

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
5 February 2009 (05.02.2009)

PCT

(10) International Publication Number
WO 2009/017821 A1

(51) International Patent Classification:
C12N 15/87 (2006.01)

(21) International Application Number:
PCT/US2008/009310

(22) International Filing Date: 31 July 2008 (31.07.2008)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/953,337 1 August 2007 (01.08.2007) US

(71) Applicant (for all designated States except US):
BIOGLOW INC. [US/US]; 175 Harbor Road, Head
Of The Harbor, NY 11780 (US).

(72) Inventor; and

(75) Inventor/Applicant (for US only): **KRICHEVSKY,**
Alexander [US/US]; 19 Orchid Drive, Port Jefferson
Station, NY 11776 (US).

(74) Agents: **FEIT, Irving, N.** et al.; Hofman & Baron, LLP,
6900 Jericho Turnpike, Syosset, NY 11791 (US).

(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, 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, 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, MT, NL,
NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG,
CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

- with international search report
- before the expiration of the time limit for amending the
claims and to be republished in the event of receipt of
amendments
- with sequence listing part of description published sepa-
rately in electronic form and available upon request from
the International Bureau

(54) Title: BIOLUMINESCENT PLANTS COMPRISING BACTERIAL LUX OPERON AND METHODS OF MAKING SAME

(57) Abstract: In one aspect, the invention relates to a transgenic bioluminescent plant including an expressible heterologous nu-
cleotide sequence comprising a bacterial LUX operon, which includes LUX A, LUX B, LUX C, LUX D, LUX E, and LUX G genes,
wherein the heterologous nucleotide sequence is expressed to render the plant bioluminescent.

WO 2009/017821 A1

BIOLUMINESCENT PLANTS COMPRISING BACTERIAL LUX OPERON
AND METHODS OF MAKING SAME

5 **CROSS-REFERENCE TO RELATED APPLICATION**

This application claims the benefit of U.S. Provisional Application No. 60/953,337, filed August 1, 2007, which is incorporated herein by reference in its entirety.

BACKGROUND OF THE INVENTION

10 The phenomenon of light emission by living organisms, known as bioluminescence, is a characteristic of different species including, for example, insects and bacteria. Most of the world's luminescent bacteria are represented by marine species, including *Vibrio*, *Photobacterium*, *Photorhabdus* and *Shewanella* families. Some of these organisms live as free water-born bacteria, while others as
15 symbionts in light organs of marine animals.

 Non-bacterial organisms such as plants that are capable of bioluminescence would be useful for many purposes, such as for environmental and aesthetic applications. However, such organisms have not been readily achieved for many reasons. For example, the genes and mechanisms responsible for bioluminescence are
20 complex.

 Genetic engineering of plants has been typically limited to introduction of one or two new genes into plant genome. This limitation prevents incorporation of complex metabolic pathways, such as those involved in light emission, into transgenic plant organisms.

25 Conventionally constructed genetic transformation vectors are made using "regular" restriction endonucleases with average recognition site length of six nucleotides. Such construction is a laborious and time consuming process, which involves sequential cloning of different functional plasmid elements (*e.g.*, promoters, terminators, and integration sequences) by a series of multiple cloning steps.

After cloning several vector segments, commonly used restriction endonucleases sites will be present in the newly inserted DNA fragments (due to high statistical probability of occurrence of a six base pair sequence in an extended DNA fragment). The presence of such sites significantly limits the number and size of additional DNA sequences that may need to be added to achieve a desired biological property such as bioluminescence. Moreover, if initially cloned DNA elements are required to be exchanged to yield another property, the entire vector typically must be reconstructed. It is often impossible to remove previously cloned genetic elements without affecting later cloned sequences.

Incorporation of multiple transgenes into a single plant organism has been attempted using standard breeding techniques. However, such an approach is time consuming and largely ineffective. Accordingly, a genetic vectors system allowing for efficient incorporation of multiple transgenes is needed. An approach for allowing rapid exchange of multiple transgenes in the genome of any selected plant species (especially the plastid genome) is also needed.

SUMMARY OF THE INVENTION

The present invention addresses these and other objectives.

In one aspect, the invention relates to a transgenic bioluminescent plant that includes an expressible heterologous nucleotide sequence comprising a bacterial LUX operon, comprising LUX A, LUX B, LUX C, LUX D, LUX E, and LUX G genes, wherein the heterologous nucleotide sequence is expressed to render the plant bioluminescent.

In another aspect, the invention relates to a method for producing a transgenic bioluminescent plant. The method includes transfecting a plant cell with a vector that includes an expressible heterologous nucleotide sequence including a bacterial LUX operon, comprising LUX A, LUX B, LUX C, LUX D, LUX E, and LUX G genes.

The present invention provides methods for producing luminescence in plants comprising transfection of a plant cell with genes corresponding to the LUX operon. In one embodiment of the invention, the transfected genes comprise the genes from the LUX operon which encode for bacterial luciferase, i.e., the LUX A and B genes;

the genes related to production of substrates required for the bacterial luciferase luminescent reaction, i.e., LUX genes: C, D, and E; and LUX G.

In one aspect, the present invention provides methods of producing luminescence in plants comprising transfection of a plant cell with genes encoding for bacterial luciferase encoded by LUX A and B genes and any other gene related to production of substrates required for the bacterial luciferase luminescent reaction including, but not limited to LUX genes: C, D, E, G, H, F, as well as any component of the RIB operon (such as LS or RS) involved in riboflavin synthesis and donors of fatty acids for the aldehyde synthesis, such as bacterial or plant acyl carrier protein (ACP). In another aspect, the invention provides corresponding transgene expression in all and any part of the plant cell from nuclear, chloroplast, plasmid or any other plant genetic sequences, as well as means of regulation for the genes expression, such as inducible promoters, etc. In a further aspect, the invention provides transfection of at least one plant cell with at least one vector encoding for the luciferase system mentioned herein and growing at least one plant cell into plant.

In yet another aspect, the invention relates to a vector for transforming a plastid. The vector includes a first and a second homologous recombination site, wherein each of said homologous recombination site is flanked by a unique rare cutting enzyme sequence; and a multiple cloning site comprising at least one unique rare cutting enzyme sequences.

In a further aspect, the invention relates to a method for producing a vector for transforming a plastid. The method includes providing an acceptor vector, which includes (i) a first and a second homologous recombination site, wherein each of said homologous recombination site is flanked by a unique rare cutting enzyme sequence; and (ii) a multiple cloning site that includes at least one unique rare cutting enzyme sequence. The method further includes providing at least one donor vector, wherein the donor vector includes a heterologous nucleotide sequence flanked by a rare cutting enzyme sequence that is sufficiently identical to the at least one unique rare cutting enzyme sequence of the multiple cloning site of the acceptor vector. The method further includes sequentially subcloning the heterologous nucleotide sequence from the donor vector into the acceptor vector, wherein a vector for transforming a plastid is produced.

In yet another aspect, the invention relates to a vector system kit. The kit includes an acceptor vector that includes (i) a first and a second homologous recombination site, wherein each of said homologous recombination site is flanked by a unique rare cutting enzyme sequence; and (ii) a multiple cloning site comprising at least one unique rare cutting enzyme sequence. The kit further includes at least one donor vector, wherein the donor vector comprises a heterologous nucleotide sequence flanked by a rare cutting enzyme sequence that is sufficiently identical to the at least one unique rare cutting enzyme sequence of the multiple cloning site of the acceptor vector.

