a plant plastid transformation unit (PTU); ii) a reverse transcriptase fused to a plant-derived chloroplast transit peptide sequence; and iii) an RNA binding protein fused to a plant plastid transit peptide. Such plastid fusion systems do not appear to have been described or alluded to in the prior art. Further simplified modifications of this kind of plant plastid transformation unit

include those that comprise nucleic acid sequences that encode i) a plant plastid transformation unit [PTU, for example, a chloroplast transformation unit (CTU)]; a plant plastid translocation sequence (PPS-5'), for example, a chloroplast translocation sequence (CTS-5'), fused to the 5' end of the PTU; a further plant plastid translocation sequence (PPS-3'), for example a chloroplast translocation sequence (CTS-3') fused to the 3'- end of the CTU; and a primer binding domain designed for reverse transcription in plastids using plastid tRNA-Met, such as chloroplast tRNA-Met(PBD-CHL). By placing the PBD-CHL next to the 3' end of the CTS-3', that is to say, outside of the LtrB intron as depicted in Figure 3(A), the LtrA protein is able to function as both a translocation protein and as a source of reverse transcriptase. In such a variant, there is no need to introduce a second gene for reverse transcriptase functionality. In a second variant of this system, where the PBD (PBD-CYT) is designed to interact with endogenous cytoplasmic tRNA-Met, the PBD may be located adjacent to the 3'-end of the PTU (or preferably, a CTU) and a plastid translocation sequence, preferably a chloroplast translocation sequence, is fused to it downstream. In this second variant, where a PBD is employed that is able to bind with cytoplasmic tRNA-Met as primer, reverse transcription is initiated by endogenous reverse transcriptase in the cytoplasm using cytoplasmic tRNA-Met. Thus, the second variant of the system does not require the co-delivery of a reverse transcriptase nucleic acid sequence to the plastids, such as chloroplasts. The use of such plastid transformation systems provides for an improved yield of RNA and hence protein of interest from plastid sources than has been hitherto achievable in the prior art.

According to the present invention there is provided a method of transforming a plant cell that comprises:

1) introducing into the said plant cell a nucleic acid sequence that comprises a plant nuclear promoter operably linked to a first nucleic acid sequence that comprises a plastid transgene cassette, a plastid translocation sequence (PTS), and a primer binding domain (PBD);

2) introducing into the said plant cell a second nucleic acid sequence that encodes for a plastid translocation sequence binding protein fused to a first plastid transit peptide (PTSBP-TP) wherein said second nucleic acid sequence is operably linked to a plant nuclear promoter; and

3) introducing into the said plant cell a third nucleic acid sequence that encodes for a reverse transcriptase protein fused to a second plastid transit peptide wherein the third nucleic acid sequence is operably linked to a plant nuclear promoter that drives expression in a plant cell nucleus.

The word "plastid" for the purposes of the present invention encompasses chloroplasts, proplastids, etioplasts, chromoplasts, amyloplasts, leucoplasts and elaioplasts. Preferably, "plastid" refers to chloroplasts. For the purposes of the description, the terms "chloroplast" and "chloroplasts" are used interchangeably unless context demands otherwise, as are the terms "plastid" and "plastids".

The skilled addressee will appreciate that where there are native proteins present in a plant cell that are capable of binding to a plastid translocation sequence, such as a chloroplast translocation sequence, and which are capable of translocating RNA nucleic acid sequences to the plastid, such as viroid proteins, the method of the invention for transforming a plant cell comprises:

1) introducing into the said plant cell a nucleic acid sequence that comprises a plant nuclear promoter operably linked to a first nucleic acid sequence that comprises a plastid transgene cassette, a plastid translocation sequence (PTS), and a primer binding domain (PBD); and

2) introducing into the said plant cell a second nucleic acid sequence that encodes for a reverse transcriptase protein fused to a second plastid transit peptide wherein the said second nucleic acid sequence is operably linked to a plant nuclear promoter that drives expression in a plant cell nucleus.

In a preferment of the above method, the plastid transgene cassette is a chloroplast transgene cassette, and the plastid translocation sequence (PTS) is a chloroplast translocation sequence (CTS), and the reverse transcriptase protein is a reverse transcriptase from a retrotransposon or a retrovirus which is fused to a chloroplast transit peptide for targeting into the chloroplast.

In a further aspect of the invention there is provided a method of transforming a plant cell that comprises introducing into the plant cell a nucleic acid sequence that comprises a plant nuclear promoter operably linked to a first nucleic acid sequence that comprises a first plastid translocation sequence (PTS-5') fused to the 5'-end of the plastid transgene unit (PTU), a second plastid translocation sequence (PTS-3')fused to the 3'end of the PTU, and a primer binding domain designed for reverse transcription in plastids, using tRNA-Met located within the plastids.

Preferably, the first plastid translocation sequence at the 5'-end is a chloroplast translocation sequence (CTS-5'), that is fused to a the 5'-end of a chloroplast transformation unit(CTU), and a second plastid translocation sequence is fused to the 3'-end of the CTU. Preferably, the second plastid translocation sequence is a chloroplast translocation sequence (CTS-3')that is fused to a primer binding domain that is designed for reverse transcription in chloroplast plastids (PBD-CHL), using tRNA-Met as primer that are located within the chloroplasts. The two plastid translocation sequences may be the same or different depending on design. In this variant reverse transcription can be effected when the PBD is located downstream of the CTU, that is to say 3' to a chloroplast translocation sequence (CTS-3'). Such a combination allows both translocation of the CTU into the chloroplast and reverse transcription of the CTU by the LtrA protein and does not require the co-delivery of a nucleic acid sequence for reverse transcriptase functionality.

In a still further variant of the invention, there is provided a method of transforming a plant cell that comprises introducing into the plant cell a nucleic acid sequence that comprises a plant

nuclear promoter operably linked to a first nucleic acid sequence that comprises a first plastid translocation sequence (PTS-5') fused to the 5'-end of the plastid transgene unit (PTU), a second plastid translocation sequence (PTS-3')fused to the 3'- end of a primer binding domain for binding tRNA-Met as primer that uses tRNA-Met that is located within the cytoplasm.

Preferably, the first plastid translocation sequence at the 5'-end is a chloroplast translocation sequence (CTS-5'), that is fused to the 5'-end of a chloroplast transformation unit (CTU). The second plastid translocation sequence is a chloroplast translocation sequence (CTS-3') that is fused to a primer binding domain that is capable of utilising native, endogenous reverse transcriptase located in the cytoplasm (PBD-CYT) for reverse transcription using cytoplasmic tRNA-Met as primer. Again, the two plastid translocation sequences may be the same or different depending on design. In this variant, there is also no need to co-deliver a nucleic acid sequence to the chloroplasts for reverse transcriptase functionality.

As another aspect of the invention there is provided a plant cell obtained by any one of the methods of the invention as described herein above.

In a further aspect of the invention there is provided a method of producing at least a heterologous or exogenous RNA species in a plant that comprises:

1) introducing into a regenerable plant cell a nucleic acid sequence that comprises a plant nuclear promoter operably linked to a first nucleic acid sequence that comprises a plastid transgene cassette, a plastid translocation sequence (PTS), and a primer binding domain (PBD);

2) introducing into the said regenerable plant cell a second nucleic acid sequence that encodes for a plastid translocation sequence binding protein fused to a first plastid transit peptide (PTSBP-TP) wherein said second nucleic acid sequence is operably linked to a plant nuclear promoter; and

3) introducing into the said regenerable plant cell a third nucleic acid sequence that encodes for a reverse transcriptase protein fused to a second plastid transit peptide wherein the third nucleic acid sequence is operably linked to a plant nuclear promoter that drives expression in a plant cell;

4) growing said regenerable plant cell of steps 1) to 3);

5) selecting a plant cell of (4) wherein the transgene comprised within the plastid transgene cassette is integrated into the plastid genome;

6) regenerating a plant from the plant cell of (5); and

7) growing the plant of (6).

Preferably, the plant obtained according to the above method is grown under conditions wherein the said heterologous or exogenous RNA species encoded by the transgene integrated into the plastid is expressed as heterologous or exogenous protein.

Again, and with reference to the method of obtaining a plant above, the skilled addressee will appreciate that where there are native proteins present in a plant cell that are capable of binding to a plastid translocation sequence, such as a chloroplast translocation sequence, and which are capable of translocating RNA nucleic acid sequences to the plastid, such as viroid proteins, step 2) of the said method may be omitted. In such an instance, there is provided a a method of producing at least a heterologous or exogenous RNA species in a plant that comprises:

1) introducing into a regenerable plant cell a nucleic acid sequence that comprises a plant nuclear promoter operably linked to a first nucleic acid sequence that comprises a plastid transgene cassette, a plastid translocation sequence (PTS), and a primer binding domain (PBD);

2) introducing into the said regenerable plant cell a second nucleic acid sequence that encodes for a reverse transcriptase protein fused to a second plastid transit peptide wherein the

second nucleic acid sequence is operably linked to a plant nuclear promoter that drives expression in a plant cell;

3) growing said regenerable plant cell of steps 1)and 2);

4) selecting a plant cell of (3) wherein the transgene comprised within the plastid transgene cassette is integrated into the plastid genome;

5) regenerating a plant from the plant cell of (4); and

6) growing the plant of (5).

In a further aspect of the invention there is provided a method of producing at least a heterologous or exogenous RNA species in a plant that comprises:

1)introducing into a regenerable plant cell a nucleic acid sequence that comprises a plant nuclear promoter operably linked to a first nucleic acid sequence that comprises a first plastid translocation sequence (PTS-5') fused to the 5'-end of the plastid transgene unit (PTU), a second plastid translocation sequence (PTS-3')fused to the 3'-end of the PTU, and a primer binding domain designed for reverse transcription in plastids;

2) growing said regenerable plant cell of step 1);

3) selecting a plant cell of (2) wherein the transgene comprised within the plastid transgene cassette is integrated into the plastid genome;

4) regenerating a plant from the plant cell of (3); and

5) growing the plant of (4).

Preferably, the first plastid translocation sequence at the 5'-end is a chloroplast translocation sequence (CTS-5'), that is fused to the 5'-end of a chloroplast transformation unit(CTU), and a second plastid translocation sequence is fused to the 3'-end of the CTU. Preferably, the second plastid translocation sequence is a chloroplast translocation sequence (CTS-3')that is fused to a primer binding domain that is designed for reverse transcription

in chloroplast plastids (PBD-CHL), using tRNA-Met as primer that are located within the chloroplasts. The two plastid translocation sequences may be the same or different depending on design.

In a further variant of this aspect of the invention there is provided a method of producing at least a heterologous or exogenous RNA species in a plant that comprises:

1) introducing into a regenerable plant cell a nucleic acid sequence that comprises a plant nuclear promoter operably linked to a first nucleic acid sequence that comprises a first plastid translocation sequence (PTS-5') fused to the 5'-end of the plastid transgene unit (PTU), a second plastid translocation sequence (PTS-3') fused to the 3'- end of a primer binding domain for binding tRNA-Met as primer that uses tRNA-Met that is located within the cytoplasm.

2) growing said regenerable plant cell of step 1);

3) selecting a plant cell of (2) wherein the transgene comprised within the plastid transgene cassette is integrated into the plastid genome;

4) regenerating a plant from the plant cell of (3); and

5) growing the plant of (4).

Preferably, the first plastid translocation sequence at the 5'-end is a chloroplast translocation sequence (CTS-5'), that is fused to the 5'-end of a chloroplast transformation unit (CTU). The second plastid translocation sequence is a chloroplast translocation sequence (CTS-3') that is fused to a primer binding domain that is capable of utilising native, endogenous reverse transcriptase located in the cytoplasm (PBD-CYT) for reverse transcription using cytoplasmic tRNA-Met as primer. Again, the two plastid translocation sequences may be the same or different depending on design.

Naturally, the person skilled in the art will understand that the plant nuclear promoter by being operably linked to the nucleic

acid sequences provided for herein drives expression of such sequences in the plant nucleus.

The "plastid transgene cassette" comprises a left flanking sequence (LFS) and a right flanking sequence (RFS) which are used for homologous recombination of the cassette into the plastid genome. In between the LFS and RFS are located at least one plastid specific promoter sequence (such as a chloroplast specific promoter, e.g. Prrn) and at least one plastid specific terminator sequence (such as a chloroplast specific terminator, e.g. 3'UTR sequence of psbA gene from tobacco) which in turn flanks at least one isolated gene or isolated nucleic acid sequence of interest, such as a recombinant DNA sequence (e.g. cDNA) or an introduced native DNA sequence. The LFS and RFS may include the chloroplast specific promoter and terminator sequences, respectively, if for example, the isolated nucleic acid of interest is fused to a native chloroplast nucleic acid of interest. Thus, the promoter and the terminator sequences are not necessarily included within the LFS or RFS, respectively *per se*, or between the LFS and RFS if a transgene is inserted into the chloroplast genome as a cistron unit or if a transgene is translationally fused to a native gene. In such an instance, when a transgene is fused to a native chloroplast coding sequence it is after the transformation event has taken place that the promoter may be found upstream of the sequence that is homologous to the LFS in the chloroplast genome and is available to drive expression of the gene fused to the transgene of interest. For the purposes of the present invention "transgene" includes isolated nucleic acid sequences that may ultimately give rise to the expression of proteins or peptides of interest in the plastid (e.g. chloroplast) as herein described. Thus, the isolated nucleic acid sequence may be one that gives rise to an RNA sequence of interest which may not encode or give rise to the expression of a translatable product, or the isolated nucleic acid sequence may give rise to an RNA sequence that does encode or give rise to the expression of a translatable product such as a protein or peptide of interest. The person skilled in the art will also appreciate that the transgene that is carried on the isolated nucleic acid may also be designed to give rise to an

RNA sequence that gives rise to the expression of a translatable product or products, and untranslatable RNAs. Such RNAs that do not give rise to the expression of proteins may give rise to RNA sequences that contain deletions or other mutations and these may find use as research tools for studying gene function in the plastid, e.g. chloroplast. Where the "transgene" gives rise to the expression of proteins or peptides, suitable transgenes of interest include plant proteins capable of conferring desired traits to plant crops, and pharmaceutical proteins for use in mammals, including man, such as insulin, preproinsulin, proinsulin, glucagon, interferons such as α-interferon, β-interferon, γ-interferon, blood-clotting factors selected from Factor VII, VIII, IX, X, XI, and XII, fertility hormones including luteinising hormone, follicle stimulating hormone growth factors including epidermal growth factor, platelet-derived growth factor, granulocyte colony stimulating factor and the like, prolactin, oxytocin, thyroid stimulating hormone, adrenocorticotropic hormone, calcitonin, parathyroid hormone, somatostatin, erythropoietin (EPO), enzymes such as β-glucocerebrosidase, haemoglobin, serum albumin, collagen, biotic and abiotic stress proteins, such as insecticidal and insect toxic proteins, for example from, or derived from Bacillus thuringiensis, nematicidal proteins, herbicide resistance proteins, (e.g. to glyphosate), salt-tolerance proteins, drought tolerant proteins, nutritional enhancement proteins involved in the biosynthesis of phenolics, starches, sugars, alkaloids, vitamins, and edible vaccines, and the like. Furthermore, the method of the invention can be used for the production of specific monoclonal antibodies or active fragments thereof and of industrial enzymes or active fragments thereof.

All proteins mentioned hereinabove are of the plant and human type. Other proteins that are contemplated for production in the present invention include proteins for use in veterinary care and may correspond to animal homologues of human proteins, such as the human proteins mentioned hereinabove.

In a further aspect of the invention there is provided a plant cell that comprises plastids, such as chloroplasts, that are permanently transformed with an exogenous or a heterologous nucleic acid sequence that encodes for a protein or RNA of interest. Suitable proteins and peptides and nucleic acids of interest are provided herein.

The LFS and RFS may be selected from any nucleotide sequences that may be used for homologous recombination in the plastid. Suitable examples include coding sequences such as the sequence coding for psbA, rbcL genes from chloroplasts.

The plant plastid promoter may be selected from the group consisting of the RNA polymerase promoter, rpo B promoter element, atpB promoter element, the clpP promoter element, the 16S rDNA promoter element, PrbcL, Prps16, the Prrn16, Prrn-62, Pycf2-1577, PatpB-289, Prps2-152, Prps16-107, Pycf1-41, PatpI-207, PclpP-511, PclpP-173, PaccD-129, PaccD-129 promoter of the tobacco accD gene, the PclpP-53 promoter of the clpP gene, the Prrn-62 promoter of the rrn gene, the Prps16-107 promoter of the rps16 gene, the PatpB/E-290 promoter of the tobacco atpB/E gene, and the PrpoB-345 promoter of the rpoB gene. Furthermore, all those promoters which belong to class III (Hajdukiewicz P T J et al. (1997) EMBO J 16:4041-4048) and all fragments of the class II promoters which control the initiation of transcription by NEP may be utilized in the method of the invention. Such promoters or promoter moieties are not generally known to be highly conserved. ATAGAATAAA is given as consensus near the transcription initiation site of NEP promoters (Hajdukiewicz P T J et al (1997) EMBO J 16:4041-4048).

The plant plastid terminator, such as a chloroplast transcription terminator may be selected from any plastid terminator such as psbA, atpA, rbcL 3'-UTR region, and bacterial transcription terminators such as rrnB described by Orosz A., et al., Eur. J. Biochemistry, 2005, Volume 201, Issue 3, pp 653-659.

Naturally, the man skilled in the art will appreciate that other terminator DNA sequences may be present in constructs used in the invention.

The plant plastid (e.g. chloroplast) transgene cassette also comprises a primer binding domain (PBD) that once inside the plastid (e.g. chloroplast) is able to capture tRNAs as primers to form template RNA to initiate reverse transcription of introduced plant chloroplast transformation units of the invention. A suitable tRNA for use in the present invention as a primer is tRNA-fMet which forms a template RNA ready for reverse transcription. The skilled person in the art will appreciate that PBDs are found naturally on retroelements including retroviruses and retrotransposons. PBDs comprise specific RNA domains that anneal to specific sequences on tRNA molecules. The tRNA itself does not serve as a PBD but as a primer for reverse transcription, the template for reverse transcription is the RNA molecule that carries a PBD. Novel PBDs can be readily engineered that can anneal to other tRNAs. PBDs can be designed to bind other types of tRNAs such as, tRNA-lys and tRNA-Met of tobacco and others which are known in the art (http://www.unibayreuth.de/departments/biochemie/trna/).

Certain elements of retroelements such as retroviruses or retrotransposons, have native PBDs possessing conserved domains that anneal with complementary domains from tRNA (usually tRNA-met, or tRNA-trp); because of the conserved structures of all tRNAs (the so-called clover-leaf structure), PBDs can be engineered so that they carry specific domains that will anneal with a tRNA of choice.

A "plastid translocation sequence" (PTS, for example a chloroplast translocation sequence (CTS))is an RNA sequence that is capable of being bound to a plant PTS binding protein and thereby, the PTS and other RNA sequences that may be associated with it or fused with it can be transported across and into the plastid (e.g. chloroplast). The CTS can be selected from naked RNA

viruses, including viral RNAs such as those from positive stranded RNA viruses such as potato virus X (PVX), tobacco mosaic virus (TMV), tomato mosaic virus (ToMV), and viral RNAs from negative stranded RNA viruses, such as tomato spotted wilt virus (TSWV) and Impatiens necrotic spotted virus (INSV), viroids such as peach latent mosaic viroid (PLMVd)or avocado sunblotch viroid (ASBV), satellite viruses such as satellite tobacco mosaic virus (STMV) and the like. Other sources of the PTS/CTS include group I and group II intron RNAs or modified versions thereof in which cryptic splicing sites have been eliminated that may be derived from a bacterium, a fungus or a plastid/chloroplast from a plant, such as an LTRB intron lacking the sequence coding for LTRA (the protein encoded by an LTRA sequence being capable of serving as an PTS/CTS-binding protein in the method of the invention).

Preferably, the intron is a group II intron, such as the Lactococcus lactis Ll.ltrB intron or a modified version of it in which cryptic splicing sites have been eliminated as outlined herein. Group II introns are widely represented in the organelles of plants and fungi, and in bacteria. Group II introns useful in the method of the invention are mobile, highly structural retroelements that encode multifunctional protein (intron encoded protein or IEP) which possesses reverse transcriptase (RT) activity. The IEP facilitates splicing of intron RNA by stabilization of the catalytically active RNA structure, performs reverse transcription and insertion of the intron into specific DNA target sites of the bacterial genome at high frequency (Moran et al. (1995) Mol Cell Biol 15:2828-2838; Cousineau et al. (1998) Cell 94:451-462).

Group II introns of bacterial origin, such as those derived from Lactococcus that comprise a modified LtrA gene, are preferably used in the method of the invention. The LtrA polynucleotide sequence of a *Lactococcus* bacterium, such as *Lactococcus lactis* may be modified for optimum expression in plants by inserting into it at least one polynucleotide sequence comprising one or more introns from at least one plant nucleic acid sequence, such as from one or more plant genes and by substituting certain selected

codons having a low frequency of usage in native plants with codons that occur with a higher frequency in such plants. Typically, the bacterial LtrA sequence of interest is analysed with reference to plant codon usage using *in silico* comparisons such as those found at the website www.kazusa.or.jp/codon for bacterial codons that occur with low frequency in plants. Such codons may then be substituted with codons that have a high frequency of occurrence in plants, and an *in silico*-derived modified polynucleotide sequence is generated. From this optimised LtrA sequence a synthetic LtrA polynucleotide sequence corresponding to the *in silico* generated sequence is made using standard polynucleotide synthesis procedures known in the art, and may then be used in the preparation of constructs of use in the present invention as outlined herein. It is thought that by using a modified sequence that comprises plant codon substitutions as outlined above more plant cell environment stable polynucleotide RNA sequences are generated.

Other types of introns that may be used in the method of the invention include, for example, the group I intron from Tetrahymena (GenBank Acc. No.: X54512; Kruger K et al. (1982) Cell 31:147-157; Roman J and Woodson S A (1998) Proc Natl Acad Sci USA 95:2134-2139), the group II rIl intron from Scenedesmus obliquus (GenBank Acc. No.: X17375.2 nucleotides 28831 to 29438; Hollander V and Kuck U (1999) Nucl Acids Res 27: 2339-2344; Herdenberger F et al. (1994) Nucl Acids Res 22: 2869-2875; Kuck U et al. (1990) Nucl Acids Res 18:2691-2697), and the Ll.LtrB intron (GenBank Acc. No.: U50902 nucleotides 2854 to 5345).

Aside from heterologous introns described herein, endogenous introns that occur naturally in the plastid, for example, in the chloroplast, such as group II introns from plant chloroplasts, for example the atpF, rpl, trnA, trnI, trnK, petD, petB (Jenkins B.D. et al., The Plant Cell, Vol. 9, 283-296, March 1997).

Introns which occur naturally in the plastids, such as chloroplasts of the plant of interest may be modified such that they have a sequence homology of about 50%, 60%, 70%, 75%, 80%,

85%, 90% or 95%, or of any percentage sequence homology therebetween, with the sequence of the starting intron, while retaining functionality, may also be employed in the method of the invention. Other MTS include RNA domains found on tobacco TNT1, yeast Ty1- and Ty3-like retrotransposons or other RNA that harbours a domain that is recognised by an RNA binding protein that is driven into the chloroplasts.

A "plastid translocation sequence binding protein" (PTS-BP, for example, a CTS-BP) can be any RNA binding protein that recognises and binds to specific RNA domains of interest and is fused to a plastid transit peptide, such as a chloroplast transit peptide. Examples of suitable PTS-BP/CTS-BP proteins may be selected from the Ltra protein from the group II intron II LtrB, coat proteins that bind to RNA viruses such as the coat protein from potato virus X (PVX), the coat protein of TMV, RNA-dependent RNA polymerases (RdRpS) of RNA viruses such as the replicases of PVX or TMV, native plant proteins that are responsible for translocation of viroid RNA (such as PLMVd and ASBV viroids) into the chloroplasts, reverse transcriptase protein from retrotransposons, such as tobacco TnT1, yeast Tyl-1 which recognise structures on the retrotransposon RNA molecule, and proteins that bind to cellular RNAs. Preferably, the PTS-BP protein/CTS-BP protein is the LrtA protein from the group II intron LltrB.

A "plant chloroplast transit peptide" (TP) is one that may be derived or obtained from a plastid-targeted protein, for example transit peptide from small subunit of Rubisco (rbcS) or HSP70 proteins (Marshall & Keegstra (1992) Plant Physiology, 100, 1048-1054), and those that may be predicted by chloroplast localisation sequences programmes (http://www.psort.org).

The "reverse transcriptase" protein, if employed, may be selected from a retrovirus source, such as from plant retroviruses such as SIRE-1 from soybean, or from a retrotransposon source such as from the yeast Tyl1 retrotransposon, for example the reverse transcriptase-RNaseH domain (Goffeau et al., Science 274 (5287),

546-547 (1996)) or the tobacco TnT1 retrotransposon (RTRH domain) (Vernhettes et., al.; Mol. Biol. Evol. 15 (7), 827-836 (1998)).

A plant nuclear promoter (for example, an exogenous nucleus specific promoter) is one that is able to drive expression of a nucleic acid sequence such as a cDNA sequence or a full length gene sequence in the nucleus of a plant cell, forming a transcribed RNA sequence. The plant nuclear promoter is one that is introduced in front of a nucleic acid sequence of interest and is operably associated therewith. Thus a plant nuclear promoter is one that has been placed in front of a selected polynucleotide component. Typically, a plant nuclear promoter, such as an exogenous nucleus specific promoter, is one that is transferred to a host cell or host plant from a source other than the host cell or host plant.

The cDNAs encoding a polynucleotide of the invention contain at least one type of nucleus specific promoter that is operable in a plant cell, for example, an inducible or a constitutive promoter operatively linked to a first and/or second nucleic acid sequence or nucleic acid sequence component as herein defined and as provided by the present invention. As discussed, this enables control of expression of polynucleotides of the invention. The invention also provides plants transformed with polynucleotide sequences or constructs and methods including introduction of such polynucleotide nucleic acid sequences or constructs into a plant cell and/or induction of expression of said first or second nucleic acid sequence or construct within a plant cell, e.g. by application of a suitable stimulus, such as an effective exogenous inducer.

The term "inducible" as applied to a promoter is well understood by those skilled in the art. In essence, expression under the control of an inducible promoter is "switched on" or increased in response to an applied stimulus (which may be generated within a

cell or provided exogenously). The nature of the stimulus varies between promoters. Some inducible promoters cause little or undetectable levels of expression (or no expression) in the absence of the appropriate stimulus. Other inducible promoters cause detectable constitutive expression in the absence of the stimulus. Whatever the level of expression is in the absence of the stimulus, expression from any inducible promoter is increased in the presence of the correct stimulus. The preferable situation is where the level of expression increases upon application of the relevant stimulus by an amount effective to alter a phenotypic characteristic. Thus an inducible (or "switchable") promoter may be used which causes a basic level of expression in the absence of the stimulus which level is too low to bring about a desired phenotype (and may in fact be zero). Upon application of the stimulus, expression is increased (or switched on) to a level, which brings about the desired phenotype. One example of an inducible promoter is the ethanol inducible gene switch disclosed in Caddick et al (1998) Nature Biotechnology 16: 177-180. A number of inducible promoters are known in the art.

Chemically regulated promoters can be used to modulate the expression of a gene or a polynucleotide sequence of the invention in a plant through the application of an exogenous chemical regulator. Depending upon the objective, the promoter may be a chemically inducible promoter, where application of the chemical induces gene expression, or a chemical-repressible promoter, where application of the chemical represses gene expression. Chemically inducible promoters are known in the art and include, but are not limited to, the maize In2-2 promoter, which is activated by benzenesulfonamide herbicide safeners, the maize GST promoter, which is activated by hydrophobic electrophilic compounds that are used as pre-emergent herbicides, and the tobacco PR-1a promoter, which is activated by salicylic acid. Other chemically regulated promoters of interest include steroid-responsive promoters (see, for example, the glucocorticoid-inducible promoter in Schena *et al.* (1991) *Proc. Natl. Acad. Sci. USA 88*:10421-10425 and McNellis *et al.* (1998) *Plant J. 14(2)*:247-257) and tetracycline-inducible and tetracycline-repressible promoters (see, for example, Gatz *et*

al. (1991) *Mol. Gen. Genet. 227*:229-237, and U.S. Patent Nos. 5,814,618 and 5,789,156).

Where enhanced expression in particular tissues is desired, tissue-specific promoters can be utilized. Tissue-specific promoters include those described by Yamamoto *et al.* (1997) *Plant J. 12(2)*255-265; Kawamata *et al.* (1997) *Plant Cell Physiol. 38(7)*:792-803; Hansen *et al.* (1997) *Mol. Gen Genet. 254(3)*:337-343; Russell *et al.* (1997) *Transgenic Res. 6(2)*:157-168; Rinehart *et al.* (1996) *Plant Physiol. 112(3)*:1331-1341; Van Camp *et al.* (1996) *Plant Physiol. 112(2)*:525-535; Canevascini *et al.* (1996) *Plant Physiol. 112(2)*:513-524; Yamamoto *et al.* (1994) *Plant Cell Physiol. 35(5)*:773-778; Lam (1994) *Results Probl. Cell Differ. 20*:181-196; Orozco *et al.* (1993) *Plant Mol Biol. 23(6)*:1129-1138; Matsuoka *et al.* (1993) *Proc Natl. Acad. Sci. USA 90(20)*:9586-9590; and Guevara-Garcia *et al.* (1993) *Plant J. 4(3)*:495-505.

So-called constitutive promoters may also be used in the methods of the present invention. Constitutive promoters include, for example, CaMV 35S promoter (Odell *et al.* (1985) *Nature 313*:810-812); rice actin (McElroy *et al.* (1990) *Plant Cell 2*:163-171); ubiquitin (Christensen *et al.* (1989) *Plant Mol. Biol. 12*:619-632 and Christensen *et al.* (1992) *Plant Mol. Biol. 18*:675-689); pEMU (Last *et al.* (1991) *Theor. Appl. Genet. 81*:581-588); MAS (Velten *et al.* (1984) *EMBO J. 3*:2723-2730); ALS promoter (U.S. Application Serial No. 08/409,297), and the like. Other constitutive promoters include those in U.S. Patent Nos. 5,608,149; 5,608,144; 5,604,121; 5,569,597; 5,466,785; 5,399,680; 5,268,463; and 5,608,142. In a preferment, the plant nuclear promoter used in the method of the invention is a constitutive promoter.

The expression in the plastid, such as in the chloroplast, is effected by employing a plant plastid promoter such as plastid specific promoters and/or transcription regulation elements. Examples include the RNA polymerase promoter (WO 97/06250) and other promoters described in the art, eg in WO 00/07431, U.S. Pat. No. 5,877,402, WO 97/06250, WO 98/55595, WO 99/46394, WO 01/42441

and WO 01/07590; the rpo B promoter element, the atpB promoter element, the clpP promoter element (see also WO 99/46394) and the 16S rDNA promoter element. The plastid specific promoter may also have a polycistronic "operon" assigned to it (EP-A 1 076 095; WO 00/20611). Further promoters that may be used in the method of the invention also include the PrbcL promoter, the Prps16 promoter, and the Prrn16 promoter described in US Patent application 2006/0253916, the plastid specific promoters Prrn-62, Pycf2-1577, PatpB-289, Prps2-152, Prps16-107, Pycf1-41, PatpI-207, PclpP-511, PclpP-173 and PaccD-129 (WO 97/06250; Hajdukiewicz P T J et al. (1997) EMBO J 16:4041-4048), the PaccD-129 promoter of the tobacco accD gene (WO 97/06250), the PclpP-53 promoter of the clpP gene as highly active NEP promoter in chloroplasts (WO 97/06250), the Prrn-62 promoter of the rrn gene, the Prps16-107 promoter of the rps16 gene, the PatpB/E-290 promoter of the tobacco atpB/E gene (Kapoor S et al. (1997) Plant J 11:327-337), and the PrpoB-345 promoter of the rpoB gene (Liere K & Maliga P (1999) EMBO J 18: 249-257). Furthermore, all those promoters which belong to class III (Hajdukiewicz P T J et al. (1997) EMBO J 16:4041-4048) and all fragments of the class II promoters which control the initiation of transcription by NEP may be utilized in the method of the invention. Such promoters or promoter moieties are not generally known to be highly conserved. ATAGAATAAA is given as consensus near the transcription initiation site of NEP promoters (Hajdukiewicz P T J et al (1997) EMBO J 16:4041-4048).

Naturally, the man skilled in the art will appreciate that other terminator DNA sequences may be present in constructs used in the invention. A terminator is contemplated as a DNA sequence at the end of a transcriptional unit which signals termination of transcription. These elements are 3'-non-translated sequences containing polyadenylation signals, which act to cause the addition of polyadenylate sequences to the 3' end of primary transcripts. For expression in plant cells the nopaline synthase transcriptional terminator (A. Depicker et al., 1982, J. of Mol. & Applied Gen. 1:561-573) sequence serves as a transcriptional termination signal.

In another aspect of the invention there is provided a plastid transformation sequence that comprises:

i) a plant plastid translocation sequence;

ii) a plastid transgene cassette; and

iii) a primer binding domain.

The plant plastid translocation sequence and the primer binding domain are as defined herein.

The plastid transgene cassette comprises a left flanking sequence (LFS) and a right flanking sequence (RFS) as herein described, and may include a promoter region and/or a terminator region sourced from a higher or lower plant plastid, such as a chloroplast, for example from tobacco, arabidopsis, brassica sp., potato, corn(maize), canola, rice, wheat, barley, brassica sp., cotton, algae (e.g. blue green species), lemnospora ("duckweed"), or moss (e.g. *physcomitrella patens*). Preferably, the promoter and terminator regions are sourced from higher plant species. Where the LFS and RFS do not include a promoter and/or a terminator region, these components may be placed adjacent to the LFS and/or RFS, as appropriate, or there may be a spacer region therein between. Included within the plastid transgene cassette is at least one transgene or one nucleotide sequence of choice that is destined to be transcribed and/or translated in the chloroplast in accordance with the design of the method of the present invention for example, for the production of desired protein(s), RNAs of interest, or knockout of endogenous plastidal genes and regulatory sequences. Suitable transgenes of interest contemplated for protein or peptide production in a method of the present invention include plant proteins and pharmaceutical proteins for use in mammals, including man, such as insulin, preproinsulin, proinsulin, glucagon, interferons such as α-interferon, β-interferon, γ-interferon, blood-clotting factors selected from Factor VII, VIII, IX, X, XI, and XII, fertility hormones including luteinising hormone, follicle stimulating hormone growth factors including epidermal growth factor, platelet-derived growth factor,

granulocyte colony stimulating factor and the like, prolactin, oxytocin, thyroid stimulating hormone, adrenocorticotropic hormone, calcitonin, parathyroid hormone, somatostatin, erythropoietin (EPO), enzymes such as β-glucocerebrosidase, haemoglobin, serum albumin, collagen, insect toxic protein from Bacillus thuringiensis; herbicide resistance protein (glyphosate); salt-tolerance proteins; proteins involved in conferring cytoplasmic male sterility to plant breeding lines; nutritional enhancement proteins involved in the biosynthesis of phenolics, starches, sugars, alkaloids, vitamins, and edible vaccines, and the like. Furthermore, the method of the invention can be used for the production of specific monoclonal antibodies or active fragments thereof and of industrial enzymes.

All proteins mentioned hereinabove are of the plant and human type. Other proteins that are contemplated for production in the present invention include proteins for use in veterinary care and may correspond to animal homologues of human proteins, such as the human proteins mentioned hereinabove.

In a further aspect of the invention there is provided a plant cell that comprises plastids, such as chloroplasts, that are permanently transformed with an exogenous or a heterologous nucleic acid sequence that encodes for a protein of interest. Suitable proteins and peptides of interest may be selected from those provided herein. Accordingly, there is also provided a plant derived from a plant cell as described herein.

Naturally, the person skilled in the art will appreciate that where nuclear terminator DNA sequences will be present in constructs used in the methods of the invention, these are contemplated as comprising a DNA sequence at the end of a transcriptional unit which signals termination of transcription. These elements are 3'-non-translated sequences containing polyadenylation signals, which act to cause the addition of polyadenylate sequences to the 3' end of primary transcripts. For expression in plant cells the nopaline synthase transcriptional

terminator (A. Depicker et al., 1982, J. of Mol. & Applied Gen. 1:561-573) sequence serves as a transcriptional termination signal.

Those skilled in the art are well able to construct vectors and design protocols for recombinant nucleic acid sequences or gene expression. Suitable vectors can be chosen or constructed, containing appropriate regulatory sequences, including promoter sequences, terminator fragments, polyadenylation sequences, enhancer sequences, marker genes and other sequences as appropriate. For further details see, for example, *Molecular Cloning: a Laboratory Manual:* 2nd edition, Sambrook *et al,* 1989, Cold Spring Harbor Laboratory Press. Many known techniques and protocols for manipulation of nucleic acid, for example in preparation of nucleic acid constructs, mutagenesis, sequencing, introduction of DNA into cells and gene expression, and analysis of proteins, are described in detail in *Current Protocols in Molecular Biology,* Second Edition, Ausubel et al. eds., John Wiley & Sons, 1992.

Specific procedures and vectors previously used with wide success upon plants are described by Bevan (Nucl. Acids Res. 12, 8711-8721 (1984)) and Guerineau and Mullineaux (1993) (Plant transformation and expression vectors. In: Plant Molecular Biology Labfax (Croy RRD ed.) Oxford, BIOS Scientific Publishers, pp 121-148).

Naturally, the skilled addressee will appreciate that each introduced transgene in a transgene cassette will be under regulatory control of its own exogenous plastidal promoter, for example a chloroplast promoter and terminator. When two or more target proteins are destined to be produced from a single carrier RNA it is preferable if they are able to be readily separated, for example by binding to different protein-specific antibodies (monoclonal or polyclonal) in the harvesting phase of the plant cell culture system.

Selectable genetic markers may facilitate the selection of transgenic plants and these may consist of chimaeric genes that confer selectable phenotypes such as resistance to antibiotics such as spectinomycin, streptomycin, kanamycin, neomycin, hygromycin, puramycin, phosphinotricin, chlorsulfuron, methotrexate, gentamycin, spectinomycin, imidazolinones and glyphosate.

When introducing selected nucleic acid sequences according to the present invention into a cell, certain considerations must be taken into account, well known to those skilled in the art. The nucleic acid to be inserted should be assembled within a construct, which contains effective regulatory elements, which will drive transcription. There must be available a method of transporting the construct into the cell. Once the construct is within the cell, integration into the endogenous chromosomal material either will or will not occur. Finally, as far as plants are concerned the target cell type must be such that cells can be regenerated into whole plants.

Plants transformed with DNA segments containing sequences of interest as provided herein may be produced by standard techniques, which are already known for the genetic manipulation of plants. DNA can be transformed into plant cells using any suitable technology, such as a disarmed Ti-plasmid vector carried by Agrobacterium exploiting its natural gene transfer ability (EP-A-270355, EP-A-0116718, NAR 12(22) 8711 -87215 1984), particle or micro projectile bombardment (US 5100792, EP-A-444882, EP-A-434616) microinjection (WO 92/09696, WO 94/00583, EP 331083, EP 175966, Green et al. (1987) *Plant Tissue and Cell Culture*, Academic Press), electroporation (EP 290395, WO 8706614) other forms of direct DNA uptake (DE 4005152, WO 9012096, US 4684611), liposome mediated DNA uptake (e.g. Freeman et al. *Plant Cell Physiol.* **29**: 1353 (1984)), or the vortexing method (e.g. Kindle, *PNAS U.S.A.* **87**: 1228 (1990d) Physical methods for the transformation of plant cells are reviewed in Oard, 1991, *Biotech. Adv.* **9**: 1-11.

Thus once a nucleic acid sequence or gene has been identified, it may be reintroduced into plant cells using techniques well known to those skilled in the art to produce transgenic plants of the appropriate phenotype.