BRIEF DESCRIPTION OF DRAWINGS

Figure 1: Schematic illustration of the biochemistry of a bacterial luminescence reaction.

Figure 2: (A) Schematic illustration of a transgene cloned into a donor vector, pCas vector. The region from the 5'-end of the leftmost rare cutter restriction site to the 3'-end of the rightmost rare cutter restriction site (or rare cutting enzyme sequence) represents an expression cassette to be introduced into plastid genome. Fig. 2(B): Schematic illustration example of a pCas vector: use of photosystem II (*psbA*) regulatory elements for transgene expression (PI-PspI-P*psbA*-L*psbA*-MCS-T*psbA*-PI-PspI (SEQ ID NO: ____). Abbreviations: P*psbA* is a *psbA* promoter; L*psbA* is a *psbA* leader; T*psbA* is a *psbA* terminator.

Figure 3: Schematic illustration of four different examples of a heterologous nucleotide sequence in a pCas (donor vector) expression cassette.

Figure 4: Schematic illustration of pCas expression cassettes with pre-cloned antibiotic resistance and marker genes. MCS located past marker/selection genes is used to express inserted transgenes of interest on one polycistronic mRNA or as translational fusions with marker/selection genes. (A) attB/attP recombination sites flanking, selection marker expression cassette, are aimed for removal of antibiotic resistance genes from the final transgenic crop product. Spectinomycin resistance gene (*aadA*) and phiC31 phage site-specific recombinase attachment (*attP/attB*) are illustrated. phiC31 phage recombinase attachment sequences can be used to flank the selection marker gene for its future removal from commercial crops (not illustrated).

(B) MCS located past reporter gene is used to express inserted transgenes of interest on one polycistronic mRNA or as translational fusions with reporter gene. (C) Schematic illustration example of fluorescent marker gene (GFP) incorporated into pre-cloned pCas expression vector. The gene encoding for Green Fluorescent Protein has been inserted using BglII and SacI restriction enzymes. Additional genes can be cloned downstream of GFP and, upon incorporation of a ribosome binding site, translationally co-expressed as a single operon unit (PI-PspI–*PpsbA*–*LpsbA*–BglII–*GFP*–SacI–MCS–*TpsbA*–PI-PspI (SEQ ID NO: __)). Abbreviations: *aadA* – spectinomycin resistance gene; *neo* – kanamycin resistance gene; GFP – green fluorescent protein; YFP – yellow fluorescent protein; GUS – beta-glucuronidase; attB/attP – selection marker excision recombination sites; HE – homing endonuclease recognition site; MCS – multiple cloning site.

Figure 5: Schematic illustration of an assembly of modular pUniPlast vector using expression cassettes derived from pCas vectors. pCas vector containing single expression cassette; restriction sites for the homing endonuclease PI-PspI indicated by arrows. Fully assembled pUniPlast vector containing three independent expression cassettes also shown in far right bottom corner.

Figure 6: Schematic illustration of pUniPlast MCS detailed. Homologues recombination sequences – such as *trnI* and *trnA* – are flanked by a pair of rare-cutter nuclease sites, allowing sustaining directionality of the cloned insert, as well as easy exchange with sequences from different species, and thus rapid adaptation of pUniPlast for use in any plant species. pUniPlast MCS can be expended by insertion of additional rare-cutter cloning sites, such as those of homing endonucleases (HE) or zinc finger nucleases (ZFN).

Figure 7: Schematic illustration of a chimeric transcriptional activator containing the yeast GAL4 DNA binding domain [mGal4, modified for optimal activity in Arabidopsis] fused to the VP16 transcriptional activator from Herpes simplex virus. This mGal4-VP16 transcription factor expression will be controlled by an inducible promoter, such as a tissue-specific or circadian or otherwise inducible (stress, heavy metal, etc) promoter in the nucleus. When the tissue-specific promoter is activated, the mGal4- VP16 protein will transcribed and be targeted to a plastid (e.g., a chloroplast) using N-terminally fused plastid transit peptide. The LUX genes

in the chloroplast will be driven by an mGal4-VP16-inducible Gal4-UAS promoter (<http://www.plantsci.cam.ac.uk/Haseloff/Home.html>), to which mGal4-VP16 binds and thus activates LUX transcription. Thus, activation of the LUX operon is indirect.

Figure 8: Agarose gel prepared for DNA analysis. Cloning plastid promoters into pSat plasmids (Tzfira et al, 2005) to create pCas vectors.

Figure 9: Vector for nuclear expression of LUX genes. Separate LUX genes are PCR amplified and subcloned into MCS of pSAT vectors (one gene per one vector), under plant-specific regulatory promoter and terminator sequences, such as 35S promoter and terminator. Further, the cassettes containing promoter, lux gene and a terminator are subcloned from the SAT vectors into PZP-RCS binary vector (Tzfira et al, 2005) using homing endonucleases, such as PI-PspI or I-CeuI. The resulting binary PZP-RCS-Lux plasmid is used for plant transformation.

Figure 10: Vector for expression of LUX operon from chloroplast genome. LUX operon as a whole is PCR amplified and subcloned, under chloroplast specific promoter and terminator, into vector suitable for transformation of chloroplasts, such as pLDCtV (ref: De Cosa, B., Moar, W., Lee, S. B., Miller, M., and Daniell, H. (2001) "Overexpression of the Bt cry2Aa2 operon in chloroplasts leads to formation of insecticidal crystals". Nature Biotechnol., vol. 19, pp. 71–74). Vector elements: *aadA*: selection marker conferring spectinomycin resistance; *trnI/trnA*: targeting homologues recombination elements; pUniPlast: Universal Plastid transformation vector.

Figure 11: Generation of transgenic tobacco plants. Transgenic tobacco plants are resistant to kanamycin (through incorporation of *nptII* gene) and are expressing free DsRed and GFP. Photographs made using fluorescent confocal microscope.

Figure 12: Remote activation of an inducible promoter in-planta. (A) Schematic representation of beta-glucuronidase (GUS) activation by plasmid-borne mGal4-VP16. (B). mGal4-VP16 expressed from transient plant expression vector pRTL2 was biolistically delivered into leaves of an *Arabidopsis* plant carrying UAS driven, genome-integrated GUS reporter gene. Transient expression of mGal4-VP16 resulted in activation of GUS expression and was detected by histochemical staining (C) Specificity of UAS promoter activation has been confirmed when similar biolistic

delivery has been made into a wild-type Arabidopsis plants, carrying no UAS driven GUS gene. Krichevsky et al unpublished and Developmental Biology (2007).

Figure 13: pCas-3 donor vector prepared by BioGlow Inc. pCas-3 enzymatic digest displays presence of Prm (AgeI/NcoI fragment), Trps16 (NotI/XbaI fragment) and the whole Prm/MCS/Trps16 cassette (I-SceI fragment), showing that a complete pCas-3 has been constructed (BioGlow Inc, unpublished).

Figure 14: Alignment of LUX operons from three major luminescent bacteria families, namely *Photobacterium*, *Vibrio* and *Xenorhabdus* (aka *Photorhabdus*). Sequence comparison produced using ClustalW software (EBI) of three luminescent bacteria species: *Photobacterium leiognathi* (abbreviated as Pl), *Vibrio harveyi* (abbreviated as Vh) and *Xenorhabdus* (aka *Photorhabdus*) *luminescens* (abbreviated as XI).