Agrobacterium transformation is widely used by those skilled in the art to transform dicotyledonous species. Production of stable, fertile transgenic plants in almost all economically relevant monocot plants is also now routine:(Toriyama, et al. (1988) *Bio/Technology* **6**, 1072-1074; Zhang, et al. (1988) *Plant Cell Rep.* **7**, 379-384; Zhang, et al. (1988) *Theor. Appl. Genet* **76**, 835-840; Shimamoto, et al. (1989) *Nature* **338**, 274-276; Datta, et al. (1990) *Bio/Technology* **8**, 736-740; Christou, et al. (1991) *Bio/Technology* **9**, 957-962; Peng, et al. (1991) International Rice Research Institute, Manila, Philippines 563-574; Cao, et al. (1992) *Plant Cell Rep.* **11**, 585-591; Li, et al. (1993) *Plant Cell Rep.* **12**, 250-255; Rathore, et al. (1993) *Plant Molecular Biology* **21**, 871-884; Fromm, et al. (1990) *Bio/Technology* **8**, 833-839; Gordon-Kamm, et al. (1990) *Plant Cell* **2**, 603-618; D'Halluin, et al. (1992) *Plant Cell* **4**, 1495-1505; Walters, et al. (1992) *Plant Molecular Biology* **18**, 189-200; Koziel, et al. (1993) *Biotechnology* **11**, 194-200; Vasil, I. K. (1994) *Plant Molecular Biology* **25**, 925-937; Weeks, et al. (1993) *Plant Physiology* **102**, 1077-1084; Somers, et al. (1992) *Bio/Technology* **10**, 1589-1594; WO92/14828). In particular, *Agrobacterium* mediated transformation is now a highly efficient alternative transformation method in monocots (Hiei et al. (1994) *The Plant Journal* **6**, 271-282).

The generation of fertile transgenic plants has been achieved in the cereals rice, maize, wheat, oat, and barley (reviewed in Shimamoto, K. (1994) *Current Opinion in Biotechnology* **5**, 158-162.; Vasil, et al. (1992) *Bio/Technology* 10, 667-674; Vain et al., 1995, *Biotechnology Advances* **13** (4): 653-671; Vasil, 1996, *Nature Biotechnology* **14** page 702). Wan and Lemaux (1994) *Plant Physiol.* **104**: 37-48 describe techniques for generation of large numbers of independently transformed fertile barley plants.

Micro projectile bombardment, electroporation and direct DNA uptake are preferred where Agrobacterium is inefficient or ineffective. Alternatively, a combination of different techniques may be employed to enhance the efficiency of the transformation process, e.g. bombardment with Agrobacterium coated micro particles (EP-A-486234) or micro projectile bombardment to induce wounding followed by co-cultivation with Agrobacterium (EP-A-486233).

Following transformation, a plant may be regenerated, e.g. from single cells, callus tissue or leaf discs, as is standard in the art. Almost any plant can be entirely regenerated from cells, tissues and organs of the plant. Available techniques are reviewed in Vasil et al., *Cell Culture and Somatic Cell Genetics of Plants, Vol. I, II and III, Laboratory Procedures and Their Applications*, Academic Press, 1984, and Weiss Bach and Weiss Bach, *Methods for Plant Molecular Biology*, Academic Press, 1989.

The particular choice of a transformation technology will be determined by its efficiency to transform certain plant species as well as the experience and preference of the person practising the invention with a particular methodology of choice. It will be apparent to the skilled person that the particular choice of a transformation system to introduce nucleic acid into plant cells is not essential to or a limitation of the invention, nor is the choice of technique for plant regeneration.

The invention further encompasses a host cell transformed with vectors or constructs as set forth above, especially a plant or a microbial cell. Thus, a host cell, such as a plant cell, including nucleotide sequences of the invention as herein

indicated is provided. Within the cell, the nucleotide sequence may be incorporated within the chromosome.

Also according to the invention there is provided a plant cell having incorporated into its genome at least a nucleotide sequence, particularly heterologous nucleotide sequences, as provided by the present invention under operative control of regulatory sequences for control of expression as herein described. The coding sequence may be operably linked to one or more regulatory sequences which may be heterologous or foreign to the nucleic acid sequences employed in the invention, such as those not naturally associated with the nucleic acid sequence(s) for its(their) expression. The nucleotide sequence according to the invention may be placed under the control of an externally inducible promoter to place expression under the control of the user. A further aspect of the present invention provides a method of making such a plant cell involving introduction of nucleic acid sequence(s) contemplated for use in the invention or a suitable vector including the sequence(s) contemplated for use in the invention into a plant cell and causing or allowing recombination between the vector and the plant cell genome to introduce the said sequences into the genome. The invention extends to plant cells containing a nucleotide sequence according to the invention as a result of introduction of the nucleotide sequence into an ancestor cell.

The term "heterologous" may be used to indicate that the gene/sequence of nucleotides in question have been introduced into said cells of the plant or an ancestor thereof, using genetic engineering, ie by human intervention. A transgenic plant cell, i.e. transgenic for the nucleotide sequence in question, may be provided. The transgene may be on an extra-genomic vector or incorporated, preferably stably, into the genome. A heterologous gene may replace an endogenous equivalent gene, ie one that normally performs the same or a similar function, or the inserted

sequence may be additional to the endogenous gene or other sequence. An advantage of introduction of a heterologous gene is the ability to place expression of a sequence under the control of a promoter of choice, in order to be able to influence expression according to preference. Furthermore, mutants, variants and derivatives of the wild-type gene, e.g. with higher activity than wild type, may be used in place of the endogenous gene. Nucleotide sequences heterologous, or exogenous or foreign, to a plant cell may be non-naturally occurring in cells of that type, variety or species. Thus, a nucleotide sequence may include a coding sequence of or derived from a particular type of plant cell or species or variety of plant, placed within the context of a plant cell of a different type or species or variety of plant. A further possibility is for a nucleotide sequence to be placed within a cell in which it or a homologue is found naturally, but wherein the nucleotide sequence is linked and/or adjacent to nucleic acid which does not occur naturally within the cell, or cells of that type or species or variety of plant, such as operably linked to one or more regulatory sequences, such as a promoter sequence, for control of expression. A sequence within a plant or other host cell may be identifiably heterologous, exogenous or foreign.

Plants which include a plant cell according to the invention are also provided, along with any part or propagule thereof, seed, selfed or hybrid progeny and descendants. Particularly provided are transgenic crop plants, which have been engineered to carry genes identified as stated above. Examples of suitable plants include tobacco (*Nicotiana tabacum*) and other *Nicotiana* species, carrot, vegetable and oilseed *Brassicas*, melons, Capsicums, grape vines, lettuce, strawberry, sugar beet, wheat, barley, corn(maize), rice, soybean, peas, sorghum, sunflower, tomato, cotton, and potato. Especially preferred transgenic plants of the invention include cotton, rice, oilseed Brassica species such as canola, corn(maize) and soybean.

In addition to a plant, the present invention provides any clone of such a plant, seed, selfed or hybrid progeny and descendants, and any part of any of these, such as cuttings, seed. The invention provides any plant propagule that is any part which may be used in reproduction or propagation, sexual or asexual, including cuttings, seed and so on. Also encompassed by the invention is a plant which is a sexually or asexually propagated offspring, clone or descendant of such a plant, or any part or propagule of said plant, offspring, clone or descendant.

The present invention also encompasses the polypeptide expression product of a nucleic acid molecule according to the invention as disclosed herein or obtainable in accordance with the information and suggestions herein. Also provided are methods of making such an expression product by expression from a nucleotide sequence encoding therefore under suitable conditions in suitable host cells e.g. *E.coli*. Those skilled in the art are well able to construct vectors and design protocols and systems for expression and recovery of products of recombinant gene expression.

The heterologous or exogenous target protein is contemplated to be any protein of interest that may be produced by the method of the invention.

A polypeptide according to the present invention may be an allele, variant, fragment, derivative, mutant or homologue of the(a) polypeptides as mentioned herein. The allele, variant, fragment, derivative, mutant or homologue may have substantially the same function of the polypeptides alluded to above and as shown herein or may be a functional mutant thereof.

"Homology" in relation to an amino acid sequence or polypeptide sequence produced by the method of the invention may be used to

refer to identity or similarity, preferably identity. As noted already above, high level of amino acid identity may be limited to functionally significant domains or regions.

In certain embodiments, an allele, variant, derivative, mutant derivative, mutant or homologue of the specific sequence may show little overall homology, say about 20%, or about 25%, or about 30%, or about 35%, or about 40% or about 45%, with the specific sequence. However, in functionally significant domains or regions, the amino acid homology may be much higher. Putative functionally significant domains or regions can be identified using processes of bioinformatics, including comparison of the sequences of homologues.

Functionally significant domains or regions of different polypeptides may be combined for expression from encoding nucleic acid as a fusion protein. For example, particularly advantageous or desirable properties of different homologues may be combined in a hybrid protein, such that the resultant expression product, may include fragments of various parent proteins, if appropriate.

Similarity of amino acid sequences may be as defined and determined by the TBLASTN program, of Altschul *et al.* (1990) *J. Mol. Biol.* **215**: 403-10, which is in standard use in the art. In particular, TBLASTN 2.0 may be used with Matrix BLOSUM62 and GAP penalties: existence: 11, extension: 1. Another standard program that may be used is BestFit, which is part of the Wisconsin Package, Version 8, September 1994, (Genetics Computer Group, 575 Science Drive, Madison, Wisconsin, USA, Wisconsin 53711). BestFit makes an optimal alignment of the best segment of similarity between two sequences. Optimal alignments are found by inserting gaps to maximize the number of matches using the local homology algorithm of Smith and Waterman (*Adv. Appl. Math.* (1981) 2: 482-489). Other algorithms include GAP, which uses the Needleman and

Wunsch algorithm to align two complete sequences that maximizes the number of matches and minimizes the number of gaps. As with any algorithm, generally the default parameters are used, which for GAP are a gap creation penalty = 12 and gap extension penalty = 4. Alternatively, a gap creation penalty of 3 and gap extension penalty of 0.1 may be used. The algorithm FASTA (which uses the method of Pearson and Lipman (1988) *PNAS USA* **85**: 2444-2448) is a further alternative.

Use of either of the terms "homology" and "homologous" herein does not imply any necessary evolutionary relationship between compared sequences, in keeping for example with standard use of terms such as "homologous recombination" which merely requires that two nucleotide sequences are sufficiently similar to recombine under the appropriate conditions. Further discussion of polypeptides according to the present invention, which may be encoded by nucleic acid according to the present invention, is found below.

There now follow non-limiting examples and figures illustrating the invention.

FIGURES

Figure 1: the major components of chloroplast transformation system.

(1) Transformation vector contains (i) chloroplast translocation sequence (CTS); (ii) chloroplast transgene cassette comprising left flanking sequence (LFS) and right flanking sequence (RFS) to facilitate insertion of the cassette into the chloroplast genome using homologous recombination, promoter region from tobacco chloroplast rrn16 gene (Prrn), aadA gene as a selectable marker (aadA), transcription terminator from chloroplast genome (term); and (iii) primer binding domain (PBD). (2) Reverse Transcriptase-

RNase H gene translationally fused to the chloroplast transit peptide from small subunit of tobacco Rubisco gene (rbcS-cTP). (3) CTS-Binding peptide translationally fused to the chloroplast transit peptide from Arabidopsis HSP60 gene (HSP60-cTP).

Figure 2: set of constructs used for chloroplast transformation in tobacco (ALG298, ALG327 and ALG 344) and in Arabidopsis (ALG347 and ALG327)

The chloroplast transformation cassette contains left and right flanking sequences (**LFS** and **RFS**), Prrn promoter (**Prrn**), aadA gene for spectinomycin selection (**aadA**), and rrnB transcription terminator (**rrnB ter**). Primer binding domain (**PBD**) from yeast Ty1 retrotransposon designed for capturing tRNA-Met from chloroplasts was fused to chloroplast transgene cassette. The resulting cassette was inserted within domain IV of *LtrB* (**LtrB5'** and **LtrB3'**) intron from *Lactococcus lactis* (ALG298 and ALG347) or fused to chloroplast translocation sequence from *Avocado sunblotch viroid* (**ASB-CTS** in ALG344). The chloroplast transgene cassette was expressed from nuclear inserted cassette and resultant RNA was translocated into the chloroplast using LtrASi protein for vectors (ALG298 and ALG347), or using native plant proteins for vector ALG344. Reverse transcription of the RNA was performed by Reverse transcriptase-RNaseH fused to chloroplast transit peptide (**cTP-RTRH**) from HSP60 gene (ALG327). **Ubiq3 Pro**- Arabidopsis promoter from ubiquitin 3 gene; **35S Pro**- promoter from Cauliflower Mosaic Virus 35S gene, **TAF2 Pro**- Arabidopsis promoter from TAF 2 gene; **nos ter**- transcription terminator from Agrobacterium nos gene.

Figure 3: modifications of the chloroplast transformation cassette were made by designing primer binding domain and positioning of building blocks on the transgene cassette.

CTU- chloroplast transformation unit; **CTS-5'**-chloroplast translocation sequence located at the 5'-end of the transformation cassette; **CTS-3'**- chloroplast translocation sequence located at the 3'-end of the transformation cassette; **PDB-CHL**- primer binding domain designed for reverse transcription in the chloroplasts using tRNA-Met from chloroplasts; **PBD-CYT**- primer binding domain

designed for reverse transcription in the cytoplasm using cytoplasmic tRNA-Met.

The modifications detailed in Example section 1B hereinafter and corresponding figures include a first modification of the use of PBD for the binding of cytoplasmic tRNA-Met as primer [Fig 3(C)]. As a second modification CTS can be located at both the 5'- and 3'- ends of the transformation cassette, such as in the case with the LtrB intron. The transgene cassette is inserted inside of the LtrB intron (domain IV). The PDB-CHL is located downstream of the LtrB 3'-end of the cassette (CTS-3'), so that the LtrA protein is able to function as both a translocation protein and reverse transcriptase. The LtrA protein has three major functions: (1) as a maturase (it binds to LtrB RNA and stabilises the secondary structure of the RNA, and assists splicing); (2) as an endonuclease (it induces single-stranded DNA breaks on target site); and (3) as a reverse transcriptase (it performs reverse transcription of the intron RNA after insertion of the LtrB intron RNA into the donor site).

The LtrA protein is unable to perform the reverse transcription reaction efficiently if the PBD-CYT is located adjacent to and in front of a chloroplast translocation sequence at the 3'-end of the CTU (CTS-3') as in Fig 3(B), but can efficiently reverse transcribe RNA if the PBD is located downstream of a chloroplast translocation sequence (CTS-3') as shown in Fig 3A. Such a positioning or the combination of components of the transformation cassette as shown in Fig 3(A) allows both the translocation of the CTU into the chloroplast and reverse transcription of the CTU by the LtrA protein. Thus, by positioning of the CTS components and of the PBD-CHL as shown in Fig 3(A) the procedure of transformation is simplified since there is no requirement to co-deliver another gene to provide a reverse transcriptase function.

A similar simplification of the procedure is achieved if a PBD-CYT is used, since there is a significant amount of native endogenous reverse transcriptase in the cytoplasm, and reverse transcription is initiated by endogenous reverse transcriptase using cytoplasmic

tRNA-Met. This also eliminates the necessity for the co-delivery of another gene for reverse transcription in the chloroplasts.

The case in Fig 1A and B is attributed to LtrB intron, the case in Fig 1C attributed to ASB-CTS.

Figure 4: schematic presentation of constructs based on the LtrB-CTS for chloroplast transformation in tobacco.

Nos ter- nos transcription terminator, **LtrB3-** 3'-prime end of LtrB intron, **PBD-CHL-** primer binding domain for chloroplast tRNA-Met, **PDB- CYT-** primer binding domain for cytoplasmic tRNA-Met, **trnA flank-** left flank of the transgene cassette, **psbA ter-** chloroplast transcription terminator from tobacco, **mGFP-** mGFP4 gene, **aadA-** aadA gene, **Trrn-** rrn16 chloroplast promoter from tobacco, **trnI flank-** right flank of transgene cassette, **LtrB5- 5'-**prime end of LtrB intron, **35S Pro-** 35S promoter from cauliflower mosaic virus (CaMV), **TAF2 Pro-** promoter from Arabidopsis TAF2 gene, **cTP-** chloroplast transit peptide from rbcS gene of tobacco, **LtrA-** gene encoded by open reading frame of LtrB intron, **ags ter-** ags gene transcription terminator.

Figure 5: schematic presentation of constructs based on the LtrB-CTS for chloroplast transformation in rice.

Nos ter- nos transcription terminator, **LtrB3-** 3'-prime end of LtrB intron, **PBD-CHL-** primer binding domain for chloroplast tRNA-Met, **PDB-CYT-** primer binding domain for cytoplasmic tRNA-Met, **trnA flank-** left flank of the transgene cassette, **atpA ter-** chloroplast transcription terminator from wheat, **mGFP-** mGFP4 gene, **aadA-** aadA gene, **Wrrn-** rrn16 chloroplast promoter from wheat, **trnI flank-** right flank of transgene cassette, **LtrB5- 5'-**prime end of LtrB intron, **35S Pro-** 35S promoter from cauliflower mosaic virus (CaMV), **Act1 Pro-** actin 1 gene promoter from rice, **cTP-** chloroplast transit peptide from rbcS gene of tobacco, **LtrA-** gene encoded by open reading frame of LtrB intron, **ags ter-** ags gene transcription terminator.

Figure 6: Schematic presentation of constructs based on ASB-CTS for chloroplast transformation in tobacco.

Nos ter- nos transcription terminator, **ASB**- sequence from Avocado sunblotch viroid (ASBVd) as CTS, **PBD-CHL**- primer binding domain for chloroplast tRNA-Met, **PDB- CYT**- primer binding domain for cytoplasmic tRNA-Met, **trnA flank**- left flank of the transgene cassette, **psbA ter**- chloroplast transcription terminator from tobacco, **mGFP**- mGFP4 gene, **aadA**- aadA gene, **Trrn**- rrn16 chloroplast promoter from tobacco, **trnI flank**- right flank of transgene cassette, **35S Pro**- 35S promoter from cauliflower mosaic virus (CaMV), **TAF2 Pro**- promoter from Arabidopsis TAF2 gene, **cTP**- chloroplast transit peptide from rbcS gene of tobacco, **RT-ty1**- reverse transcriptase gene from yeast ty1 retrotransposon, **ags ter**- ags gene transcription terminator.

Figure 7: PCR amplification of left flanking junction in tobacco transformed by the LtrB-CTS-based vectors.

M- DNA marker, **1-6**- independent transgenic lines, **wt**- non-transgenic tobacco, **NC**- negative control without DNA.

Figure 8: Southern hybridisation for tobacco transformed with ASB-CTS and LtrB-CTS based vectors.

Expected size of wild type DNA band is ~1.3 kb, and band with transgene insertion ~3.6 kb. Chloroplast probe upstream of LFS was used as a probe. **M**- DNA marker, **wt**- DNA from non-transgenic line, **1-3**- ASB-CTS lines, **4-8**- LtrB-CTS transgenic lines.

Figure 9: Northern analysis for tobacco plants transformed with LtrB-CTS based vector.

The aad-GFP DNA probe was used for hybridisation. Expected size of the band is ~1.5 kb. **Lane 1**- RNA from plants transformed with 35S-aadA-GFP-nos cassette; **lane 2**- WT RNA; **lanes 3-8**- independent transgenic lines.

EXPERIMENTAL SECTION 1A

A novel approach for efficient chloroplast transformation

A new method for chloroplast transformation in plants comprises (1) a transformation vector consisting of 3 major domains: (i)

chloroplast translocation sequence (CTS), (ii) chloroplast transgene cassette, (iii) primer binding domain (PBD) which uses chloroplast tRNA-fMet or any other chloroplast tRNAs as a primer for reverse transcription;

(2) Reverse Transcriptase- RNase H (RT-RH) from retrotransposon or retroviruses fused to chloroplast transit peptide for targeting into chloroplasts;

(3) RNA binding protein that binds to chloroplast translocation sequence (CTS) of the transformation vector, fused to chloroplast transit peptide (Fig.1).

Technology Rationale

The process of chloroplast transformation comprises two steps:

(1) Targeting of RNA-protein complex to the chloroplasts.

After delivery of the chloroplast transformation construct into the plant cell a strong expression of the RNA which contains the chloroplast translocation sequence (CTS) transgene cassette and primer binding domain (PBD) is achieved from the nuclear specific promoter. The CTS binding protein (CTS-BP) fused to a chloroplast transit peptide, will be also over-expressed on co-delivery from the same or a different vector and then will bind to the CTS, and facilitate translocation of the RNA into the chloroplasts.

Once the chloroplast transformation vector is presented in the plant cell via nuclear transformation, the chloroplast will then be permanently bombarded by the expressed CTS-BP-RNA complex. Such stable and continuous pumping of the complex into the targeted organelle is a prerequisite for achieving a high efficiency of organelle transformation. The technology exploits the finding that the chloroplast transit sequence is sufficient to permit the whole CTS-BP-RNA complex to be then taken up by the chloroplast.