DETAILED DESCRIPTION OF THE INVENTION

In the present invention, bacterial luminescence machinery is exploited for the production of bioluminescent reactions in plants. The metabolically engineered plants preferably produce high levels of luminescence which is detectable by a human eye.

5 The invention is useful: (1) as a commodity product for ornamental flowers and horticultural markets, (2) as an environmental pollution control tool and (3) for basic research, e.g., as an improved reporter system.

The following abbreviations are used throughout the specification. PCR – polymerase chain reaction; RB – right border; LB – left border (note: the DNA
10 sequence located between right and left borders of the binary plasmid is integrated into host plant nuclear genome); MCS – multiple cloning site; pSAT vectors – series of vectors used for subcloning of cassettes, carrying genetic sequences, into binary PZP-RCS plasmid; PZP-RCS plasmid – binary vector used for plant transformation; pLDCtV – chloroplast transformation vector (Ref: De Cosa et al, 2001.); pUniPlast –
15 a plastid transformation system developed by BioGlow Inc; HE – homing endonuclease; ZNF – zinc finger nuclease.

Biochemistry of bacterial luminescence:

The bacterial luminescence reaction involves bacterial luciferase; two substrates, i.e., reduced flavin mononucleotide (FMNH₂) and a long-chain aldehyde;
20 and oxygen. See Figure 1.



Flavin mononucleotide (FMNH₂) cannot be substituted by another type of nucleotide (such as NADH). However, the aldehyde specificity is much less restricted. It has been shown that a wide range of aldehydes with chain length longer
25 than seven carbon atoms, including palmitaldehyde (derived from kidney cortex), tetradecanal (the original bacterial substrate) or decanal, can produce strong luminescence. (Meighen et al. *Adv. Microb. Physiol.*, 34:1-67 (1993); Meighen E.A *FASEB*, 7: 1016-22 (1993).) This was demonstrated in so-called “Standard Assay,”

where FMNH₂ is injected into solution containing an aldehyde and bacterial luciferase, resulting in rapid light emission.

Related Plant Biochemistry

Precursors of fatty acyl acids, such as acyl-ACP, are produced in both prokaryotes and eukaryotes. However, Vitamin B2 is naturally produced only in bacteria and plants, but not animals, which receive it from their diet. Intriguingly, biochemical pathways of Vitamin B2 and fatty acids synthesis are very strongly conserved amongst plants and eubacteria. (Fischer et al., *JBC*, 279:36299–36308 (2004); Slabas et al., *Biochemical Society Transactions*, 30, part 6 (2002).) Therefore, similarly to bacteria, plants possess all the basic components for bioluminescent reaction based on bacterial luciferase system. Also, transgenic plants expressing bacterial luciferase subunits A and B have been shown to produce low levels of luminescence, detectable with low-light imaging equipment (X-ray film or photomultiplier), when exposed to a substrate supplied from an external source. (Greer III LF and Szalay *Luminesc.*, 17:43-74 (2002)).

Transgenic Bioluminescent Plant

In one aspect, the invention relates to a transgenic bioluminescent plant. The plant includes an expressible heterologous nucleotide sequence, which includes a bacterial LUX operon. The LUX operon includes LUX A, LUX B, LUX C, LUX D, LUX E, and LUX G genes (“the six LUX genes”). The heterologous nucleotide sequence is expressed to render the plant cell bioluminescent.

The terms “transgenic,” “transformed,” “transfected” as used herein includes any cell, cell line, callus, tissue, plant tissue, or plant into which a nucleic acid heterologous to the host cell has been introduced. The term “transgenic” as used herein does not encompass an alteration of the genome (chromosomal or extra-chromosomal) by conventional plant breeding methods or by naturally occurring events, such as random cross-fertilization, non-recombinant viral infection, non-recombinant bacterial transformation, non-recombinant transposition, or spontaneous mutation. The term “transgenic plant” refers to a plant or plant tissue that contains an inheritable heterologous nucleotide sequence.

The term “bioluminescent” as used herein refers to luminescence that occurs in a plant or plant tissue, in which energy from a chemical reaction is transformed into light energy. The term “bioluminescent” further refers to the production of light in a recombinant plant or plant tissue engineered to contain chemical compounds
5 necessary for bioluminescence in the plant or plant tissue. Preferably, the transgenic plant is “stably” bioluminescent, which refers to the introduction and integration of a heterologous nucleotide sequence for bioluminescence into the genome of a transfected cell.

10 An example of a bioluminescent chemical reaction is a reaction in which a chemical substrate, such as luciferin, reacts with oxygen in the presence of an enzyme, such as luciferase. See the above description of bacterial luminescence.

The term “plant” is used broadly herein to refer to a eukaryotic organism containing plastids, particularly chloroplasts, and includes any such organism at any stage of development. The term “plant” as used herein refers to a whole plant or a
15 part of a plant (*e.g.*, a plant cutting, a plant cell, a plant cell culture, a plant organ, a plant seed, and a plantlet), a seed, a cell- or a tissue- culture derived from a plant, plant organ (*e.g.*, embryos, pollen, ovules, seeds, leaves, flowers, branches, fruit, kernels, ears, cobs, husks, stalks, roots, root tips, anthers, *etc.*).

Any plant may be used for the invention. For example, *Nicotiniana*
20 *benthamiana*, *Arabidopsis thaliana*, or *Nicotiana tabacum* (tobacco) can be used, as they are frequently employed as model organisms in plant research and a large amount of data regarding their biology has been accumulated. Also, a good model organism for the bioluminescent plants of the present invention are plants which intrinsically express large amounts of Vitamin B2, such as, for example, asparagus or
25 chard. To minimize costs, the luminescent system from several highly luminescent bacterial species, such as *Photobacterium leiognathi* or *Vibrio harveyi*, as well as *Shewanella hanedai*, is transferred into several plant species. Since this approach requires application of essentially same technology in parallel on several gene transfers, this strategy does not significantly increase costs. Other preferred plants
30 include ornamental plants, useful or ornamental trees, flowers, cut flowers, shrubs, or turf.

The transgenic bioluminescent plant, as used herein, includes at least one plant cell. A "plant cell" refers to any cell of a plant, either taken directly from a seed or plant, or derived through culture from a cell taken from a plant. A plant cell includes, for example, cells from undifferentiated tissue (*e.g.*, callus), plant seeds, propagules, gametophytes, sporophytes, pollen, microspores, and embryos. A plant cell typically contains a "plastid," which refers to an organelle with its own genetic machinery in a plant cell. Examples of a plastid include chloroplasts, chromoplasts, etioplasts, gerontoplasts, leucoplasts, proplastids, etc.

The transgenic bioluminescent plant further includes an expressible heterologous nucleotide sequence. The term "expressible," "expressed," and variations thereof refer to the ability of a cell to transcribe a nucleotide sequence to mRNA and translate the mRNA to synthesize a peptide that provides a biological or biochemical function. Preferably, the cell is a plant cell.

As used herein, "heterologous" refers to that which is foreign or non-native to a particular host or genome. Accordingly, a "heterologous nucleotide sequence" or "transgene" refers to a nucleotide sequence that originates from a species foreign to the host organism, or if the nucleotide sequence originates from the same species as the host, the nucleotide sequence is substantially modified from its native form in composition and/or genomic locus by deliberate genetic manipulation. The term "nucleotide sequence" refers to a sequence of two or more nucleotides, such as RNA or DNA. A "heterologous protein" refers to a protein that is foreign or non-native to a host cell and is typically encoded by a heterologous nucleotide sequence.

The LUX operon

The luminescence in bacteria is based on a functional LUX operon, highly conserved amongst different bacterial families. The bacteria do not constantly emit light and the operon is induced by different physiological conditions, such as quorum sensing in light organs of marine animals. LUX contains 6 basic luminescence genes in the following order: C-D-A-B-E-G.

The LUX A and B genes encode luciferase subunits. The LUX C, D and E genes encode fatty-reductase complex which produces aldehyde for the reaction. The LUX G gene encodes an exchange factor, facilitating FMNH₂ turnover at the enzyme

active site. (Lin et al., *BBRC*, 246(2):446-52 (1998)). LUX G is the “molecular switch” for FMNH₂ turnover. Preferably, the heterologous nucleotide sequence does not include a gene encoding *Vibrio harveyi* FRP.

The LUX CDE enzymatic complex diverts a range of fatty-acyl acids, such as acyl-CoA, acyl-ACP and others from the basic fatty acids biosynthesis cycle, converts them to the aldehyde substrate and channels them to the luminescence reaction. The other substrate, the FMNH₂, is naturally produced in bacteria. One of the pathways for FMNH₂ production in the luminescent bacteria is encoded by the RIB operon, immediately adjacent to the LUX operon. (Lee et al., *Journal of Bacteriology* 176(7):2100-2104 (1994); Jones et al. *Biochem. J.*, 347:205-209 (2000).)

In one embodiment, the heterologous nucleotide sequence includes a bacterial LUX operon. Use of the complete bacterial LUX operon allows for intrinsic bioluminescence, which refers to the ability of a cell to contain all of the required elements for production of light in the transgenic cell, without the requirement for exogenous addition of chemical compounds. Preferably, the cell is a cell from a plant.

The term “operon” refers to a nucleotide sequence which codes for a group of genes transcribed together. The term “gene” refers to chromosomal DNA, plasmid DNA, cDNA, synthetic DNA, or other DNA that encodes a peptide, polypeptide, protein, or RNA molecule, and regions flanking the coding sequence involved in the regulation of expression. Some genes can be transcribed into mRNA and translated into polypeptides (structural genes); other genes can be transcribed into RNA (e.g., rRNA, tRNA); and other types of genes function as regulators of expression (regulator genes).

The term “LUX operon” as used herein refers to an operon that includes at least six genes for intrinsic bioluminescence. The six genes include LUX A, LUX B, LUX C, LUX D, LUX E, and LUX G genes.

For purposes of the invention, the genes corresponding to the LUX operon, and any other gene required for proper functioning of LUX luciferase in a plant, are isolated from the genome of luminescent bacteria. For example, the LUX operon and LUX A, LUX B, LUX C, LUX D, LUX E, and LUX G genes may be derived from

any bioluminescent bacteria that express the LUX operon or six LUX genes to generate bioluminescence.

Examples of a nucleotide sequence encoding the full LUX operon is presented in GenBank under accession numbers AY341062 (*Vibrio fischeri* [Vibrio fischeri strain ATCC 7744 lux operon, complete sequence])(SEQ ID NO: 7); EU192082 (*Vibrio harveyi* [Vibrio harveyi BCB440 lux operon, complete sequence]) (SEQ ID NO: 8); AF403784 (*Photorhabdus luminescens*, (formally referred as *Xenorhabdus luminescens* [*Photorhabdus luminescens* lux operon, complete sequence]) (SEQ ID NO: 9); and AB261992 (*Shewanella hanedai* [*Shewanella hanedai* lux operon (luxC, luxD, luxA, luxB, luxE, luxG) genes and flanking regions, strain: NCIMB 2157]) (SEQ ID NO: 10); and M63594 (*Photobacterium leiognathi* [*Photobacterium leiognathi* lux operon (luxC, luxD, luxA, luxB, luxE, luxG) genes, complete cds]) (SEQ ID NO: 11); and DQ988873 (*Photobacterium phosphoreum* [*Photobacterium phosphoreum* strain ATCC 11040, complete LUX and RIB operons]) (SEQ ID NO: 12).

Examples of a nucleotide sequence encoding LUX A, LUX B, LUX C, LUX D, LUX E, and LUX G genes are included in the nucleotide sequences encoding the full LUX operon, listed above. For example, the following LUX genes were derived from GenBank accession number M63594 (*Photobacterium leiognathi* [*Photobacterium leiognathi* lux operon (luxC, luxD, luxA, luxB, luxE, luxG) genes, complete cds]) (SEQ ID NO: 11): LUX A (SEQ ID NO: 13), LUX B (SEQ ID NO: 14), LUX C (SEQ ID NO: 15), LUX D (SEQ ID NO: 16), LUX E (SEQ ID NO: 17), and LUX G (SEQ ID NO: 18) genes.

Further examples of a LUX E gene is presented in GenBank accession number M62812 for *Vibrio fischeri* [Vibrio fischeri LuxE (luxE) gene, partial cds; and LuxG (luxG) gene, complete cds]) (SEQ ID NO: 19). Further examples of a LUX G gene is presented in the sequences of SEQ ID NO: 53 (*Photobacterium leiognathi* (derived from GenBank # M63594); SEQ ID NO: 54 (*Photobacterium phosphoreum* (derived from DQ988873); SEQ ID NO: 55 (*Vibrio harveyi* (derived from EU192082); SEQ ID NO: 56 (*Vibrio fischeri* (derived from M62812); and SEQ ID NO: 57 (*Shewanella hanedai* (derived from AB261992).

The polymerase chain reaction can be used to amplify the isolated LUX genes. Such genes may be cloned into a vector, such as a donor or acceptor vector described below.

5 The nucleotide sequence of the LUX operon and LUX A, LUX B, LUX C, LUX D, LUX E, and LUX G genes may be derived from wild-type organisms. Wild-type refers to the normal gene or organism found in nature without any known mutation. Other nucleotide sequences within the invention include a nucleotide sequence that encodes variants of LUX A, LUX B, LUX C, LUX D, LUX E, and
10 LUX G proteins, and a nucleotide sequence that encodes mutant forms, recombinant forms, or non-naturally occurring variant forms of these proteins.

In some additional preferred embodiments, the heterologous nucleotide sequence (the transfected gene) includes additional genes related to production of substrates, such as, for example, LUX H, LUX F, LUX Y, and/or LUX L. In another
15 embodiment, the heterologous nucleotide sequence further includes a LUX H, LUX F, LUX Y, and/or a LUX L gene, or combinations thereof. In some species, an additional gene not conserved amongst other luminescent bacteria, such as LUX F or LUX H, can be present but is not directly related the luminescence. These genes mainly encode for proteins with support function, such as luminescence wave length
20 shift. (Karatani et al., *Photochemistry and Photobiology* 71(2): 230–236 (2000)).