Chloroplast translocation sequence (CTS) can be selected from a number of RNA sequences such as viroid RNA, groupI and groupII intron RNA, viral coat protein binding domains, retrotransposon

primer binding sites, which are recognised by corresponding native RNA binding proteins.

(2) Reverse transcription of the transgene cassette and insertion into the chloroplast genome.

Once the RNA of the transformation vector is inside of the organelle, primer binding domain (PBD) of the vector RNA captures tRNA-fMet as a primer, and the over-expression of the reverse transcriptase (RT-RH) fused to the chloroplast transit peptide facilitates reverse transcription of RNA into single stranded DNA. This is followed by insertion of the reverse transcribed cassette into the chloroplast genome using homologous recombination between flanking sequences of the transgene cassette and the homologous regions in the chloroplast genome.

Primer binding domain (PBD) is designed to capture RT-RH protein and chloroplast tRNA-fMet (or other chloroplast tRNAs) as a primer, and initiate reverse transcription of chloroplast transgene cassette RNA into single-stranded DNA.

Once the population of organelle genomes has been transformed in the initial plant line, the nuclear encoded transgenes are no longer required and they can then be removed through segregation in subsequent plant generations, leaving a clean organelle transformed plant line.

Materials and Methods

Preparation of group II intron based chloroplast translocation sequence (CTS).

LtrB intron from *Lactococcus lactis* was synthesised by commercial DNA synthesis provider. Potential splicing sites were eliminated from this sequence as described in our previous patent. The domain for insertion of transgene cassette (AscI-MluI-NotI sites) is underlined and shown in bold letters.

LtrB intron sequence

GGATCCCTCGAGGTGCGCCCAGATAGGGTGTTAAGTCAAGTAGTTTAAGGTACTACTCAGTAAGAT
AACACTGAAAACAGCCAACCTAACCGAAAAGCGAAAGCTGATACGGGAACAGAGCACGGTTGGAAA
GCGATGAGTTAGCTAAAGACAATCGGCTACGACTGAGTCGCAATGTTAATCAGATATAAGCTATAA
GTTGTGTTTACTGAACGCAAGTTTCTAATTTCGGTTATGTGTCGATAGAGGAAAGTGTCTGAAACC
TCTAGTACAAAGAAAGCTAAGTTATGGTTGTGGACTTAGCTGTTATCACCACATTTGTACAATCTG
TTGGAGAACCAATGGGAACGAAACGAAAGCGATGGCGAGAATCTGAATTTACCAAGACTTAACACT
AACTGGGGATAGCCTAAACAAGAATGCCTAATAGAAAGGAGGAAAAAGGCTATAGCACTAGAGCTT
GAAAATCTTGCAAGGCTACGGAGTAGTCGTAGTAGTCTGAGAAGGCTAACGGCCTTTACATGGCAA
AGGGCTACAGTTATTGTGTACTAAAATTAAAAATTGATTAGGGAGGAAAACCTCAAAATGAAACCA
ACAATGGCAATTTTAGAAAGAATCAGTAAAAATTCACAAGAAAATATAGACGAAGTTTTTACAAGA
CTTTATCGTTATCTTTTACGTCCTGATATTTATTACGTGGCG**GGCGCGCCACGCGTGCGGCCGC**TG
GGAAATGGCAATGATAGCGAAAGAACCTAAAACTCTGGTTCTATGCTTTCATTGTCATCGTCACGT
GATTCATAAACACAAGTGAATTTTTACGAACGAACAATAACAGAGCCGTATACTCCGAGAGGGGTA
CGTACGGTTCCCGAAGAGGGTGGTGCAAACCAGTCACAGTAATGTGAACAAGGCGGTACCTCCCTA
CTTCACCATATCATTTTTAATTCTACGAATCTTTATACTGGCAAACAATTTGACTG

SEQ ID NO.1

The chloroplast translocation sequence (CTL) from Avocado sunblotch viroid (Bank Accession No. J02020) was synthesised by PCR using the set of the following overlapping primers:

AS839 GAACTAATTTTTTTAATAAAAGTTCACCACGACTCCTCCTTCTCTCACAA

SEQ ID NO.2

AS840 TAAAAAAATTAGTTCACTCGTCTTCAATCTCTTGATCACTTCGTCTCTTC

SEQ ID NO.3

AS841 TGCGAGACTCATCAGTGTTCTTCCCATCTTTCCCTGAAGAGACGAAGTGA

SEQ ID NO.4

AS842 CTGATGAGTCTCGCAAGGTTTACTCCTCTATCTTCATTGTTTTTTTACAA

SEQ ID NO.5

AS843 GGGCGCGCCAAGATTTTGTAAAAAAACAATGAAGA SEQ ID NO.6

AS844 GCTCGAGACTTGTGAGAGAAGGAGGAGTC SEQ ID NO.7

The CTL sequence from Avocado sunblotch viroid

GCTCGAGACTTGTGAGAGAAGGAGGAGTCGTGGTGAACTTTTATTAAAAAAATTAGTTCACTCGTC
TTCAATCTCTTGATCACTTCGTCTCTTCAGGGAAAGATGGGAAGAACACTGATGAGTCTCGCAAGG
TTTACTCCTCTATCTTCATTGTTTTTTTACAAAATCTTGGGCGCGCCC SEQ ID NO.8

The expression of chloroplast translocation sequence and chloroplast cassette fused to it was driven by 35S promoter from Cauliflower mosaic virus obtained by DNA synthesis

35S promoter sequence

CAATCCCACAAAAATCTGAGCTTAACAGCACAGTTGCTCCTCTCAGAGCAGAATCGGGTATTCAAC
ACCCTCATATCAACTACTACGTTGTGTATAACGGTCCACATGCCGGTATATACGATGACTGGGGTT
GTACAAAGGCGGCAACAAACGGCGTTCCCGGAGTTGCACACAAGAAATTTGCCACTATTACAGAGG
CAAGAGCAGCAGCTGACGCGTACACAACAAGTCAGCAAACAGACAGGTTGAACTTCATCCCCAAAG
GAGAAGCTCAACTCAAGCCCAAGAGCTTTGCTAAGGCCCTAACAAGCCCACCAAAGCAAAAAGCCC
ACTGGCTCACGCTAGGAACCAAAAGGCCCAGCAGTGATCCAGCCCCAAAAGAGATCTCCTTTGCCC
CGGAGATTACAATGGACGATTTCCTCTATCTTTACGATCTAGGAAGGAAGTTCGAAGGTGAAGTAG
ACGACACTATGTTCACCACTGATAATGAGAAGGTTAGCCTCTTCAATTTCAGAAAGAATGCTGACC
CACAGATGGTTAGAGAGGCCTACGCAGCAGGTCTCATCAAGACGATCTACCCGAGTAACAATCTCC
AGGAGATCAAATACCTTCCCAAGAAGGTTAAAGATGCAGTCAAAAGATTCAGGACTAATTGCATCA
AGAACACAGAGAAAGACATATTTCTCAAGATCAGAAGTACTATTCCAGTATGGACGATTCAAGGCT
TGCTTCATAAACCAAGGCAAGTAATAGAGATTGGAGTCTCTAAAAAGGTAGTTCCTACTGAATCTA
AGGCCATGCATGGAGTCTAAGATTCAAATCGAGGATCTAACAGAACTCGCCGTGAAGACTGGCGAA
CAGTTCATACAGAGTCTTTTACGACTCAATGACAAGAAGAAAATCTTCGTCAACATGGTGGAGCAC
GACACTCTGGTCTACTCCAAAAATGTCAAAGATACAGTCTCAGAAGACCAAAGGGCTATTGAGACT
TTTCAACAAAGGATAATTTCGGGAAACCTCCTCGGATTCCATTGCCCAGCTATCTGTCACTTCATC
GAAAGGACAGTAGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCTATC
ATTCAAGATCTCTCTGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAA
AAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGATGTGACATCTCCACTGACGTAAGG
GATGACGCACAATCCCACTATCCTTCGCAAGACCCTTCCTCTATATAAGGAAGTTCATTTCATTTG
GAGAGGACACG SEQ ID NO.9

The chloroplast transgene cassette contains left and right flanking sequences (LFS and RFS) for insertion of whole cassette

into the chloroplast genome using homologous recombination, Prrn16 promoter region from tobacco, aadA gene as a selectable marker, and 3'UTR sequence of psbA gene as transcription terminator (Fig. 1).

LFS sequences for tobacco and Arabidopsis were amplified using the following PCR primers:

AS699 GGCGCGCCGTGGGATCCGGGCGGTCCG SEQ ID NO.10

AS700 GGCATGCTGGCGCAGCTGGGCCATCC SEQ ID NO.11

Tobacco LFS sequence

GGCGCGCCATGGGATCCGGGCGGTCCGGGGGGGACCACCACGGCTCCTCTCTTCTCGAGAATCCAT
ACATCCCTTATCAGTGTATGGACAGCTATCTCTCGAGCACAGGTTTAGCAATGGGAAAATAAAATG
GAGCACCTAACAACGCATCTTCACAGACCAAGAACTACGAGATCGCCCCTTTCATTCTGGGGTGAC
GGAGGGATCGTACCATTCGAGCCGTTTTTTTCTTGACTCGAAATGGGAGCAGGTTTGAAAAAGGAT
CTTAGAGTGTCTAGGGTTGGGCCAGGAGGGTCTCTTAACGCCTTCTTTTTTCTTCTCATCGGAGTT
ATTTCACAAAGACTTGCCAGGGTAAGGAAGAAGGGGGGAACAAGCACACTTGGAGAGCGCAGTACA
ACGGAGAGTTGTATGCTGCGTTCGGGAAGGATGAATCGCTCCCGAAAAGGAATCTATTGATTCTCT
CCCAATTGGTTGGACCGTAGGTGCGATGATTTACTTCACGGGCGAGGTCTCTGGTTCAAGTCCAGG
ATGGCCGCATGCC SEQ ID NO.12

Arabidopsis LFS sequence

GGCGCGCCGTGGGATCCGGGCGGTCCGGAGGGGACCACTATGGCTCCTCTCTTCTCGAGAATCCAT
ACATCCCTTATCAGTGTATGGACAGCTATCTCTCGAGCGCAGGTTTAGGTTCGGCCTCAATGGGAA
AATAAAATGGAGCACCTAACAACGTATCTTCACAGACCAAGAACTACGAGATCACCCCTTTCATTC
TGGGGTGACGGAGGGATCGTACCGTTCGAGCCTTTTTTTCATGTTATCTATCTCTTGACTCGAAAT
GGGAGCAGGTTTGAAAAAGGATCTTAGAGTGTCTAGGGTTAGGCCAGTAGGGTCTCTTAACGCCCT
CTTTTTTCTTCTCATCGAAGTTATTTCACAAATACTTCCTATGGTAACGAAGAGGGGGGGAACAAG
CACACTTGGAGAGCGCAGTACAACGGAGAGTTGTATGCTGCGTTCGGGAAGGATGAATCGCTCCCG
AAAAGGAATCTATTGATTCTCTCCCAATTGGTTGGACCATAGGTGCGATGATTTACTTCACGGGCG
AGGTCTCTGGTTCAAATCCAGGATGGCCCAGCTGCGCCAGCATGC SEQ ID NO.13

RFS sequences were amplified using the following PCR primers:

AS764 TGATATCGGATGGCCCTGCTGCGCCAGGGAAAAGAAT SEQ ID NO.14

AS845 GCCGCGGATTGCCCTTCTCCGACCCTGAC SEQ ID NO.15

Tobacco RFS sequence

GATATCGGATGGCCCTGCTGCGCCAGGGAAAAGAATAGAAGAAGCATCTGACTACTTCATGCATGC
TCCACTTGGCTCGGGGGGATATAGCTCAGTTGGTAGAGCTCCGCTCTTGCAATTGGGTCGTTGCGA
TTACGGGTTGGATGTCTAATTGTCCAGGCGGTAATGATAGTATCTTGTACCTGAACCGGTGGCTCA
CTTTTTCTAAGTAATGGGGAAGAGGACCGAAACGTGCCACTGAAAGACTCTACTGAGACAAAGATG
GGCTGTCAAGAACGTAGAGGAGGTAGGATGGGCAGTTGGTCAGATCTAGTATGGATCGTACATGGA
CGGTAGTTGGAGTCGGCGGCTCTCCCAGGGTTCCCTCATCTGAGATCTCTGGGGAAGAGGATCAAG
TTGGCCCTTGCGAACAGCTTGATGCACTATCTCCCTTCAACCCTTTGAGCGAAATGCGGCAAAAGA
AAAGGAAGGAAAATCCATGGACCGACCCCATCATCTCCACCCCGTAGGAACTACGAGATCACCCCA
AGGACGCCTTCGGCATCCAGGGGTCACGGACCGACCATAGAACCCTGTTCAATAAGTGGAACGCAT
TAGCTGTCCGCTCTCAGGTTGGGCAGTCAGGGTCGGAGAAGGGCAATCCGCGG

SEQ ID NO.16

Arabidopsis RFS sequence

GATATCGGATGGCCCTGCTGCGCCAAGGAAAAGAATATAAGAAGGATCTGACTCCTTCATGCATGC
TCCACTTGGCTCGGGGGATATAGCTCAGTTGGTAGAGCTCCGCTCTTGCAATTGGGTCGTTGCGAT
TACGGGTTGGGTGTCTAATTGTCCAGGCGGTAATGATAGTATCTTGTACCTGAACCGGTGGCTCAC
TTTTTCTAAGTAATGGGGAAAAGGACCGAAACATGCCACTGAAAGACTCTACTGAGACAAAGATGG
GCTGTCAAGAACGTAGAGGAGGTAGGATGGTCAGTTGGTCAGATCTAGTATGGATCGTACATGGAC
GGTAGTTGGAGTCGGCGGCTCTCCTAGGGTTCCCTCGTCTGGGATTGATCCCTGGGGAAGAGGATC
AAGTTGGCCCTTGCGAACAGCTTGATGCACTATCTCCCTTCAACCCTTTGAGCGAAATGCGGCAAA
AGGAAGGAAAATCCATGGACCGACCCCATCGTCTCCACCCCGTAGGAACTACGAGATCACCCCAAG
GACGCCTTCGGTATCCAGGGGTCGCGGACCGACCATAGAACCCTGTTCAATAAGTGGAATGCATTA
GCTGTCCGCTCGCAGGTTGGGCAGTAAGGGTCGGAGAAGGGCAATCCGCGG SEQ ID NO.17

Prrn promoter was amplified from tobacco genomic DNA cv. Petite Gerard using following PCR primers:

AS750 GGCATGCCGCAATGTGAGTTTTTGTAGTTG SEQ ID NO.18

Prrn-R ACTTGTATCGATGCGCTTCATATTCGCCCGGA SEQ ID NO.19

Prrn16 promoter sequence

GCATGCCGCAATGTGAGTTTTTGTAGTTGGATTTGCTCCCCCGCCGTCGTTCAATGAGAATGGATA
AGAGGCTCGTGGGATTGACGTGAGGGGGCAGGGATGGCTATATTTCTGGGAGCGAACTCCGGGCGA
ATATGAAGCGCATCGATACAAGT SEQ ID NO.20

aadA gene was synthesised by commercial DNA synthesis provider. Three introns from Arabidopsis gene At2g29890 were inserted into the coding sequence to optimise expression of the aadA in the cytoplasm of plant cells. The introns are underlined and shown in bold letters.

aadA gene sequence

ATGGCAGAAGCGGTGATCGCCGAAGTATCGACTCAACTATCAGAG**GTAAGTAACTTTTAGCTCTCA**
GCTGCTGTTTACTAAGTTCATGCCATACATTGATTCTGGTTTATTAAGGGTTATGTTCAGTATTAC
TAGTAACAAAATCTATTTCTTCGTTTCCGTCTGCAGGTAGTTGGCGTCATCGAGCGCCATCTCGAA
CCGACGTTGCTGGCCGTACATTTGTACGGCTCCGCAGTGGATGGCGGCCTGAAGCCACACAGTGAT
ATTGATTTGCTGGTTACGGTGACCGTAAGGCTTGATGAAACAACGCGGCGAGCTTTGATCAACGAC
CTTTTGGAAACTTCGGCTTCCCCTGGAGAGAGCGAGATTCTCCGCGCTGTAGAG**GTAATTTTCATC**
TTTGTTTGGCCTTCCAAGTGCTTTTTTTGCTGTTTACGGGTGGAACTTCAGTAAAAATGGGATCAA
AACATCATATGGCATAAATAAATTTTAAGAATGGCGAACTCGGGGTTACCGAATATGGCTTCCTTT
TTCAGTGTTTCTTAGTCCATTGTACTTATGAGATTGCAGGTCACCATTGTTGTGCACGACGACATC
ATTCCGTGGCGTTATCCAGCTAAGCGCGAACTGCAATTTGGAGAATGGCAGCGCAATGACATTCTT
GCAGGTATCTTCGAGCCAGCCACGATCGACATTGATCTGGCTATCTTGCTGACAAAAGCAAGAGAA
CATAGCGTTGCCTTGGTAGGTCCAGCGGCGGAGGAACTCTTTGATCCGGTTCCTGAACAGGATCTA
TTTGAGGCGCTAAATGAAACCTTAACGCTATGGAACTCGCCGCCCGACTGGGCAG**GTAAGAAATCT**
TTTCCCATCTTGAAGTCACCTCAAACCGAACGTTAGGAAATTCCAAAATGTTTTGATAGTAGTCTA
CTTAGTTTCAAGTTTTGGGTTTGTGTATACTTTCACTAATAATATGCGTGGAAACATTGCAGGTGA
TGAGCGAAATGTAGTGCTTACGTTGTCCCGCATTTGGTACAGCGCAGTAACCGGCAAAATCGCGCC
GAAGGATGTCGCTGCCGACTGGGCAATGGAGCGCCTGCCGGCCCAGTATCAGCCCGTCATACTTGA
AGCTAGACAGGCTTATCTTGGACAAGAAGAAGATCGCTTGGCCTCGCGCGCAGATCAGTTGGAAGA
ATTTGTCCACTACGTGAAAGGCGAGATCACCAAGGTAGTCGGCAAATAA

SEQ ID NO.21

Thé psbA 3'UTR terminator was amplified from the tobacco genomic DNA cv Petite Gerard using the following primers:

AS749 GGATATCAAACAAATACAAAATCAAAATAGA SEQ ID NO.22

AS778 GGAATTCTGAGCGCGCTAGAGCGATCCTG SEQ ID NO.23

psbA 3'UTR sequence

GAATTCTGAGCGCGCTAGAGCGATCCTGGCCTAGTCTATAGGAGGTTTTGAAAAGAAAGGAGCAAT
AATCATTTTCTTGTTCTATCAAGAGGGTGCTATTGCTCCTTTCTTTTTTTCTTTTTATTTATTTAC
TAGTATTTTACTTACATAGACTTTTTTGTTTACATTATAGAAAAAGAAGGAGAGGTTATTTTCTTG
CATTTATTCATGATTGAGTATTCTATTTTGATTTTGTATTTGTTTGATAT SEQ ID NO.24

Primer Binding Domain (PBD) was designed as described by Friant et al., ((1998) Mol. Cellul. Biology, 18: 799-806) and amplified by PCR using the set of following overlapping primers:

AS830 CCGCGGTATCTCACATTCACCCAATTGTCATGGTT SEQ ID NO.25

AS831 TTAGAAGTATCCTGTGCACATCCGCAACCATGACAATTGG SEQ ID NO.26

AS832 ACAGGATACTTCTAAGGAAGTCCACACAAATCAAGAACCCTTAGA SEQ ID NO.27

AS833 TCACATTCTTCTGTTTTGGTAGCTGAAACGTCTAAGGGTTCTTGA SEQ ID NO.28

AS834 ACAGAAGAATGTGAGAAGGCTTCCACTAAGGCTAACTCTCAACAG SEQ ID NO.29

AS835 CGCGGCCGCGTTGTCTGTTGAGAGTTAGC SEQ ID NO.30

PBD sequence

CCGCGGTATCTCACATTCACCCAATTGTCATGGTTGCGGATGTGCACAGGATACTTCTAAGGAAGT
CCACACAAATCAAGAACCCTTAGACGTTTCAGCTACCAAAACAGAAGAATGTGAGAAGGCTTCCAC
TAAGGCTAACTCTCAACAGACAACGCGGCCGC SEQ ID NO.31

LtrA gene from *Lactococcus lactis* encoded by the LtrB intron was synthesised by commercial DNA synthesis provider. The sequence of the LtrA protein was first optimised for codon usage in plants and 5 plant introns were inserted into the coding sequence to improve LtrA expression in plants. Plant introns inserted in the coding sequence of LtrA gene are underlined and shown in bold letters. The introns 1,2 4 are from Arabidopsis gene At5g01290, intron 3 and 5 were selected from Arabidopsis gene At5g43940. The clone was named as LtrASi.