LUX Y encodes a Yellow Fluorescent Protein (YFP). LUX Y can be isolated from, for example, *Vibrio fischeri* strain Y-1. LUX L gene codes for a lumazine protein with an amino acid identity of riboflavin synthase, and is located upstream of LUX C gene in some *Photobacterium* species. An example of a nucleotide sequence
25 encoding *Photobacterium leiognathi* LUX F gene is presented at EBI # CAA30833 (SEQ ID NO: 20). An example of nucleotide sequence encoding *Aliivibrio fischeri* yellow fluorescent protein LUX Y is presented at EBI # AAA27544 (SEQ ID NO: 21). An example of a nucleotide sequence encoding *Vibrio sp.* BCB494 LUX H is presented at UniProtKB/TrEMBL entry B0F6S8, EBI # ABY56819 (SEQ ID NO:
30 22).

In another embodiment, the heterologous nucleotide sequence includes a plastid targeting sequence. A "plastid targeting sequence" as used herein refers to a nucleotide sequence that encodes a polypeptide sequence, which can direct a second polypeptide to a plastid of the plant cell. Preferably, the plastid targeting sequence is a chloroplast targeting sequence.

It is known in the art that non-chloroplast proteins may be targeted *in vivo* and *in vitro* to the chloroplast by use of protein fusions with a peptide encoded by a chloroplast targeting sequence. For example, luciferase genes of a heterologous nucleotide sequence may be fused with a plastid targeting sequence. When the luciferase gene is expressed, the targeting sequence is included in the translated polypeptide. The targeting sequence then directs the polypeptide into a plastid, such as a chloroplast.

Typically, the chloroplast targeting sequence encodes a polypeptide extension (called a chloroplast transit peptide (CTP)). The polypeptide extension is linked to the N-terminus of the heterologous peptide encoded by the heterologous nucleotide sequence.

Those skilled in the art will appreciate that there are a variety of known targeting sequences that can be added to the expressed polypeptide. (See, for example, Gnanasambandam A, Polkinghorne IG, and Birch RG. "Heterologous signals allow efficient targeting of a nuclear-encoded fusion protein to plastids and endoplasmic reticulum in diverse plant species." *Plant Biotechnol J.*, 2007 5(2):290-6; and Rensink WA, Pilon M, and Weisbeek P. "Domains of a transit sequence required for *in vivo* import in *Arabidopsis* chloroplasts." *Plant Physiol.* 1998 118(2):691-9).

Examples of a chloroplast targeting sequence include a sequence that encodes the *Arabidopsis thaliana* EPSPS chloroplast transit peptide, the *Petunia hybrida* EPSPS chloroplast transit peptide, and the rice *rbcS* gene chloroplast targeting sequence.

Further examples of a chloroplast target peptide include the small subunit (SSU) of ribulose-1,5,-biphosphate carboxylase, and the light harvesting complex protein I and protein II. Incorporation of a suitable chloroplast target (transit) peptide

has been shown to target heterologous protein sequences to chloroplasts in transgenic plants. Those skilled in the art will recognize that various chimeric constructs can be made, if needed, that utilize the functionality of a particular CTP to import a given gene product into a chloroplast.

5 Other CTPs that may be useful in practicing the present invention include PsRbcS-derived CTPs (*Pisum sativum* Rubisco small subunit CTP); AtRbcS CTP (*Arabidopsis thaliana* Rubisco small subunit 1A CTP; CTP1); AtShkG CTP (CTP2); AtShkGZm CTP (CTP2synthetic; codon optimized for monocot expression); PhShkG CTP (*Petunia hybrida* EPSPS; CTP4; codon optimized for monocot expression);
 10 TaWaxy CTP (*Triticum aestivum* granule-bound starch synthase CTPsynthetic, codon optimized for corn expression); OsWaxy CTP (*Oryza sativa* starch synthase CTP); NtRbcS CTP (*Nicotiana tabacum* ribulose 1,5-bisphosphate carboxylase small subunit chloroplast transit peptide); ZmAS CTP (*Zea mays* anthranilate synthase alpha 2 subunit gene CTP); and RgAS CTP (*Ruta graveolens* anthranilate synthase
 15 CTP). Other transit peptides that may be useful include maize cab-m7 signal sequence and the pea (*Pisum sativum*) glutathione reductase signal sequence.

Additional examples of such targeting sequences may include: spinach lumazine synthase (SEQ ID NO: 23), *Chlamydomonas* ferredoxin (SEQ ID NO: 24), and Rubisco activase (SEQ ID NO: 25) transit peptides, and others.

20 The chloroplast targeting sequence may be used to target any peptide encoded by a heterologous nucleotide sequence to the chloroplast or other plastid. In one embodiment, the chloroplast targeting sequence is linked to a 5'- or a 3'- end of the LUX A, LUX B, LUX C, LUX D, LUX E, or LUX G genes. In another embodiment, the chloroplast targeting sequence is linked to a 5'- or a 3'- end of a gene encoding a
 25 fluorescent protein.

In one embodiment, the heterologous nucleotide sequence can be placed in a single vector. For example, the heterologous nucleotide sequence can include the six LUX genes in a single vector. See Figure 10. In another embodiment, a heterologous nucleotide sequence encoding one of the six LUX genes can be placed in a different
 30 vector for each LUX gene, resulting in multiple different vectors. See Figure 9. The

heterologous nucleotide sequence can additionally include at least one gene encoding a cofactor for enhancing bioluminescence.

The term “vector” as used herein refers to a vehicle used for introduction of a nucleotide sequence into a host. A vector may be a plasmid, cosmid, phage, transposon, virus, or any other suitable vehicle. Preferably, the vector is a plasmid. A vector may include regulatory sequences useful for expression of a gene product in a host, including but not limited to a promoter, ribosomal binding site, and termination sequences. In one preferred embodiment, the vector is a vector for transforming a plastid as described below in another aspect of the invention.

Numerous vectors are suitable for stable transformation of a plant cell or a plastid. Accordingly, the LUX genes may be delivered into nuclear or chloroplast genomes.

In one embodiment, the vector is a binary vector. A “binary vector” refers to a vector that includes a modified T-region from Ti plasmid, which allows replication in *E. coli* and in *Agrobacterium* cells, and usually includes selection marker genes. Preferably, the vector is a binary pPZP-RCS vector (Goderis et al, 2002), assembled employing expression cassettes derived from the pSAT vectors (Tzfira et al., (2005) “pSAT vectors: a modular series of plasmids for autofluorescent protein tagging and expression of multiple genes in plants.” *Plant Mol. Biol.*, 57(4):503-16 (2005)).

The pSAT vectors contain a plant promoter, an MCS and a plant terminator, which allows for subcloning and expression of one transgene. Cassettes, containing promoter/gene of interest/terminator sequence are derived from pSAT vectors using homing endonucleases and subcloned into the same sites of the pPZP-RCS vector. The pPZP-RCS is a binary vector that includes homing endonuclease enzyme recognition sites in its MCS and allows for cloning of multiple (from 6 or more) pSATs derived cassettes into it, thus serving as a single binary (acceptor) vector. This vector system allows for multiple nuclear transgene expression without requiring bicistronic RNAs or internal ribosome binding sites (IRES). Accordingly, use of pSAT vectors allows introduction of multiple genes into a single acceptor vector. The single pPZP-RCS acceptor vector containing the multiple genes may then be

introduced in a single transformation event into a plant, without requiring three or more subsequent plant transformations.