LtrASi gene sequence:

GCATGCATGAAGCCAACAATGGCAATCCTCGAACGAATCTCTAAGAACTCACAGGAGAACATCGAC
GAG**GTACAATAACCCATATATATGAATTGATTCATGTGTTACTCGTACTTGTTTGAATATGTTTGG**
AGCAAGTTTGATACTTTTGGATGATGATATCGCAAATTCGTTATCTTTTTGGCGTTATAGGTCTTC
ACAAGACTTTACCGTTACCTTCTCCGTCCTGACATCTACTACGTGGCATATCAGAACCTCTACTCT
AACAAGGGAGCTTCTACAAAGGGAATCCTCGATGATACAGCTGATGGATTCTCTGAGGAGAAGATC
AAGAAGATCATCCAATCTTTGAAGGACGGAACTTACTACCCTCAGCCTGTCCGAAGAATGTACATC
GCAAAGAAGAACTCTAAGAAGATGAGACCTCTTGGAATCCCAACTTTCACAGACAAGTTGATCCAG
GAGGCTGTGAGAATCATCCTTGAATCTATCTATGAGCCTGTCTTCGAGGATGTGTCTCACGGTTTC
CGACCTCAGCGAAGCTGTCACACAGCTTTGAAGACAATCAAGAGAGAGTTCGGAG**GTAAATTATAT**
GCTTTGCCACTTCCTCAAAAGATCATTTTAGGTTCATTGGTATGTGGTTTTTTTCTTAACAGGTGC
AAGATGGTTCGTGGAGGGAGATATCAAGGGATGCTTCGATAACATCGACCACGTCACACTCATCGG
ACTCATCAACCTTAAGATCAAGGATATGAAGATGAGCCAGTTGATCTACAAGTTCCTCAAGGCAGG
TTACCTCGAAAACTGGCAGTACCACAAGACTTACAGCGGAACACCTCAGGGCGGAATCCTCTCTCC
TCTCCTCGCTAACATCTATCTTCATGAATTGGACAAGTTCGTTCTCCAACTCAAGATGAAGTTCGA
CCGAGAGAGTCCAGAGAGAATCACACCTGAATACCGGGAGCTTCACAACGAGATCAAAAGAATCTC
TCACCGTCTCAAGAAGTTGGAGGGCGAGGAGAAGGCTAAGGTTCTCTTGGAATACCAGGAGAAGAG
GAAGAGGTTGCCTACACTCCCTTGTACATCACAAACAAACAAG**GTTCGTTCTCTCCATTTTCATTC**
GTTTGAGTCTGATTTAGTGTTTTGTGGTTGATCTGAATCGATTTATTGTTGATTAGTGAATCAATT
TGAGGCTGTGTCCTAATGTTTTGACTTTTGATTACAGGTCTTGAAGTACGTCCGATACGCTGACGA
CTTCATCATCTCTGTTAAGGGAAGCAAGGAGGACTGTCAATGGATCAAGGAGCAATTGAAGCTCTT
CATCCATAACAAGCTCAAGATGGAATTGAGTGAGGAGAAGACACTCATCACACATAGCAGTCAGCC
TGCTCGTTTCCTCGGATACGACATCCGAGTCAGGAGAAGTGGAACTATCAAGCGATCTGGAAAG**GT**
TCAATTCTTTCTTTCACATTTGTACTTGTTCACTCGTTTTATTAATCCTCTTTAGAATGGAGATTC
TTACCTCTGTGTGGCCTTTGGCAGGTCAAGAAGAGAACACTCAACGGGAGTGTGGAGCTTCTCATC
CCTCTCCAAGACAAGATCCGTCAATTCATCTTCGACAAGAAGATCGCTATCCAGAAGAAGGATAGC
TCATGGTTCCCAGTTCACAGGAAGTACCTTATCCGTTCAACAGACTTGGAGATCATCACAATCTAC
AACTCTGAATTGAGAG**GTAAGCTGCTACCTCAAACTTTCTAGTGCTTCCATATTTCCTTTCTTCTG**
CAAGGCAGAGAACCATTGTGGTTAAGTGTTTTAAATTGTGAATGTATAGGTATCTGCAACTACTAC
GGTCTCGCAAGTAACTTCAACCAGCTCAACTACTTCGCTTACCTTATGGAATACTCTTGCTTGAAG
ACTATCGCATCTAAGCATAAGGGAACACTCTCAAAGACCATCTCTATGTTCAAGGATGGAAGTGGT
TCTTGGGGAATCCCTTACGAGATCAAGCAGGGGAAGCAGAGGAGATACTTCGCCAACTTCAGTGAA
TGCAAATCTCCTTACCAATTCACTGATGAGATCAGTCAAGCTCCTGTGCTTTACGGATACGCTCGG
AACACTCTTGAGAACAGACTTAAGGCTAAGTGTTGTGAGCTTTGTGGAACATCTGATGAGAACACA
TCTTACGAGATCCACCACGTCAACAAGGTCAAGAACCTTAAGGGAAAGGAGAAGTGGGAGATGGCA
ATGATCGCTAAGCAGCGGAAGACTCTTGTTGTTTGCTTCCATTGTCATCGTCACGTGATCCATAAG
CACAAGTGAACTAGTAA SEQ ID NO.32

The LtrA gene was translationally fused to the chloroplast transit peptide (rbcS-cTP) from tobacco Rubisco small subunit gene (Bank Access. No. AY220079) which was amplified using the following PCR primers:

AS794 GCTCGAGACAATGGCTTCCTCAGTTCTTTCCTCT SEQ ID NO.33

AS639 CGCATGCTACCTGCATACATTGCACTCTTCCACCAT SEQ ID NO.34

rbcS-cTP sequence

CTCGAGACAATGGCTTCCTCAGTTCTTTCCTCTGCAGCAGTTGCCACTCGCACCAATGTTGCTCAA
GCTAACATGGTTGCACCTTTCACTGGTCTTAAGTCAGCTGCCTCATTCCCTGTTTCAAGGAAGCAA
AACCTTGACATCACTTCCATTGCTAGCAATGGTGGAAGAGTGCAATGTATGCAGGTAGCATGC

SEQ ID NO.34

The 5′ promoter region from Arabidopsis ubiquitin 3 gene was amplifies with the following primers:

AS724 CGGTACCTACCGGATTTGGAGCCAAGTC SEQ ID NO.35

AS726 GTGTTTGGTGACCTGAAATAAAACAATAGAACAAGT SEQ ID NO.36

Arabidopsis ubiq3 promoter sequence

TACCGGATTTGGAGCCAAGTCTCATAAACGCCATTGTGGAAGAAAGTCTTGAGTTGGTGGTAATGT
AACAGAGTAGTAAGAACAGAGAAGAGAGAGAGTGTGAGATACATGAATTGTCGGGCAACAAAAATC
CTGAACATCTTATTTTAGCAAAGAGAAAGAGTTCCGAGTCTGTAGCAGAAGAGTGAGGAGAAATTT
AAGCTCTTGGACTTGTGAATTGTTCCGCCTCTTGAATACTTCTTCAATCCTCATATATTCTTCTTC
TATGTTACCTGAAAACCGGCATTTAATCTCGCGGGTTTATTCCGGTTCAACATTTTTTTTGTTTTG
AGTTATTATCTGGGCTTAATAACGCAGGCCTGAAATAAATTCAAGGCCCAACTGTTTTTTTTTTTA
AGAAGTTGCTGTTAAAAAAAAAAAAAGGGAATTAACAACAACAACAAAAAAAGATAAAGAAAATAA
TAACAATTACTTTAATTGTAGACTAAAAAAACATAGATTTTATCATGAAAAAAAGAGAAAAGAAAT
AAAAACTTGGATCAAAAAAAAAACATACAGATCTTCTAATTATTAACTTTTCTTAAAAATTAGGTC
CTTTTTCCCAACAATTAGGTTTAGAGTTTTGGAATTAAACCAAAAAGATTGTTCTAAAAAATACTC
AAATTTGGTAGATAAGTTTCCTTATTTTAATTAGTCAATGGTAGATACTTTTTTTTCTTTTCTTTA
TTAGAGTAGATTAGAATCTTTTATGCCAAGTATTGATAAATTAAATCAAGAAGATAAACTATCATA
ATCAACATGAAATTAAAAGAAAAATCTCATATATAGTATTAGTATTCTCTATATATATTATGATTG

CTTATTCTTAATGGGTTGGGTTAACCAAGACATAGTCTTAATGGAAAGAATCTTTTTTGAACTTTT
TCCTTATTGATTAAATTCTTCTATAGAAAAGAAAGAAATTATTTGAGGAAAAGTATATACAAAAAG
AAAAATAGAAAAATGTCAGTGAAGCAGATGTAATGGATGACCTAATCCAACCACCACCATAGGATG
TTTCTACTTGAGTCGGTCTTTTAAAAACGCACGGTGGAAAATATGACACGTATCATATGATTCCTT
CCTTTAGTTTCGTGATAATAATCCTCAACTGATATCTTCCTTTTTTTGTTTTGGCTAAAGATATTT
TATTCTCATTAATAGAAAAGACGGTTTTGGGCTTTTGGTTTGCGATATAAAGAAGACCTTCGTGTG
GAAGATAATAATTCATCCTTTCGTCTTTTTCTGACTCTTCAATCTCTCCCAAAGCCTAAAGCGATC
TCTGCAAATCTCTCGCGACTCTCTCTTTCAAGGTATATTTTCTGATTCTTTTTGTTTTTGATTCGT
ATCTGATCTCCAATTTTTGTTATGTGGATTATTGAATCTTTTGTATAAATTGCTTTTGACAATATT
GTTCGTTTCGTCAATCCAGCTTCTAAATTTTGTCCTGATTACTAAGATATCGATTCGTAGTGTTTA
CATCTGTGTAATTTCTTGCTTGATTGTGAAATTAGGATTTTCAAGGACGATCTATTCAATTTTTGT
GTTTTCTTTGTTCGATTCTCTCTGTTTTAGGTTTCTTATGTTTAGATCCGTTTCTCTTTGGTGTTG
TTTTGATTTCTCTTACGGCTTTTGATTTGGTATATGTTCGCTGATTGGTTTCTACTTGTTCTATTG
TTTTATTTCAGGTCACCAAACA SEQ ID NO.37

The *nos* teminator fragment was synthesised based on gene bank sequence accession EU048864.

nos terminator sequence

TCTAGAGTCAAGCAGATCGTTCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCG
GTCTTGCGATGATTATCATATAATTTCTGTTGAATTACGTGAAGCATGTAATAATTAACATGTAAT
GCATGACGTTATTTATGAGATGGGTTTTTATGATTAGAGTCCCGCAATTATACATTTAATACGCGA
TAGAAAACAAAATATAGCGCGCAAACTAGGATAAATTATCGCGCGCGGTGTCATCTATGTTACTAG
ATCGACCTGCAG SEQ ID NO.38

The reverse transcriptase-RNase H gene from yeast Ty1-H3 clone (Boeke et al., Mol. Cellul. Biology (1988), 8: 1432-1442; bank accession No. M18706) was optimised for codon usage in plants, and by insertion of 5 introns from Arabidopsis genome (intron 1- from At1g04820, intron 2- from At2g29550, intron 3- from At1g31810, intron 4 and 5- from At1g09170). The introns are underlined and shown in bold letters. The clone was synthesised by commercial DNA synthesis provider and named as RTRHi-Ty1.

RTRHi-Ty1 sequence

ATGAACAATTCATCCCACAACATCGTTCCTATCAAGACTCCAACTACTGTTTCTGAGCAGAACACT
GAAGAATCTATCATCGCTGATCTTCCACTTCCTGATCTTCCTCCAGAATCTCCTACTGAATTTCCT
GATCCATTCAAAGAACTTCCACCTATCAACTCAAGACAAACTAACTCTTCATTGGGCGGAATTGGC
GATTCTAATGCTTACACTACTATCAACTCTAAGAAGAG**GTATTGTAGCCAGCCTCAACCAGTCTTT**
TTGCTGTTACATTTTCTTGGGCTCATCTAATGTTATTTTCCTATTTTGTTTTCAGGTCACTTGAAG
ATAATGAAACTGAAATCAAAGTTTCTAGGGATACATGGAATACTAAGAATATGAGATCACTTGAAC
CTCCAAGATCTAAGAAGAGAATCCATCTTATTGCAGCTGTTAAAGCTGTGAAATCAATCAAACCAA
TTAGAACAACTCTTAGATACGATGAAGCAATTACATACAACAAAGACATCAAGGAGAAGGAGAAAT
ACATCGAGGCTTACCACAAAGAAGTTAACCAACTTCTTAAGATGAAAACTTGGGATACTGATGAAT
ACTACGATAGAAAAGAGATTGACCCTAAGAGAGTTATCAACTCAATGTTCATCTTCAACAAGAAGA
GAGACGGAACTCACAAAGCTAGATTCGTTGCAAGAGGAGATATTCAGCATCCTGACACTTACGATT
CAG**GTAAGTATTCCAATGTTCTTCGATTATGAGTCAATGTTGTTACTGTATCTGTCTCTGTGGTTT**
ATTGTTTCAGGCTTAGTTATTGATTAGTATTGAAACTTCACTCACATATTTTTTTGTTTGTTTTCT
GAATTGTGCAGGTATGCAATCTAATACTGTTCATCACTACGCATTGATGACATCTCTTTCACTTGC
ATTGGACAATAACTACTACATTACACAACTTGACATATCTTCTGCATACCTTTACGCTGATATCAA
GGAGGAGCTTTACATTAGACCTCCACCACATTTGGGAATGAATGATAAGTTGATCCGTTTGAAGAA
ATCACTTTACGGATTGAAACAATCTGGAGCTAATTGGTACGAAACTATCAAATCATACCTTATTCA
GCAATGCGGTATGGAGGAAGTTAGGGGATGGTCATGCGTATTCAAGAACTCTCAAGTTACAATCTG
CCTCTTCGTTGATGATATGGTGCTCTTCTCTAAGAATCTTAACTCAAACAAGAGAATCATTGAGAA
GTTGAAGATGCAATACGACACTAAGATCATCAACCTTGGAGAATCTGATGAGGAAATTCAATACGA
CATTCTTGGATTGGAAATCAAATACCAAAGAG**GTGAGTTATATTTAACAGCTCATCAGTTACTTAA**
ACACTTTTTGGGACAAGCAGTTCAAACTCATGTTCCAATCCTAAAATTAATTGCAATTCACAGGTA
AGTACATGAAGTTGGGAATGGAAAACTCATTGACTGAGAAGATTCCTAAACTTAACGTTCCTTTGA
ATCCAAAGGGAAGAAAGCTCTCTGCTCCAGGACAACCAGGACTTTACATTGACCAGGATGAACTTG
AGATTGATGAGGATGAATACAAGGAGAAAGTACACGAGATGCAGAAGTTGATTGGACTTGCTTCAT
ACGTTGGATACAAATTCAGATTCGACCTTCTTTACTACATCAACACACTTGCTCAGCATATACTTT
TCCCATCTAGGCAAGTTCTTGACATGACATACGAGCTTATCCAATTCATGTGGGACACTAGAGACA
AGCAACTCATATGGCACAAGAACAAGCCTACAGAGCCAGATAACAAGCTCGTTGCAATCTCTGATG
CTTCTTACGGAAACCAACCATACTACAAATCACAAATTGGAAACATCTACTTGCTTAACGGAAAG**G**
TACTTTTCTCAAAGACTTTACCTTATTGTGGAATATTGAATTTTCTGAAAGACTTCACCTTATCTA
CATTTGTAATTTTACTATGGTAATCAGGTGATTGGAGGAAAGAGCACTAAGGCTTCACTTACATGC
ACTTCAACTACTGAGGCAGAGATCCACGCTATATCAGAATCTGTACCACTTCTTAACAACCTTTCT
TACCTTATCCAAGAGCTTAACAAGAAGCCAATCATCAAGGGACTTCTTACTGACTCAAGATCAACA
ATCTCTATCATTAAGTCTACAAATGAAGAGAAATTCAGAAACAGATTCTTCGGAACAAAGGCAATG
AGACTTAGAGATGAAGTTTCAG**GTAAGTATTAACTTACCAAATGATCAATATTATTTTGAAATGCA**
GGTTTTAGAATAATACTCTCTGCCGTTCTTGTTTATTTCCAGGTAACAACCTTTACGTTTACTACA

TCGAGACTAAGAAGAACATTGCTGACGTTATGACAAAGCCTCTTCCTATCAAGACCTTCAAGTTGC
TTACTAACAAATGGATTCATTAA SEQ ID NO.39

The RT-RH-Ty1 sequence was translationally fused to the chloroplast transit peptide from pea chloroplast HSP60 heat shock protein (Accession No. L03299). The sequence for the transit peptide (HSP60-cTP) was amplified from pea genomic DNA using the following PCR primers:

AS293 TCTCGAGTTGATGGCTTCTTCTGCTCAAATA SEQ ID NO.40

AS294 GGCATGCAACTCTCAAAGTGAAACCCTTC SEQ ID NO.41

HSP60-cTP sequence

CTCGAGATGGCTTCTTCTGCTCAAATACACGGTCTCGGAACCGCTTCTTTCTCTTCCCTCAAAAAA
CCCTCTTCCATTTCCGGTAATTCCAAAACCCTTTTCTTCGGTCAGCGACTCAATTCCAACCACTCT
CCCTTCACCCGCGCCGCATTCCCTAAGTTAAGTAGCAAAACCTTTAAGAAGGGTTTCACTTTGAGA
GTTGCATGC SEQ ID NO.42

The expression of the RTRHi-Ty1 and HSP60-cTP fusion was driven by TAF2 promoter from Arabidopsis *taf2* gene. It was amplified from Arabidopsis genomic DNA (Col-0) using the following set of primers:

AG3 GGTACCATGATCGCTTCATGTTTTTATC SEQ ID NO.43

AG4 CTCGAGGTTCCTTTTTTGCCGATATGTTAG SEQ ID NO.44

TAF2 promoter sequence

GTACCATGATCGCTTCATGTTTTTATCTAATTTGTTAGCATATTGAATGATTGATTTTCTTTTAAT
TTGGATATGTTGATTGTCTTGTTGCATCATCAATGTATGTTTTATTTAACACCGGAAGATCTTATG
ATGGGTTCATTACTTCATAATAATCTCCGAGTTCTACAAGACTACAACTTTCACGTGACTTTTACA
GCGACAAAAAATGCATCTAGCGAAAATTAATCCACAACCTATGCATTTTTGTCACTCTTCACACGC
GTATGTGCATAAATATATAGTATATACTCGACAATCGATGCGTATGTGTACACAATTACCAAAACA
ATTATTTGAATATTCAGACATGGGTTGACATCACCAAGTAATATTCACAGTATCTGAAAACTATGT
TTTGACATCCCTAAATAGTTTGACTAACCAGTTTAATATGAGAGCATTTGTAAGAGGCAAGAGCCA
TGGTTTTGTTGGCTCGTTTAATATGCTCATTTAACCCCCCCAAAAAATACTATTAGATTTAAACGT

AAAAGAATTAACGAACACAAGAACTGCTAAAACAAAAAAAAATCAATGGCCGACATTTCATAGTTC
ATACATCACTAATACTAAAAGATGCATCATTTCACTAGGGTCTCATGAAATAGGAGTTGACATTTT
TTTTTGTAACGACAGAAGTTGACATGTTAAGCATCAATTTTTTTAAGAGTGGATTATACTAGTTTT
TTTTTTTTTTTTTAATGTATGGTATGATACAACAACAAAAACTATAAAATAGAAAAAGTCAGTGAA
ACCTCAAATTGAAGGAAAAACTTTTGCACAAAAAGAGAGAGAGAGAGAAAGAATGTAAATCCAAAT
AAATGGGCCTAATTGAGAATGCTTTAACTTTTTTTTTTTGGCTAAAAGAGAATGCTTTAACTAAGC
CCATAAAATGAACATCAAACTCAAAGGGTAAGATTAATACATTTAGAAAACAATAGCCGAATATTT
AATAAGTTTAAGACATAGAGGAGTTTTATGTAATTTAGGAACCGATCCATCGTTGGCTGTATAAAA
AGGTTACATCTCCGGCTAACATATCGGCAAAAAAGGAACCTCGAG SEQ ID NO.45

The agropine synthase polyA signal (*ags* terminator) was synthesized based on the gene bank sequence EU181145.

The ags terminator sequence

GAATTAACAGAGGTGGATGGACAGACCCGTTCTTACACCGGACTGGGCGCGGGATAGGATATTCAG
ATTGGGATGGGATTGAGCTTAAAGCCGGCGCTGAGACCATGCTCAAGGTAGGCAATGTCCTCAGCG
TCGAGCCCGGCATCTATGTCGAGGGCATTGGTGGAGCGCGCTTCGGGGATACCGTGCTTGTAACTG
AGACCGGATATGAGGCCCTCACTCCGCTTGATCTTGGCAAAGATATTTGACGCATTTATTAGTATG
TGTTAATTTTCATTTGCAGTGCAGTATTTTCTATTCGATCTTTATGTAATTCGTTACAATTAATAA
ATATTCAAATCAGATTATTGACTGTCATTTGTATCAAATCGTGTTTAATGGATATTTTTATTATAA
TATTGATGAT SEQ ID No.46

Plant Transformation

Transformation of Arabidopsis Plants

Transformation of *Arabidopsis* plants was performed as described by Clough & Bent (Clough & Bent (1998) Plant Journal 16:735-743). *Agrobacterium tumefacience* strain GV3101 (Koncz & Schell (1986) Mol Gen Genet 204:383-396) was used for transformation. Transformation of plants was carried out with chloroplast transformation constructs (Figure 2) based on the pGreen 0029 binary vector (Hellens et al (2000) Plant Mol. Biol **42**: 819-832). In brief, a chloroplast transformation cassette containing *trnI* flank, Prrn promoter, aadA gene, psbA 3′ UTR, *trnA* flank and primer binding domain (PBD) was inserted into domain IV of the LtrB or fused to CTL from ASB using AscI-NotI enzymes. The resulting DNA fragment was fused to the 35S promoter and nos terminator and introduced into the pGreen0029 binary vector

(EU048864). The fragment of LtrASi was fused to a chloroplast transit peptide (rbcS-cTP) and ubiq3 promoter from *Arabidopsis*. Resulting cassette was inserted into pGreen 0029 together with the chloroplast transformation cassette. The reverse transcriptase-RNase H (RTRHi-Ty1) was fused to HSP60-cTP transit peptide, TAF2 promoter and *ags* terminator. The resulted cassette was inserted in pSOUP vector (EU048870) carrying T-DNA from pGreen0179 vector (EU048866). The construct carrying the chloroplast cassette and LtrASi was co-transform with construct carrying RTRHi-Ty1 cassette in the same stain of Agrobacterium and used for *Arabidopsis (Col-0)* transformation.

Transgenic lines were recovered on selection medium supplemented with 100mg/l of spectinomycin.

Transformation of tobacco plants

Tobacco plants were transformed as described by Horsch et al.,(1985) Science 227: 1229-1231, using *Agrobacterium* strain AGL1 (see protocol, below).

The constructs were similar to the constructs used for *Arabidopsis* transformation with exception that *trnI* and *trnA* flanking sequences of the chloroplast cassette were amplified from tobacco genomic DNA (Figure 2).

Transgenic tobacco plants were regenerated on selection medium supplemented with 500 mg/l of spectinomycin.

Transformation of tobacco leaf explants with Agrobacterium strain AGL1

All items are autoclave-sterilised prior to use.

Filter sterilize antibiotics to prevent fungal growth, keep antibiotics for plant tissue culture in separate box

Sterilize plant material: take plants of about 9cm high, they should not have started to flower. Cut leaves with cuticle (4-6 leaves per construct, enough to cut 100 explants), dip in 70% Ethanol and immediately dip in 1% Na-hypochlorite (cat. No

01032500; use bottle of bleach that is no more than 3 months old because the chlorine gas evaporates), hold leaves with forceps and stir in for 20 min. Avoid damaging the cuticle otherwise bleach will enter the vascular system. Rinse briefly in sterile water 5-6 times and leave in water until ready to be cut.

Co-cultivation of agro with tobacco explants: grow AGL1 in LB or L broth with appropriate antibiotics overnight at 28-30°C, the next day re-suspend agro in co-cultivation solution so that the final concentration is around 0.4-0.6 OD_{600nm}. Place tobacco leaves in co-culture broth and cut squares of 1-1.5cm x 1-1.5cm with a rounded sterile scalpel using a rolling action. Dip the leaf explants in the agro solution with sterile forceps (stored in 100% ethanol, flamed and let to cool prior to touching the leaf tissue) blot on sterile Whatman™ paper and transfer on non-selective TSM plates (6 explants per plate) need to prepare about 15 plates per construct. Repeat this procedure for each construct, making sure that the scalpel and forceps are dipped in ethanol and flamed between each construct to prevent cross-contamination. Leave for 2 days only for AGL1 (3-4 days for other agro strains)

Transfer on selective TSM plates: use sterile flamed forceps to pick up and wash explants in 100 mls co-cultivation broth supplemented with timentin 320mg/l (one pot per construct), shake well, blot on sterile whatman paper and place the washed explants on selective TSM plates supplemented with appropriate selective antibiotics and timentin 320mg/l to kill agrobacterium.

Shoot regeneration: takes around 1 month to see shoots appear, explants should be transferred on fresh plates every 10-14 days. Watch out for AGL1 recurrent growth, if Timentin is not enough to kill agro, add cefotaxime at 250mg/l.

Root regeneration: Takes around 1 week. Shoots are cut from the

explants and place in growth boxes containing TRM supplemented with the appropriate selective antibiotics and timentin 320mg/l + cefotaxime 250mg/l to prevent agrobacterium recurrent growth.

Maintain plants in TRM boxes: sub them every two weeks until ready to be transferred into glasshouse

Adaptation to glasshouse conditions: soak peat pellets in sterile water until they swell to normal size and carefully place one plant per pellet, incubate the plants under 100% humidity conditions in a propagator, gradually opening the little windows until plants adapt to normal atmosphere over several days.

Recipes:

Co-culture: MS with vitamins and MES + 0.1mg/l NAA + 1mg/l BA + 3% sucrose, pH 5.7

TSM: MS with vitamins and MES + 0.1mg/l NAA + 1mg/l BA + 3% sucrose, pH5.7, 0.2% gelrite

TRM: ½ MS salts with vitamins and MES + 0.5% sucrose, pH5.7, 0.2% gelrite.

Autoclave.

Antibiotics concentration

For agrobacterium LB or L cultures:

To grow AGL1 carrying pGreen/pSOUP: Carbenicillin 100mg/l, Tetracycline 5mg/ml, Rifampicin 50mg/ml, Kanamycin 50mg/ml

AGL1 carrying pSOUP: Carbenicilin 100mg/l, Tetracycline 5mg/ml, Rifampicin 50mg/ml.

AGL1 empty: Carbenicillin 100mg/l, Rifampicin 50mg/ml.

For plant culture:

Kanamycin: 300mg/l (100mg/l if using benthamiana)

Hygromycin: 30mg/l (10mg/l if using benthamiana)

PPT: 20mg/l (2mg/l if using benthamiana)

Spectinomycin: 500mg/l

Timentin: 320mg/l. It is used to kill agro, fairly unstable make up small amount of stock, store in freezer for up to 1 month after that the antibiotic is no more efficient.

Cefotaxime: 250mg/l. Also used to kill agro, add to TS

PCR analysis of transgenic plants.

The following primers have been used for amplification of flanking junction sequences:

LFS1	GAGATGTGGATCATCCAAGGCA	SEQ ID NO.47
RFS1	CTACCATAGAGGCCAACGATAG	SEQ ID NO.48
AS527	AACGTCGGTTCGAGATGG (aadA-R1)	SEQ ID NO.49
aadA-F1	CGAAGGATGTCGCTGCCGACT;	SEQ ID NO.50

and nested primers:

LFS2	CTCCTCCTCAGGAGGATAGATG	SEQ ID NO.51
RFS2	AACTTTCATCGTACTGTGCTCTC	SEQ ID NO.52
AS526	GAGTCGATACTTCGGCGATC(aadA-R2)	SEQ ID NO.53
aadA-F2	CTAGACAGGCTTATCTTGGACA	SEQ ID NO.54

The following primers were used for amplification of chloroplast probe for Southern hybridisation:

LP-F	CGTGTTTAGTTGCCATCGTTGA	SEQ ID NO.55
LP-R	GCTGAGAGCCCTCACAGCCCA	SEQ ID NO.56
RP-F	TGTCAGCGGTTCGAGTCCGCTTA	SEQ ID NO.57
RP-R	TAACCAAGCCACTGCCTATGAGT	SEQ ID NO.58

The following primers were used for amplification of *aadA* gene as a probe for Northern hybridisation:

aadA1	GTGATCGCCGAAGTATCGACT	SEQ ID NO.59
aadA2	ATCTCGCCTTTCACGTAGTGG	SEQ ID NO.60

Results and Discussion

The transformation of *Arabidopsis* and tobacco with our vectors containing transgene cassettes generated chloroplast transgenic plants by selection on medium supplemented with 100mg/l of spectinomycin for *Arabidopsis* and 500 mg/l for tobacco (Fig. 2). In all cases we were able to detect insertion of the transgene cassette into the chloroplast genome using PCR amplification of junction regions. Five independent transgenic lines were analysed for all constructs and we could amplify correct size DNA fragment for insertion junctions in all lines. The amplified fragments were sequenced and correct insertion sites were confirmed.

Southern and Northern analysis was also performed to confirm presence of insertion and the chloroplast transcripts.

EXPERIMENTAL SECTION 1B

Modifications of the chloroplast transformation method used in Experimental section 1A can be improved using PBD designed for reverse transcription in the cytoplasm or in plastids, and by re-positioning of the building blocks on the transformation cassette (Fig 3).

The set of constructs was prepared for tobacco and rice transformation with LtrB intron (LtrB-CTS) or with ASB sequences

(ASB-CTS) as the CTS (Fig 4 - 6). The positioning of transgene cassette building blocks was designed as described in Figure 3, A-B for LtrB-CTS and Figure 3, C-D for ASB-CTS.

The PBD-CHL was designed as described previously.

The primer binding domain of the tobacco tnt1 retrotransposon was used as the PBD-CYT, and it was amplified from genomic DNA of tobacco cv Petit Gerard using the following primers:

AS912	GCCGCGGCTTTATTACCGTGAATATTA	SEQ ID NO.61
AS913	CGCGGCCGCTCTGATAAGTGCAACCTGATT	SEQ ID NO.62

PBD-CYT

CTTTATTACCGTGAATATTATTTTGGTAAGGGGTTTATTCCCAACAACTGGTATCAGAGCACAGGT
TCTGCTCGTTCACTGAAATACTATTCACTGTCGGTAGTACTATACTTGGTGAAAAATAAAAATGTC
TGGAGTAAAGTACGAGGTAGCAAAATTCAATGGAGATAACGGTTTCTCAACATGGCAAAGAAGGAT
GAGAGATCTGCTCATCCAACAAGGATTACACAAGGTTCTAGATGTTGATTCCAAAAAGCCTGATAC
CATGAAAGCTGAGGATTGGGCTGACTTGGATGAAAGAGCTGCTAGTGCAATCAGGTTGCACTTATC
AGA

SEQ ID NO.63

In the first case, the PBD-CHL was fused to the 3'end of the LtrB intron (Fig 3A, Fig 4A for tobacco and Fig 5A for rice). As LtrA protein possesses both LtrB-CTS-binding feature and reverse transcription activity it can fulfil both functions of the transgene RNA translocation into plastids and reverse transcription of the RNA cassette using plastid tRNA-Met as a primer.

In the second case, the PBD-CYT was fused to CTU (Fig 3B, Fig 4B for tobacco and Fig 5B for rice), so that reverse transcription of the transgene cassette is initiated and performed by endogenous reverse transcriptases in the cytoplasm using cytoplasmic tRNA-Met. The LtrA protein serves as CTS-binding peptide for translocation of RNA:DNA complex initiated by the reverse transcriptases into the plastids.

The ASB-CTS was fused to the CTU with PBD-CHL or PDB-CYT (Fig3, C-D, Fig 6, A and B). The reverse transcriptase from yeast *tyl* retrotransposon was co-delivered with construct containing PBD-CHL to facilitate reverse transcription reaction in the plastids.

The chloroplast cassette for rice transformation was designed using rice-specific sequences (Fig 5).

The trnI fragment of the rice chloroplast genome was utilised as the LFS, and it was amplified using the following primers:

AS699	GGCGCGCCGTGGGATCCGGGCGGTCCG	SEQ ID NO.64
AS700	GGCATGCTGGCGCAGCTGGGCCATCC	SEQ ID NO.65

Rice trnI-LFS

Gggatccgggcggtccgggggggggcactacggctcctctcttctcgagaatccatacatcccttat
cagtgtatggagagctatctctcgagcacaggttgaggttcgtcctcaatgggaaaatggagcacc
taacaacgcatcttcacagaccaagaactacgagatcacccttcattctggggtgacggagggat
cgtaccattcgagccttttttcatgcttttcccggcggtctggagaaagcagcaatcaataggac
ttccctaatcctcccttcctgaaaggaagaacgtgaaattctttttccttccgcagggaccagga
ggttggatctagccataagaggaatgcttggtataaataagccacttcttggtcttcgactcccta
agtcactacgagcgccctcgatcagtgcaatgggatgtggctatttatctatctcttgactcgaaa
tgggagcagagcaggtttgaaaaaggatcttagagtgtctagggttgggccaggagggtctcttaa
cgccttcctttttctgcccatcggagttatttcccaaggacttgccatggtaaggggggagaagggg
aagaagcacacttgaagagcgcagtacaacggagagttgtatgctgcgttcgggaaggatgaatcg
ctcccgaaaaggagtctattgattctctcccaattggttggatcgtaggggcgatgatttacttca
cgggcgaggtctctggttcaagtccaggatggcccagctgcgcca SEQ ID NO.66

The trnA fragment of the rice chloroplast genome was used as the RFS, and it was amplified using following primers:

AS701	gatatcggatggcccagctgcgcca	SEQ ID NO.67
AS702	Gcggccgcattgcccttctccgaccct	SEQ ID NO.68

Rice trnA-RFS

Ggatggcccagctgcgccagggaaaagaatagaagaagcatctgactctttcatgcatactccact
tggctcgggggatatagctcagttggtagagctccgctcttgcaattgggtcgttgcgattacgg
gttggctgtctaattgtccaggcggtaatggtagtatcttgtacctgaaccggtggctcacttttt
ctaagtaatggggaagaggactgaaacatgccactgaaagactctactgagacaaaaagatgggct
gtcaaaaaggtagaggaggtaggatgggcagttggtcagatctagtatggatcgtacatggacgat
agttggagtcggcggctctcctaggcttccctcatctgggatccctggggaagaggatcaagttgg
cccttgcgaatagcttgatgcactatctcccttcaacccttgagcgaaatgtggcaaaaggaagg
aaaatccatggaccgaccccattatctccaccccgtaggaactacgagatcaccccaaggacgcct
tcggcgtccaggggtcacggaccgaccatagaccctgttcaataagtggaacacattagccgtccg
ctctccggttgggcagtaagggtcggagaagggcaat

SEQ ID NO.69

The chloroplast-specific rrn16 promoter from wheat cv. Pavon was amplified using PCR with the following primers:

AS518 TATCGATAACATTCCTCTAATTTCATTGCA SEQ ID NO.70

AS720 GGCATGCAGGCTTGTGGGATTGACGTGATAG SEQ ID NO.71

Wheat rrn promoter sequence (Wrrn)

Aggcttgtgggattgacgtgatagggtagggttggctatactgctggtggcgaactccaggctaat
aatctgaagcgcatggatacaagttatccttggaaggaaagacaattccgaatctgctttgtctac
gaataaggaagctataagtaatgcaactatgaatctcatg

SEQ ID NO.72

aadA-mGFP4 fusion sequence

atggcagaagcggtgatcgccgaagtatcgactcaactatcagaggtagttggcgtcatcgagcgc
catctcgaaccgacgttgctggccgtacatttgtacggctccgcagtggatggcggcctgaagcca
cacagtgatattgatttgctggttacggtgaccgtaaggcttgatgaaacaacgcggcgagctttg
atcaacgacctttttggaaacttcggcttcccctggagagagcgagattctccgcgctgtagaagtc
accattgttgtgcacgacgacatcattccgtggcgttatccagctaagcgcgaactgcaatttgga
gaatggcagcgcaatgacattcttgcaggtatcttcgagccagccacgatcgacattgatctggct
atcttgctgacaaaagcaagagaacatagcgttgccttggtaggtccagcggcggaggaactcttt
gatccggttcctgaacaggatctatttgaggcgctaaatgaaaccttaacgctatggaactcgccg

cccgactgggctggcgatgagcgaaatgtagtgcttacgttgtcccgcatttggtacagcgcagta
accggcaaaatcgcgccgaaggatgtcgctgccgactgggcaatggagcgcctgccggcccagtat
cagcccgtcatacttgaagctagacaggcttatcttggacaagaagaagatcgcttggcctcgcgc
gcagatcagttggaagaatttgtccactacgtgaaaggcgagatcaccaaggtagtcggcaaatca
ggatccatgagtaaaggagaagaacttttcactggagttgtcccaattcttgttgaattagatggt
gatgttaatgggcacaaattttctgtcagtggagagggtgaaggtgatgcaacatacggaaaactt
acccttaaatttatttgcactactggaaaactacctgttccatggccaacacttgtcactactttc
tcttatggtgttcaatgcttttcaagatacccagatcatatgaagcggcacgacttcttcaagagc
gccatgcctgagggatacgtgcaggagaggaccatcttcttcaaggacgacgggaactacaagaca
cgtgctgaagtcaagtttgagggagacaccctcgtcaacaggatcgagcttaagggaatcgatttc
aaggaggacggaaacatcctcggccacaagttggaatacaactacaactcccacaacgtatacatc
atggcagacaaacaaaagaatggaatcaaagttaacttcaaaattagacacaacattgaagatgga
agcgttcaactagcagaccattatcaacaaaatactccaattggcgatggccctgtccttttacca
gacaaccattacctgtccacacaatctgccctttcgaaagatcccaacgaaaagagagaccacatg
gtccttcttgagtttgtaacagctgctgggattacacatggcatggatgaactatacaaataatct
aga

SEQ ID NO.73

atpA terminator was amplified from wheat DNA using the following primers:

AS753 Accgcggtcaaataaattttgcatgtcta SEQ ID NO.74

AS723 Gatatctccatactccttctttatgata SEQ ID NO.75

Wheat atpA terminator

Caaataaattttgcatgtctactcttgttagtagaataggaatcgttgagaaagatttttcatttg
aatcatgcaaaaaagttttctttgtttttagtttagtatagttatttaaagaatagatagaaataa
gattgcgtccaataggatttgaacctataccaaaggtttagaagacctctgtcctatccattagac
aatggacgcttttctttcatattttattctttctttttatttttttttcttcttccgagaaaaaact
gttagaccaaaactcttttaggaaatcaaaaaatccagatacaaatgcatgatgtatatattatat
catgcatatatcataaagaaggagtatgga

SEQ ID NO.76

The LtrA gene was driven by actin1 rice promoter amplified using the following primers:

ARP1 gtcattcatatgcttgagaaga SEQ ID NO.77

ARP2 gcctacaaaaaagctccgcacg SEQ ID NO.78

Rice act1 promoter sequence

gtcattcatatgcttgagaagagagtcgggatagtccaaaataaaacaaaggtaagattacctggt
caaaagtgaaaacatcagttaaaaggtggtataagtaaaatatcggtaataaaaggtggcccaaag
tgaaatttactcttttctactattataaaaattgaggatgttttgtcggtactttgatacgtcatt
tttgtatgaattggtttttaagtttattcgcgatttggaaatgcatatctgtatttgagtcggttt
ttaagttcgttgcttttgtaaatacagagggatttgtataagaaatatctttaaaaaacccatatg
ctaatttgacataatttttgagaaaaatatatattcaggcgaattccacaatgaacaataataaga
ttaaaatagcttgcccccgttgcagcgatgggtatttttttctagtaaaataaaagataaacttaga
ctcaaaacatttacaaaaacaacccctaaagtcctaaagcccaaagtgctatgcacgatccatagc
aagcccagcccaacccaacccaacccaacccaccccagtgcagccaactggcaaatagtctccacc
cccggcactatcaccgtgagttgtccgcaccaccgcacgtctcgcagccaaaaaaaaaaaaagaaa
gaaaaaaaagaaaaagaaaaacagcaggtgggtccgggtcgtgggggccggaaaagcgaggaggat
cgcgagcagcgacgaggcccggccctccctccgcttccaaagaaacgccccccatcgccactatat
acataccccccctctcctcccatcccccccaaccctaccaccaccaccaccaccacctcctccccc
ctcgctgccggacgacgagctcctcccccctcccctccgccgccgccggtaaccaccccgccccct
ctcctctttctttctccgttttttttttcgtctcggtctcgatctttggccttggtagtttgggtg
ggcgagagcggcttcgtcgcccagatcggtgcgcgggaggggcgggatctcgcggctggcgtctcc
gggcgtgagtcggcccggatcctcgcggggaatggggctctcggatgtagatctgcgatccgccgt
tgttgggggagatgatgggggggtttaaaatttccgccatgctaaacaagatcaggaagaggggaaa
agggcactatggtttatatttttatatatttctgctgcttcgtcaggcttagatgtgctagatctt
ctttctttcttcttttttgtggtagaatttgaatccctcagcattgttcatcggtagtttttctttt
catgatttgtgacaaatgcagcctcgtgcggagctttttttgtaggc

SEQ ID NO.79

Transformation of Rice Immature Embryos.