The specific pSATs and GeneBank accession numbers are: pSAT1-EGFP-C1 (SEQ ID NO: 26), pSAT2-EGFP-C1 (SEQ ID NO: 27), pSAT3-EGFP-C1 (SEQ ID NO: 28), pSAT4-EGFP-C1 (SEQ ID NO: 29), pSAT5-EGFP-C1 (SEQ ID NO: 30), pSAT6-EGFP-C1 (SEQ ID NO: 31) and pSAT7-EGFP-C1 (SEQ ID NO: 32), respective NCBI numbers are: AY818363 (SEQ ID NO: 26), AY818365 (SEQ ID NO: 27), AY818366 (SEQ ID NO: 28), AY818367 (SEQ ID NO: 29), AY818368 (SEQ ID NO: 30), AY818377 (SEQ ID NO: 31) and AY818384 (SEQ ID NO: 32).

In another embodiment, the vector is a plastid (chloroplast) transformation vector. Typically, a transgene in a chloroplast transformation vector is flanked by a “homologous recombination site,” which is a DNA region that is homologous to a region of the plastome. The “plastome” refers to the genome of a plastid. The homologous recombination site enables site-specific integration of a transgene construct into the plastome by the process of homologous recombination. Homologous recombination is a process that naturally occurs in plastids. Homologous recombination differs from random transgene integration into plant nuclear genome. An example of a chloroplast transformation vector is pLDCtV (See De Cosa, B., Moar, W., Lee, S. B., Miller, M., and Daniell, H. (2001) “Overexpression of the Bt cry2Aa2 operon in chloroplasts leads to formation of insecticidal crystals”. *Nature Biotechnol.*, vol. 19, pp. 71–74).

In a preferred embodiment of the invention, the full or partial LUX operon is directly expressed from the chloroplast genome. Insertion of the genes into chloroplast genome is done by cloning the whole LUX operon into a chloroplast transformation vector. Such a method of cloning may include transforming chloroplasts with the vector, and bringing the population of chloroplast genomes copies to homogeneity using standard methods. (See De Cosa, B., Moar, W., Lee, S. B., Miller, M., and Daniell, H. (2001) “Overexpression of the Bt cry2Aa2 operon in chloroplasts leads to formation of insecticidal crystals”. *Nature Biotechnol.*, vol. 19, pp. 71–74.)

In another preferred embodiment, the vector is derived from the pUniPlast vector system, which is a plastid transformation vector system developed by BioGlow Inc.

The heterologous nucleotide sequence or vector described herein may include regulatory sequences useful for expression of a gene product in a host, such a promoter. The term “promoter” refers to a nucleotide sequence capable of controlling the expression of a coding sequence. A promoter drives expression of an operably linked nucleotide sequence. The term “operably linked” as used herein refers to linkage of a promoter to a nucleotide sequence such that the promoter mediates transcription of the nucleotide sequence. A “coding sequence” refers to a nucleotide sequence that encodes a specific amino acid sequence. A promoter is typically located upstream (5') to a coding sequence.

A wide variety of promoters is known in the art and may be used to facilitate expression of a gene in the heterologous nucleotide sequence. Examples of suitable promoters include constitutive promoters, plant tissue-specific promoters, plant development-specific promoters, inducible promoters, circadian cycle promoters, viral promoters, male germline-specific promoters, female germline-specific promoters, flower-specific promoters, and vegetative shoot apical meristem-specific promoters.

A “constitutive” promoter refers to a promoter that causes a gene to be expressed in all cell types at all times. An example of a constitutive plastid promoter is *psbA*, photosystem II reaction center promoter (derived from pCLT146, GeneBank # DQ463359; and *rrn*, chloroplast 16S rRNA gene promoter (derived from pN-IC101, GeneBank # AY442171).

Examples of nuclear genomic constitutive plant promoters include the cauliflower mosaic virus (CaMV) 35S promoter, which confers constitutive, high-level expression in most plant cells; the nopaline synthase promoter; the octopine synthase promoter; cauliflower mosaic virus 19S promoter; rice actin 1 promoter; manopine synthase promoter; and a histone promoter. Further suitable constitutive promoters include the Rubisco small subunit (SSU) promoter, leguminB promoter, TR dual promoter, ubiquitin promoter, and Super promoter. Different heterologous nucleotide sequences or vectors may contain different promoters to prevent gene

silencing when several consecutive genes on a chromosome are expressed from the same promoter.

An "inducible" promoter refers to a promoter that is regulated in response to a stress or stimuli. Examples of inducible promoters include a tetracycline repressor system, Lac repressor system, copper-inducible system, salicylate-inducible system (such as the PR1a system), and alcohol-inducible system. Further examples include inducible promoters that are regulated in response to environmental, hormonal, chemical, and/or developmental stress or stimuli. Such stress or stimuli include heat (*e.g.*, tomato hsp70 promoter or hsp80 promoter); light; hormones (*e.g.*, steroid-inducible MMTV LTR promoter), such as abscisic acid; chemicals, such as methyl jasmonate, salicylic acid; increased salinity; drought; pathogen (*e.g.*, promoter of the PRP1 gene); heavy metals (*e.g.*, heavy metal-inducible metallothionein I promoter and the promoter controlling expression of the tobacco gene *cdiGRP*; and wounds (*e.g.*, *pinII* promoter). Preferably, the promoter is a promoter induced by heavy metals.

A "tissue-specific" promoter as used herein refers to a promoter that drives expression of an operably linked nucleotide sequence to a particular tissue. A tissue-specific promoter drives expression of a gene in one or more cell types in a specific organ (such as leaves, or seeds), specific tissues (such as embryo or cotyledon), or specific cell types (such as seed storage cells or leaf parenchyma). Examples include *Gentiana triflora* promoter for chalcone synthase (NCBI accession AB005484), a seed-specific promoter, such as β -conglycinin, napin promoter, and phaseolin; mature leaves-specific promoter, such as the SAG promoter from *Arabidopsis*.

Promoters responsible to the circadian cycle can also be used in the heterologous nucleotide sequence or vector. Such promoters include the native ELF3 promoter and the promoter from the chlorophyll a/b binding protein (*CAB2* promoter).

The heterologous nucleotide sequence or vector may also include leader sequences, such as; *rbcL*, ribulose-bisphosphate carboxylase gene leader sequence (derived from pCLT516, GeneBank # DQ882177: (SEQ ID NO: 33); and Shine-

Dalgarno consensus ribosome binding sequence (AGGAGG); and terminators, such as *psbA*, which is a photosystem II reaction center terminator (derived from pCLT146, GeneBank # DQ463359: (SEQ ID NO: 34); and *rps16* gene *rps16* terminator (derived from pL3 vector series, GeneBank # EU520589, EU520588, EU520587: (SEQ ID NO: 35).

In addition, the heterologous nucleotide sequence or vector may include a nucleotide sequence for a selectable and/or screenable marker. A “selection marker” refers to a protein necessary for survival or growth of a transformed plant cell grown in a selective culture regimen. Typical selection markers include sequences that encode proteins, which confer resistance to selective agents, such as antibiotics, herbicides, or other toxins. Examples of selection markers include genes for conferring resistance to antibiotics, such as spectinomycin, streptomycin, tetracycline, ampicillin, kanamycin, G 418, neomycin, bleomycin, hygromycin, methotrexate, dicamba, glufosinate, or glyphosate.