Immature Embryo Excision

Day 1:

Remove milky/post-milky stage immature seeds from panicles (immature embryos 1-2 mm in size are desired).

Sterilize immature seeds: 50% sodium hypochlorite (12%) + 1 drop of tween 20. Shake 10 min.

Rinse 3-5x in sterile deionised water. Drain off surplus water. Aliquot seeds (around 40) in sterile Petri dishes.

Set up a 60 x 15 mm Petri dish containing a 50% sodium hypochlorite solution and next to this a sterile beaker on its side with a sterile filter paper in it. Use sterile forceps to aseptically remove glumes from the first seed. Immerse this seed in the 50% sodium hypochlorite. Remove glumes from a second seed and immerse the second seed into the sodium hypochlorite solution whilst removing the first seed and storing this dehusked/sterilized seed on the filter paper in the beaker. Continue in this manner with all seeds.

After all the glumes are removed:

Sterilize dehusked seeds: 50% sodium hypochlorite: 5 min. with agitation.

Rinse: 5-7 x in sterile deionized water, drain.

Place all seeds in a large sterile Petri dish. Aliquot for embryo excision (to keep seeds from drying out, work with only 50-100 in the plate at a time leaving the rest in the master plate).

Remove the embryo from each seed and place embryo, scutellum up, in a 90 x 15 mm Petri dish containing proliferation medium (40-50 embryos / plate). Culture at 28eC in the dark for 2 days prior to bombardment

Day 3:

Check each embryo for contamination before blasting

Remove the embryos from the proliferation medium. Distribute 35-40 embryos scutellum upwards in an area 1 cm^2 in the centre of a 60 x 15 mm target plate containing 10 ml of proliferation medium + osmoticum (0.6M). Check each target plate so that the scutellum is

straight. Allow enough room so the scutella do not shade each other out.

Bombardment:

Gun	14 kV
	Vacuum : 25 inches of Hg
1st bombardment	4 hours after osmoticum treatment
2nd bombardment	4 hours after 1st bombardment

Day 4:

4-16 hours after the 2nd blast transfer immature embryos to proliferation medium without osmoticum. Culture in the dark at 28°C for 2 days.

Selection:

Day 5:

Aseptically cut out with scissors the germinating shoot. Transfer 16 - 20 immature embryos to fresh proliferation medium containing 30-50 mg/l Hygromycin (depending on the genotype); culture in the dark at 28°C; record total number of embryos.

After 10 days carefully remove the callus from the scutellum by breaking it up into 2-10 small pieces; subculture onto fresh proliferation medium + hygromycin. Do not subculture brown tissue and remaining immature embryo which could inhibit further growth of healthy callus.

Subculture every 10 days by selecting healthy tissue: (embryogenic if present) and transfer it to fresh proliferation medium + hygromycin. Remove brown callus as it could be inhibiting to embryogenic callus.

30 to 40 days after bombardment change selection procedure. Instead of eliminating bad-looking tissue keep embryogenic tissue only (eliminate healthy non-embryogenic tissue)

Regeneration:

After 40 to 60 days, transfer established embryogenic callus showing differential growth on proliferation medium + hygromycin to regeneration medium + hygromycin. Culture at 28ɛC under low light for 10 days then under high light for 10 additional days. Check plates periodically in the light for the development of embryos and green shoots. As shoots develop it is sometimes beneficial to gently move the developing shoot away from the callus it originated from and remove any dead tissue from the shoot itself to prevent inhibition of growth.

Germination:

Transfer white compact embryos and green shoots initiating roots to the germination medium under high light at 28ɛC for 1 to 2 weeks. Check plates periodically. Remove necrotic tissue and divide germinating embryos if necessary.

Results

The analysis of transgenic plants was performed using PCR for insertion flanking sequences using the following primers for tobacco left flank:

AS548	ACGGTGAAGTAAGACCAAGCTCAT	SEQ ID NO.80
AS549	CTAGGTCGGAACAAGTTGATAGGAC	SEQ ID NO.81

right flank:

AS550	GGCTATGCCATCCTAAGGTGCTGCT	SEQ ID NO.82
AS551	CCATGAATGATAAATCATAGATCGAAC;	SEQ ID NO.83

for rice left flank:

RC1	CCTGACCCGAAGATGTGGATC	SEQ ID NO.84
RC2	ACATTAGCATGGCGTACTCCT	SEQ ID NO.85

right flank

RC3	AACCAGGAACGGGGAGCTCTC	SEQ ID NO.86
RC4	CGACTCTTTGATCTTAAACTT	SEQ ID NO.87

Internal primers specific for aadA gene:

AS526	GAGTCGATACTTCGGCGATC	SEQ ID NO.88
AS527	AACGTCGGTTCGAGATGG	SEQ ID NO.89

for mGFP gene

AS528	TTACCAGACAACCATTACCTGTC	SEQ ID NO.90
AS529	GCTGGGATTACACATGGCAT	SEQ ID NO.91

The expected size (1.1kb) of PCR products were obtained for all tobacco constructs (Fig 7). Sequencing analysis has confirmed junction site between transgene and plastid genome.

Southern analysis has also confirmed transgene insertions into the correct location of the tobacco chloroplast genome (Fig 8).

Northern analysis indicated presence of transgene transcript in the fraction of the chloroplast RNA (Fig 9).

WE CLAIM:

1. A method of transforming a plastid of a plant cell that comprises:

    1) introducing into the said plant cell a nucleic acid molecule that comprises a plant nuclear promoter operably linked to a first nucleic acid molecule that comprises a plastid transgene cassette, a plastid translocation molecule (PTS) and a primer binding domain (PBD);

    2) introducing into the said plant cell a second nucleic acid molecule that encodes for a plastid translocation molecule binding protein fused to a first plastid transit peptide (PTSBP-TP), wherein said second nucleic acid molecule is operably linked to a plant nuclear promoter; and

    3) introducing into the said plant cell a third nucleic acid molecule that encodes for a reverse transcriptase protein fused to a second plastid transit peptide, wherein the third nucleic acid molecule is operably linked to a plant nuclear promoter that drives expression in a plant cell nucleus.

2. The method according to claim 1, wherein the plastid transgene cassette comprises:

    i) A left flanking molecule (LFS) and a right flanking molecule (RFS), wherein the LFS and RFS are selected from any nucleotide molecules that may be used for homologous recombination in the plastid; and

    ii) at least one recombinant DNA molecule or an introduced native, isolated genomic DNA molecule.

3. The method according to claim 1, wherein the plant plastid transgene cassette comprises at least one plastid specific promoter and at least one plastid specific terminator molecule.

4. The method according to claim 2, wherein the at least one recombinant DNA molecule or the introduced native, isolated genomic DNA molecule is selected from isolated mammalian and plant nucleic acid molecules.

5. The method according to claim 2 or claim 4, wherein the at least one recombinant DNA molecule is a cDNA molecule.

Date Reçue/Date Received 2021-09-24

6. The method according to either of claims 4 or 5, wherein a transgene comprised in the transgene cassette or the isolated genomic DNA molecule is selected from insulin, preproinsulin, proinsulin, glucagon, interferons, blood-clotting factors, fertility hormones, luteinising hormone, follicle stimulating hormone, growth factors, epidermal growth factor, platelet-derived growth factor, granulocyte colony stimulating factor, prolactin, oxytocin, thyroid stimulating hormone, adrenocorticotropic hormone, calcitonin, parathyroid hormone, somatostatin, erythropoietin (EPO), enzymes, haemoglobin, serum albumin, collagen, biotic stress proteins, abiotic stress proteins, insecticidal proteins, insect toxic proteins, nematicidal proteins, herbicide resistance proteins, salt-tolerance proteins; drought tolerant proteins, proteins that confer cytoplasmic male sterility to plant breeding lines; nutritional enhancement proteins involved in the biosynthesis of phenolics, starches, sugars, alkaloids, vitamins, and edible vaccines, monoclonal antibodies, active fragments of the monoclonal antibodies, industrial enzymes, or active fragments of the industrial enzymes, wherein the blood-clotting factors are one or more of Factor VII, VIII, IX, X, XI and XII.

7. The method according to claim 6, wherein the enzymes are β-glucocerebrosidase.

8. The method according to claim 1, wherein the primer binding domain is selected from one or more of a retrotransposon and a retrovirus.

9. The method according to claim 8, wherein the primer binding domain is from the Ty1 retrotransposon from yeast.

10. The method according to claim 1, wherein the PTS molecule is selected from naked RNA viroids, viruses, viral coat protein binding domains, group I and group II intron RNA, retrotransposon primer binding sites, and RNA harbouring a domain that is recognised by RNA binding proteins.

11. The method according to claim 10, wherein the PTS is the group II intron-derived PTS from the *Lactococus lactis* LtrB intron.

12. The method according to claim 1, wherein the third nucleic acid molecule comprises a reverse transcriptase protein from a retrotransposon source or a retroviral source.

13. The method according to claim 12, wherein the reverse transcriptase nucleic acid

Date Reçue/Date Received 2021-09-24

molecule is from the yeast retrotransposon Ty1 and is reverse transcriptase-RNase H.

14. The method according to claim 1, wherein the plastid transit peptide of the second and third nucleic acid molecule is independently chloroplast transit peptide, wherein the chloroplast transit peptide is selected from the tobacco rsbc-cTP, and the Arabidopsis HSP70-cTP protein.

15. A method of producing an RNA species in a plant that comprises:

1) introducing into a regenerable plant cell a nucleic acid molecule that comprises a plant nuclear promoter operably linked to a first nucleic acid molecule that comprises a plant plastid transgene cassette, a plant plastid translocation molecule (PTS) and a primer binding domain, and expressing the nucleic acid in the plant cell from the plant nuclear promoter;

2) introducing into the said regenerable plant cell a second nucleic acid molecule that encodes for a translocation molecule binding protein fused to a plant plastid transit peptide, wherein said second nucleic acid molecule is operably linked to a plant nuclear promoter, and expressing the nucleic acid in the plant cell from the plant nuclear promoter; and

3) introducing into the said regenerable plant cell a third nucleic acid molecule that encodes for a reverse transcriptase protein fused to a plant plastid transit peptide, wherein the third nucleic acid molecule is operably linked to a plant nuclear promoter, and expressing the nucleic acid in the plant cell from the plant nuclear promoter;

4) growing said regenerable plant cell of step 3);

5) selecting a plant cell of 4) wherein a transgene comprised within the plant plastid transgene cassette is integrated into the plastid genome;

6) regenerating a plant from the plant cell of 5); and

7) growing the plant of 6).

16. The method according to claim 15, wherein the RNA species encoded by the transgene that is integrated into the plastid is expressed as a heterologous or exogenous

protein.

17. An isolated polynucleotide molecule that comprises a plant nuclear promoter operably linked to a first nucleic acid molecule that comprises a plant plastid transgene cassette, a plant plastid RNA translocation molecule (PTS), and a primer binding domain for use in a method according to claim 1.

18. An isolated polynucleotide molecule that encodes for a plastid translocation molecule-binding protein fused to a plant plastid transit peptide wherein the polynucleotide molecule is operably linked to a plant nuclear promoter for use in a method according to claim 1.

19. An isolated polynucleotide molecule that encodes for a reverse transcriptase protein fused to a plant plastid transit peptide wherein the polynucleotide molecule is operably linked to a plant nuclear promoter for use in a method according to claim 1.

20. An isolated polynucleotide molecule according to claim 17 or claim 18, comprising genomic DNA for use in a method according to claim 1.

21. An isolated polynucleotide molecule according to claim 17 or claim 18, comprising cDNA for use in a method according to claim 1.

22. An isolated polynucleotide molecule according to claim 19, comprising genomic DNA for use in a method according to claim 1.

23. An isolated polynucleotide molecule according to claim 19, comprising cDNA for use in a method according to claim 1.

24. A nucleic acid vector for use in a method according to claim 1 in the transformation of a plant cell.

25. A nucleic acid vector for use in a method according to claim 1 in the transformation of a plant cell.

26. A host cell having a transformed plastid, the host cell containing:

i) a heterologous polynucleotide molecule that comprises a plant nuclear promoter operably linked to a first nucleic acid molecule that comprises a plant plastid transgene cassette, a plant plastid RNA

translocation molecule (PTS), and a primer binding domain; and

ii) a heterologous polynucleotide molecule that encodes for a reverse transcriptase protein fused to a plant plastid transit peptide wherein the polynucleotide molecule is operably linked to a plant nuclear promoter

for use in a method according to claim 1.

27. The host cell according to claim 26 which is a plant cell.

28. A method of producing a host cell, the host cell having a transformed plastid, the method comprising:

i) incorporating into the cell a polynucleotide molecule comprising a plant nuclear promoter operably linked to a first nucleic acid molecule that comprises a plant plastid transgene cassette, a plant plastid RNA translocation molecule (PTS) and a primer binding domain; and

ii) incorporating into the cell a polynucleotide molecule that encodes for a reverse transcriptase protein fused to a plant plastid transit peptide, wherein the polynucleotide molecule is operably linked to a plant nuclear promoter.

29. A method of producing a host cell, the host cell having a transformed plastid, the method, comprising:

i) incorporating into the cell a nucleic acid vector that comprises a polynucleotide molecule comprising a plant nuclear promoter operably linked to a first nucleic acid molecule that comprises a plant plastid transgene cassette, a plant plastid RNA translocation molecule (PTS) and a primer binding domain; and

ii) incorporating into the cell a nucleic acid vector that comprises a polynucleotide molecule that encodes for a reverse transcriptase

protein fused to a plant plastid transit peptide, wherein the polynucleotide molecule is operably linked to a plant nuclear promoter.

30. A method according to claim 28 or claim 29, wherein the host cell is a plant cell and the method includes regenerating a plant from the said cell.

31. A method of producing a non-propagating plant cell, the method including incorporating a polynucleotide molecule according to claim 17 or claim 19 into a plant cell to generate a non-propagating plant cell having a transformed plastid.

32. A method of producing a non-propagating plant cell, the method including incorporating a nucleic acid vector according to claim 24 or claim 25 into a plant cell to generate a non-propagating plant cell having a transformed plastid.

33. Use of a polynucleotide molecule that comprises a plant nuclear promoter operably linked to a first nucleic acid molecule that comprises a plant plastid transgene cassette, a plant plastid RNA translocation molecule (PTS), and a primer binding domain for producing a non-propagating plant cell having a transformed plastid.

34. Use of a polynucleotide molecule that encodes for a reverse transcriptase protein fused to a plant plastid transit peptide wherein the polynucleotide molecule is operably linked to a plant nuclear promoter for producing a non-propagating plant cell having a transformed plastid.

Date Reçue/Date Received 2021-09-24

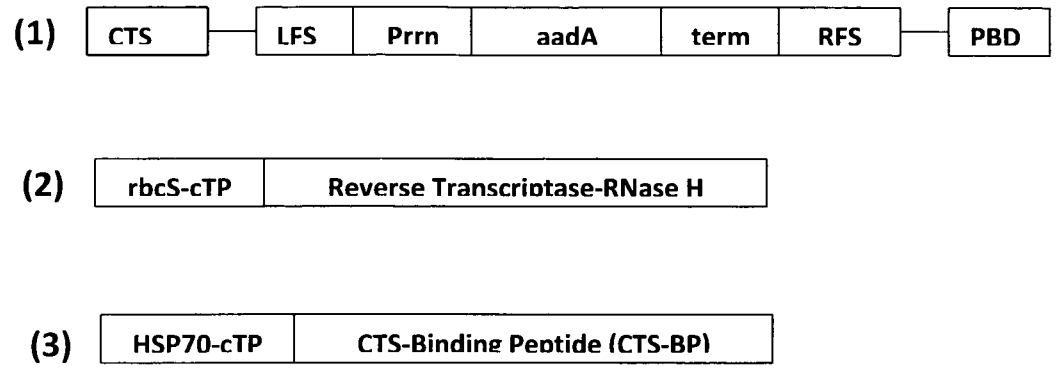


Figure 1

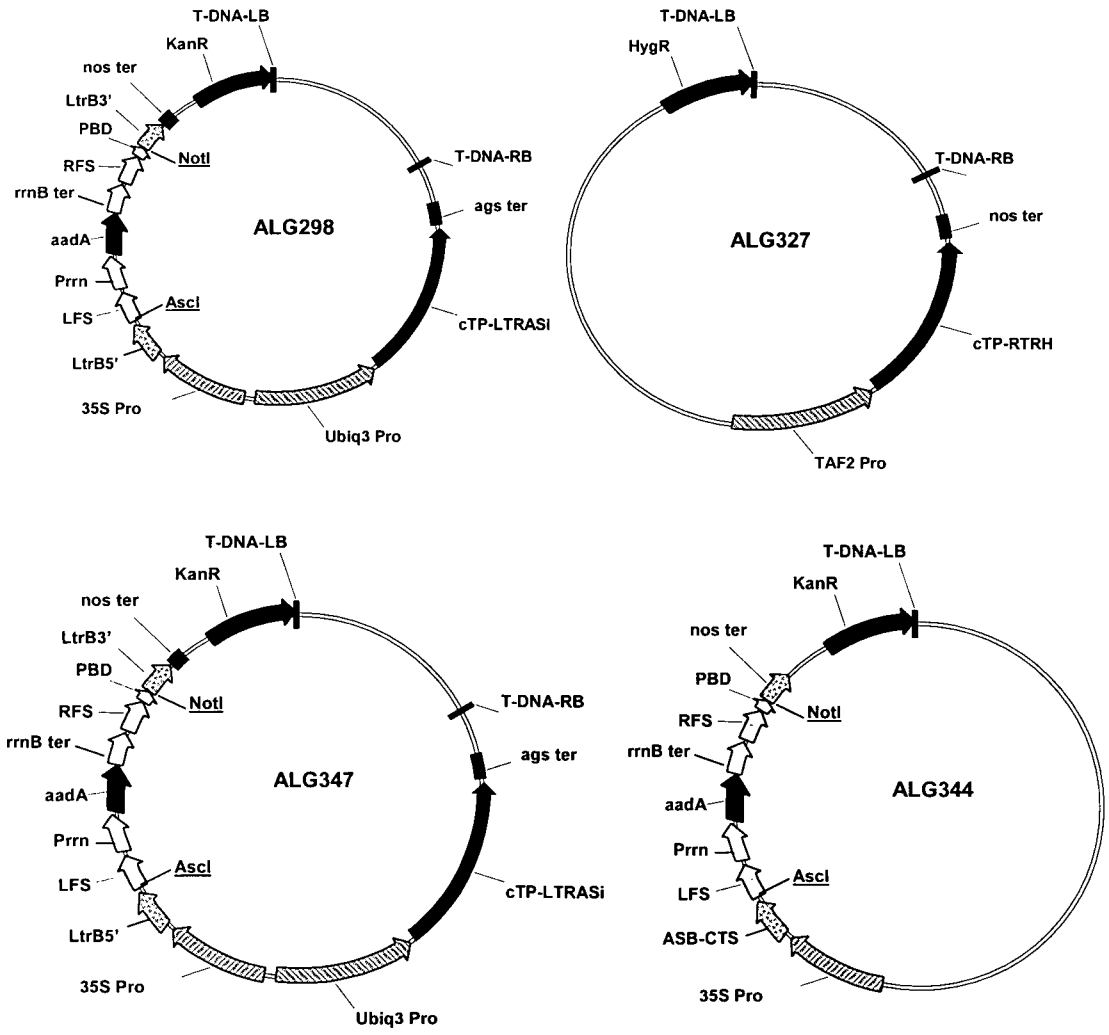


Figure 2

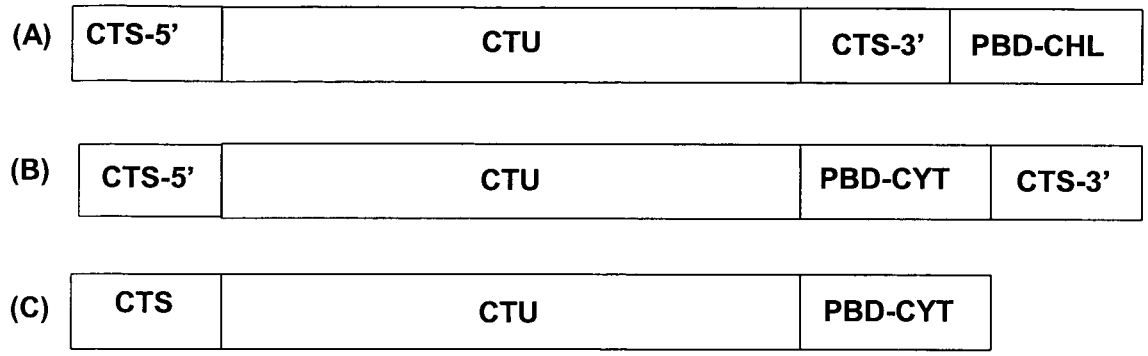


Figure 3

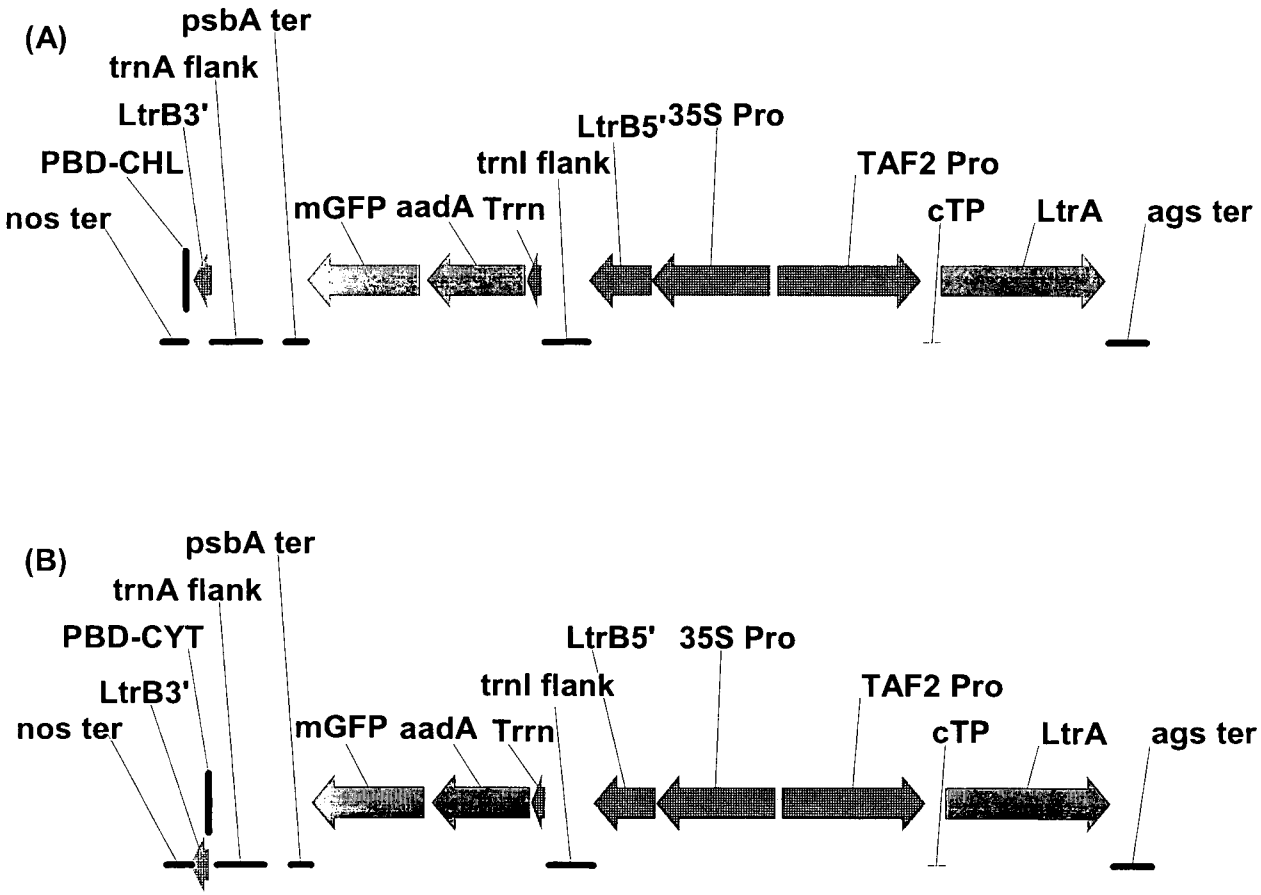


Figure 4

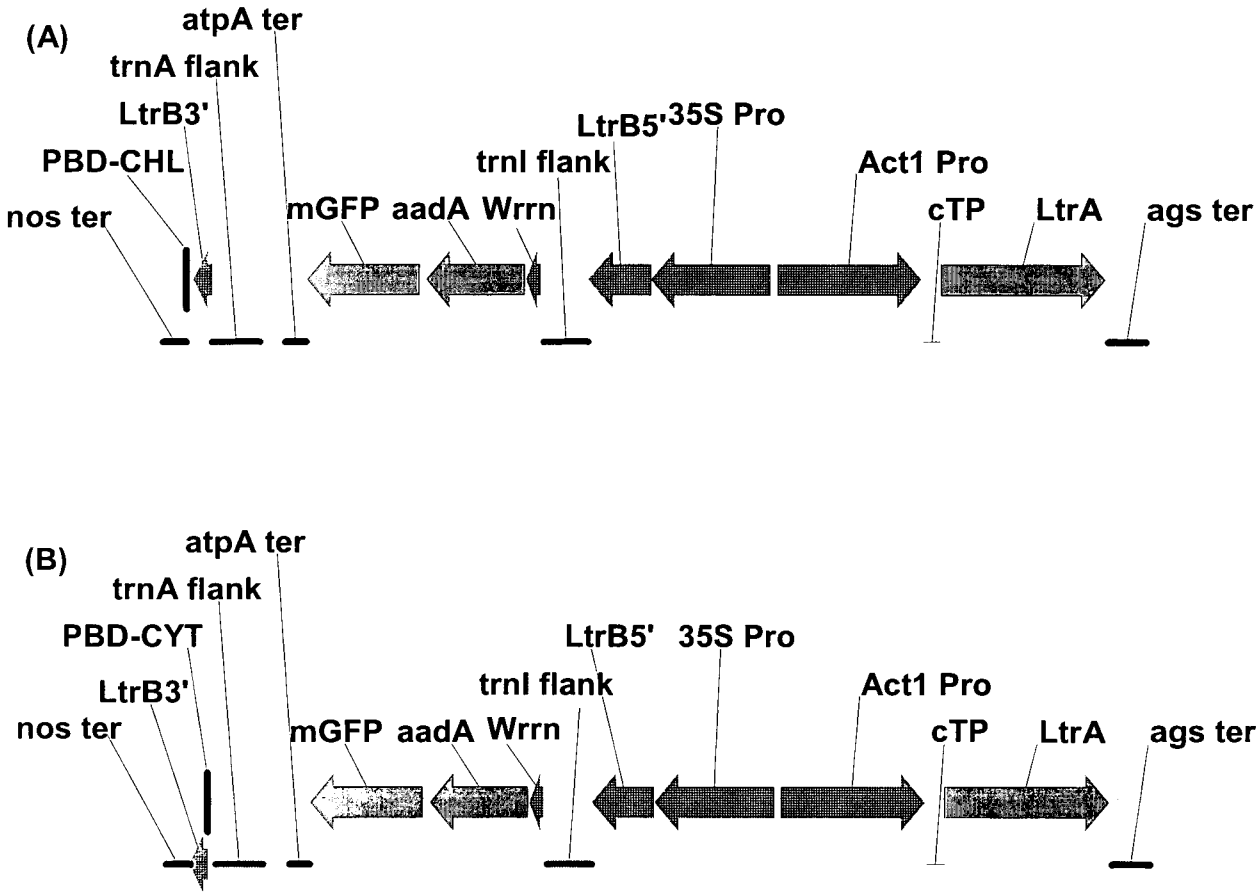

(A)
atpA ter
trnA flank
LtrB3'
PBD-CHL
nos ter
mGFP aadA Wrrn
trnI flank
LtrB5'
35S Pro
Act1 Pro
cTP
LtrA
ags ter
(B)
atpA ter
trnA flank
PBD-CYT
LtrB3'
nos ter
mGFP aadA Wrrn
trnI flank
LtrB5'
35S Pro
Act1 Pro
cTP
LtrA
ags ter

Figure 5

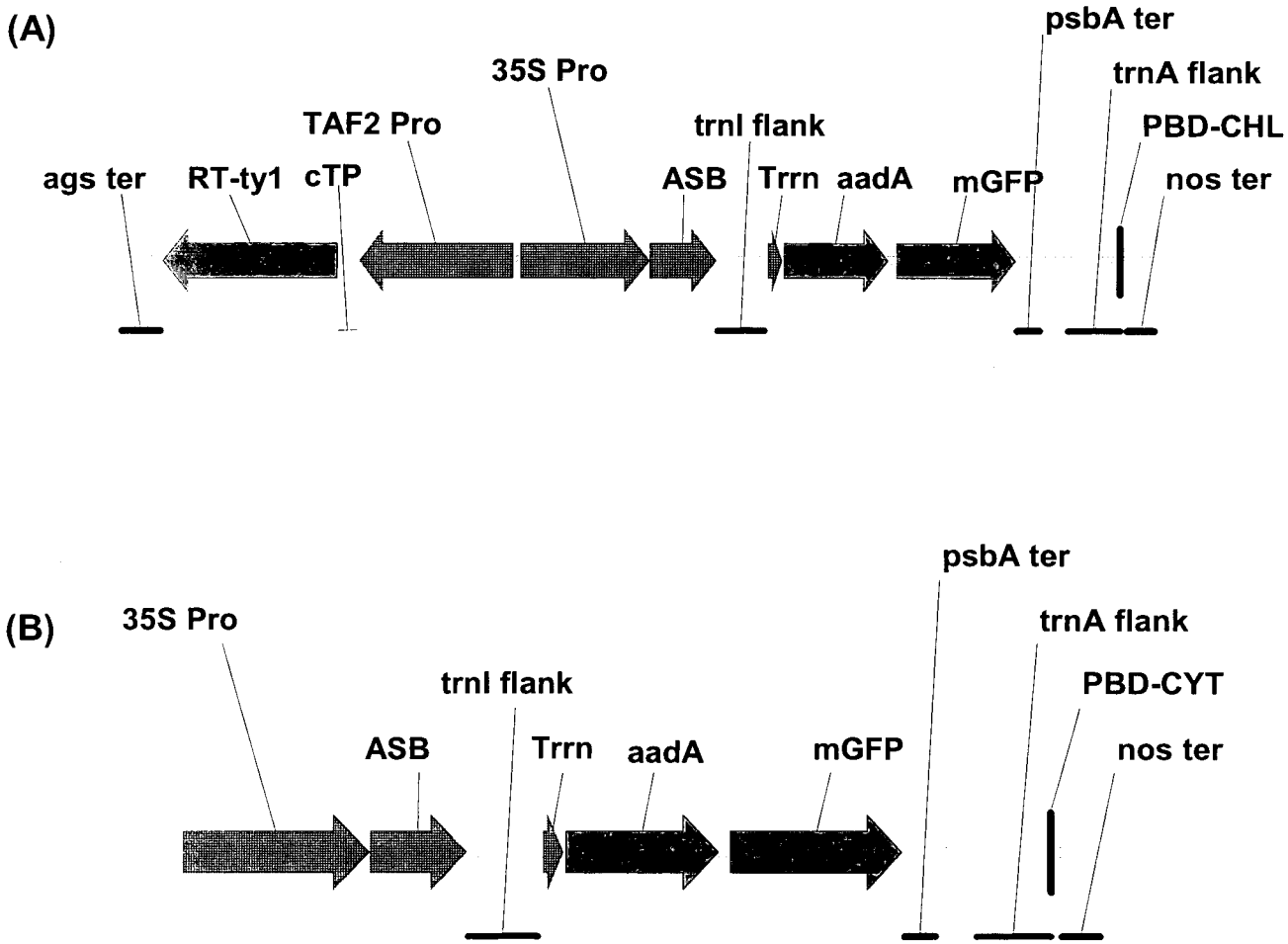

(A)
ags ter
RT-ty1
cTP
TAF2 Pro
35S Pro
ASB
trnI flank
Trrn
aadA
mGFP
psbA ter
trnA flank
PBD-CHL
nos ter
(B)
35S Pro
ASB
trnI flank
Trrn
aadA
mGFP
psbA ter
trnA flank
PBD-CYT
nos ter

Figure 6

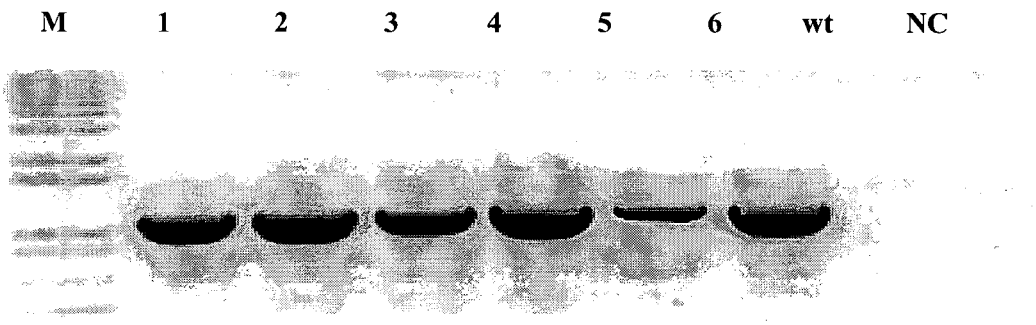


Figure 7

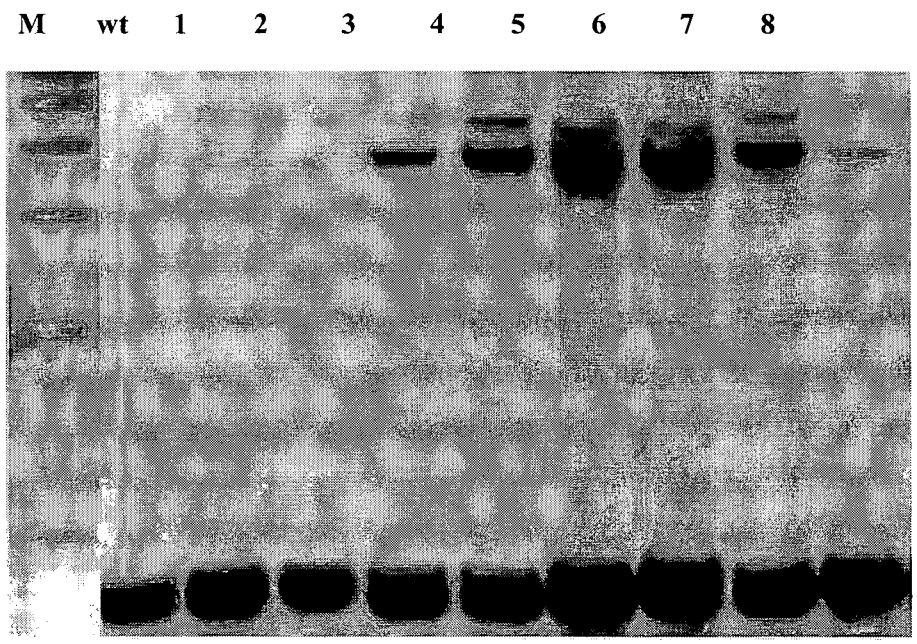


Figure 8

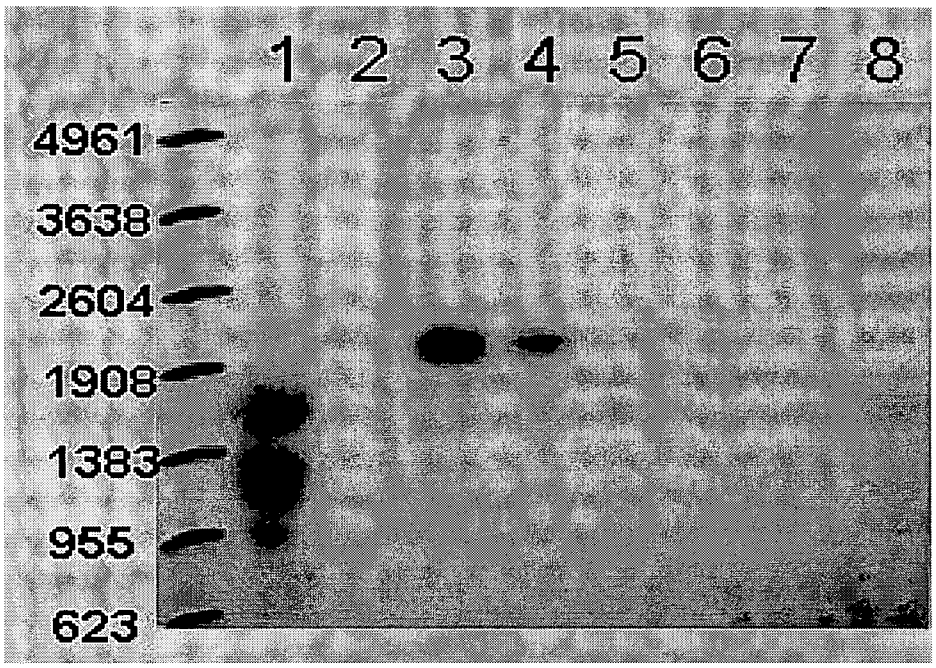
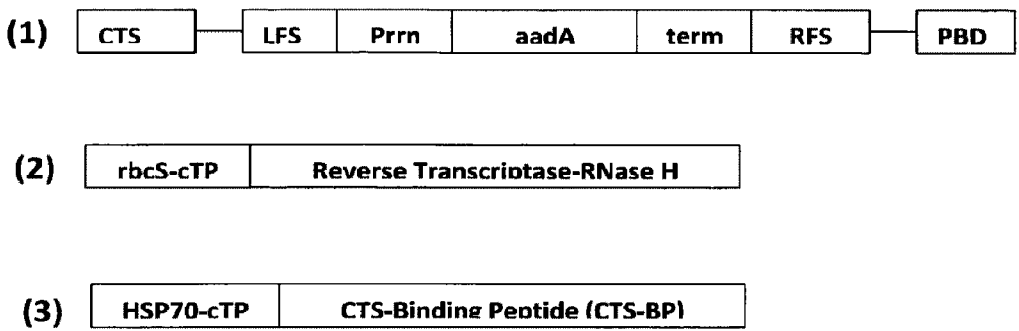


Figure 9


(1)
CTS
LFS
Prrn
aadA
term
RFS
PBD
(2)
rbcS-cTP
Reverse Transcriptase-RNase H
(3)
HSP70-cTP
CTS-Binding Peptide (CTS-BP)