Various other selection markers confer a growth-related advantage to the transformed cells over the non-transformed cells. Examples include selection markers for β -glucuronidase (in conjunction with, for example, cytokinin glucuronide), mannose-6-phosphate isomerase (in conjunction with mannose), and UDP-galactose 4-epimerase (in conjunction with, for example, galactose).

Preferably, the selection marker is functional in plastids. Such selection markers include those which confer resistance to spectinomycin (*e.g.*, encoded by the resistance gene, *aadA*), streptomycin, kanamycin, lincomycin, gentamycin, hygromycin, methotrexate, bleomycin, phleomycin, blasticidin, sulfonamide, phosphinothricin, chlorsulfuron, bromoxynil, glyphosate, 2,4-D, atrazine, 4-methyltryptophan, nitrate, S-aminoethyl-L-cysteine, lysine/threonine, aminoethyl-cysteine or betaine aldehyde. Especially preferred are the genes *aadA* (GeneBank NC_009838), *nptII* (GeneBank FM177583), *BADH* (GeneBank AY050316), *aphA-6* (GeneBank X07753).

After a heterologous nucleotide sequence has been introduced into a host cell, it may be advantageous to remove or delete certain sequences from the plastome or genome of the plant or cell. For example, it may be advantageous to remove a

selection marker gene that has been introduced into a genome if the selection marker is no longer necessarily required after the selection phase. Methods for directed deletion of sequences are known in the art. For example, the nucleotide sequence encoding a selection marker preferably includes a homology-based excision element, such as *Cre-lox* and *attB/attP* recognition sequences, which allow removal of the selection marker genes using phage site-specific recombinases.

In one embodiment, the heterologous nucleotide sequence or vector includes reporter genes. Reporter genes encode readily quantifiable proteins which, via their color or enzyme activity, allow an assessment of the transformation efficiency, the site or time of expression or the identification of transgenic plants. Examples of reporter genes include green fluorescent protein (GFP), luciferase, β -Galactosidase, β -Glucuronidase (GUS), R-Locus gene product, β -Lactamase, xylE gene product, Alpha-amylase, and Tyrosinase.

The heterologous nucleotide sequence or vector may include sequences encoding a fluorescent protein that are excited or fluoresce at different wavelengths, at different periods of time, or under different conditions. Such a fluorescent protein is DsRed (GeneBank # EU827527, DsRed-Monomer gene, synthetic construct)(SEQ ID NO: 129), which can fluoresce and emit light at red wavelengths.

The heterologous nucleotide sequence or vector may also include functional elements, which influence the generation, multiplication, function, use or value of the heterologous nucleotide sequence or vector used within the scope of the present invention. Examples of functional elements include replication origins (ORI), which make possible an amplification of the heterologous nucleotide sequence or vector according to the invention in, for example, *E. coli* or in plastids; multiple cloning sites (MCSs), which permit and facilitate the insertion of one or more nucleic acid sequences; homologous recombination sites, allowing stable recombination of transgenes into plastid genome; and border sequences, which make possible *Agrobacterium* -mediated transfer of the heterologous nucleotide sequence or vector into plant cells for the transfer and integration into the plant genome, such as, for example, the right or left border of the T-DNA or the vir region.

The heterologous nucleotide sequence or vector may optionally include RNA processing signals, *e.g.*, introns, which may be positioned upstream or downstream or within a polypeptide-encoding sequence in the heterologous nucleotide sequence. Intron sequences are known in the art to aid in the expression of heterologous nucleotide sequences in plant cells.

The heterologous nucleotide sequence or vector described herein preferably includes a rare cutting enzyme sequence. A “rare cutting enzyme” refers to an enzyme that cleaves a nucleotide sequence at a recognition site of eight or more nucleotide base pairs characteristic to the particular enzyme. The frequency of cutting for a particular rare cutting enzyme is determined in part by the length of the recognition sequence, and the distribution of bases and motifs in the target DNA. Examples of rare cutting enzymes include endonucleases, homing endonucleases, and zinc finger nucleases.

Homing endonucleases are double stranded DNases with large and extremely rare recognition sites that are longer than eight base pairs long, usually 12–40 base pairs long. Homing endonuclease I-SceI has an 18 base pair long recognition sequence, which statistically may occur only once in every 7×10^{10} base pairs, which is equivalent in length to 20 mammalian-sized genomes. On the other hand, a “regular” restriction enzyme, such as EcoRI having a six base pair recognition site, appears on average every 4×10^3 bp, and thus on average 10^6 times in a single mammalian genome. Thus, on average, homing endonucleases are approximately 20 million times less likely to cut a given DNA sequence than a “regular” DNA restriction enzyme. This property of rare-cutting enzymes allows for their use in engineering of long DNA sequences and assembly of transformation vectors from pre-arranged expression cassettes.

Examples of homing endonucleases include F-SceI, F-SceII, F-SuvI, F-TevI, F-TevII, I-Amal, I-AniI, I-CeuI, I-CeuAIP, I-ChuI, I-Cmoel, I-Cpal, I-CpaII, I-Crel, I-CrepsbIP, I-CrepsbIIP, I-CrepsbIIIP, I-CrepsbIVP, I-CsmI, I-CvuI, I-CvuAIP, I-DdiII, I-DirI, I-DmoI, I-HspNIP, I-LlaI, I-MsoI, I-NaaI, I-NanI, I-NclIP, I-NgrIP, I-NitI, I-NjaI, I-Nsp236IP, I-PakI, I-PboIP, I-PcuIP, I-PcuAI, I-PcuVI, I-PgrIP, I-PobIP, I-PorI, I-PorIIP, I-PpbIP, I-PpoI, I-SPBetaIP, I-ScaI, I-SceI, I-SceII, I-SceIII, I-SceIV, I-SceV, I-SceVI, I-SceVII, I-SexIP, I-SneIP, I-SpomCP, I-SpomIP, I-

SpomIIP, I-SquIP, I-Ssp6803I, I-SthPhiJP, I-SthPhiST3P, I-SthPhiS3bP, I-TdeIP, I-TevI, I-TevII, I-TevIII, I-UarAP, I-UarHGPAIP, I-UarHGPA13P, I-VinIP, I-ZbiIP, PI-MtuI, PI-MtuHIP, PI-MtuHIIP, PI-PfuI, PI-PfuII, PI-PkoI, PI-PkoII, PI-PspI, PI-Rma438121P, PI-SPBetaIP, PI-SceI, PI-TfuI, PI-TfuII, PI-Thyl, PI-TliI, PI-TliII.

5 A “zinc finger nuclease” or “ZFN” refers to a chimeric protein molecule that includes at least one zinc finger DNA binding domain operably linked to at least one nuclease capable of cleaving DNA. A zinc finger nuclease is capable of directing targeted genetic recombination or targeted mutation in a host cell by causing a break at a target site. A zinc finger typically includes a DNA-binding domain and a DNA-
10 cleavage domain, wherein the DNA binding domain includes at least one zinc finger and is operably linked to a DNA-cleavage domain. A preferred example of a zinc finger nuclease is ZFN QQR (TTCTTCCCCGAATTCGGGGAAGAA)(SEQ ID NO: 36)

15 A “rare cutting enzyme sequence” refers to a nucleotide sequence in a target DNA that is a recognition site for a rare cutting enzyme to cut the target DNA at the recognition site. The rare cutting enzyme sequences are typically 10 or more base pairs in length and occur rarely, if at all, in a host cell.

Examples of preferred rare cutting enzyme sequences include the following:

PI-PspI (SEQ ID NO: 37);

20 I-CeuI (SEQ ID NO: 38);

I-SceI (SEQ ID NO: 39);

I-Ppol (SEQ ID NO: 40); and

AscI; FseI; and ZFN QQR.

25 Any combination of rare cutting enzyme sequences can be included in the heterologous nucleotide sequence or vector of the invention. In the embodiments described herein, rare cutting enzymes (*e.g.*, endonucleases, homing endonucleases and gateway technology) may be used to clone LUX genes into any vector, or plant cellular vectors, or chloroplast expression vectors.

Preferably, a rare cutting enzyme sequence flanks a gene in the heterologous nucleotide sequence or vector. The rare cutting enzyme sequence may flank either the 5'-end, the 3'-end, or both the 5'- and 3'- end of a given sequence or gene.

More preferably, a unique rare cutting enzyme sequence flanks each gene. For example, in one embodiment, a unique rare cutting enzyme sequence flanks a reporter gene, another unique rare cutting enzyme sequence flanks a selection marker gene, yet another unique rare cutting enzyme sequence flanks a multiple cloning site, and a further unique rare cutting enzyme sequence flanks a regulatory sequence. A first nucleotide sequence "flanks" a second nucleotide sequence if the two sequences are adjacent to one another, or if the two sequences are separated by no more than about 100 nucleotides, preferably by no more than about 50 nucleotides.

A first nucleotide sequence is "adjacent" to a second nucleotide sequence if the two sequences are attached to or positioned next to one another, having no intervening nucleotides. By way of example, the first nucleotide sequence 5'-AAAAA-3' is adjacent to the second nucleotide sequence 5'-TTT-3' when the two are connected thus: 5'-AAAAATTT-3' or 5'-TTTAAAAA-3', but not when the two are connected thus: 5'-AAAACTTT-3'.

In another embodiment, the heterologous nucleotide sequence or vector includes at least one gene encoding a cofactor for enhancing bioluminescence. As used herein, the term "cofactor" refers to an organic molecule, an inorganic molecule, a peptide, or a protein required for enzyme activity. The protein products encoded by the LUX genes may require the cofactors for regenerating and enhancing FMNH₂ pool, and fatty acid precursors in order to induce bioluminescence. These cofactors are typically synthesized in the chloroplasts.

In some applications of the present invention, the level of luminescence may be enhanced by additional transfection of a riboflavin operon (RIB operon) and/or genes encoding for fatty acid donors (*i.e.* genes belonging to the Fatty Acids Synthase [either FASI or FASII] pathway). In particular, in some embodiments, a component of the RIB operon (such as, for example, ribE and ribH (encoding riboflavin synthase or lumazine synthase, respectively)) involved in riboflavin synthesis and/or donors of fatty acids for the aldehyde synthesis, such as bacterial or plant acyl carrier protein

(ACP), can be transfected into a plant cell as part of a heterologous nucleotide sequence or vector. In plants, ACP exists as a small cofactor protein that participates in reactions of fatty acid biosynthesis and metabolism. Addition of these genes strongly boosts substrate production and increases bioluminescence levels.

Specific examples of suitable cofactors for enhancing bioluminescence include a polypeptide encoded by a RIB operon (GeneBank accession AF364106) (SEQ ID NO: 41) or RIB genes, bacterial acyl carrier protein, plant acyl carrier protein, transcriptional activators, and FRE flavin reductases enzymes from either luminescent (*P. luminescens* (GeneBank # D17745) (SEQ ID NO: 42) and *V. fischeri* (GeneBank # D17744) (SEQ ID NO: 43) or other bacteria (*E. coli* FRE, GeneBank #NC_010473) (SEQ ID NO: 44). Further examples of suitable cofactors include riboflavin kinases (RFK) such as plant *Arabidopsis thaliana* RFK (GeneBank #NC_003075) (SEQ ID NO: 45) or bacterial *E. coli* RFK (GeneBank #NC_009801) (SEQ ID NO: 46).

As used herein, “enhancing” bioluminescence refers to increased bioluminescent intensity or brightness that is greater than that without the cofactor. Enhancing bioluminescence may further include replenishing exhausted luciferin or other substrate or cofactor or other protein in order to continue or revive the reaction for bioluminescence.

The term “RIB operon” refers to an operon containing genes coding for proteins essential to production of riboflavin (commonly known as VitaminB2). The RIB operon in the bacteria belonging to the genus *Bacillus* includes following genes: *ribO* gene coding for control element, *ribG* gene coding for deaminase/reductase, *ribB* gene coding for riboflavin synthase (α-subunit), *ribA* gene coding for GTP-cyclohydrolase/3,4-dihydroxy-2-butanone-4-phosphate synthase, *ribH* gene coding for lumazine synthetase, and *ribT* gene coding for a protein with unknown function. Nucleotide sequences of *ribG*, *ribB*, *ribA*, *ribH* and *ribT* genes of *Bacillus subtilis* are presented in GenBank under accession numbers X51510 (*B. subtilis* riboflavin biosynthesis operon *ribG*, *ribB*, *ribA*, *ribH*, and *ribT* genes) (SEQ ID NO: 47). The *rib* genes for *Escherichia coli* include *rib*, *ribA*, and *ribE* code for GTP cyclohydrolase II, 3,4-dihydroxy-2-butanone 4-phosphate (DHBP) synthetase, and riboflavin synthetase, respectively. Nucleotide sequences of *rib*, *ribA*, and *ribE* genes of *E. coli* are presented in EBI under accession numbers ABV17158 (SEQ ID NO:

48) and CAA48861 (SEQ ID NO: 49), respectively. Similarly, *Photobacterium leioognathi*, strain PL741, RIB operon, encoding for *rib* E, H, B and A genes can be found at the GeneBank under accession number AF364106 (SEQ ID NO: 50).

The term “plant acyl carrier protein” or “bacterial acyl carrier protein” refer to any acyl carrier protein having the essential functional characteristics of naturally occurring ACP molecules found in plants or bacteria, respectively. Nucleotide sequences encoding a plant or bacterial acyl carrier protein include those presented in GenBank such as *Arabidopsis thaliana* ACP (EBI# X13708) (SEQ ID NO: 51) and *Photobacterium* sp. ACP (EBI #: EAR53459) (SEQ ID NO: 52).

In one embodiment, a transgenic bioluminescent plant as described above is provided in which a first heterologous nucleotide sequence includes, for example, a LUX operon and an inducible promoter. The plant also includes a second heterologous nucleotide sequence that includes a gene encoding a promoter and a gene encoding a transcription factor. In the present embodiment, activation of transcription of the LUX genes is indirect. See figure 7 and 12.

Preferably, the promoter for the first heterologous nucleotide sequence is inducible by a transcription factor in order to activate transcription of the LUX operon. A preferred example of such a promoter is a Gal4-UAS promoter, which is inducible by mGal4-VP16 chimeric transcription factor. See Figures 7 & 12.

Preferably, the promoter for the second heterologous nucleotide sequence is an inducible promoter, such as a heavy metal sensitive promoter from tobacco *cdiGRP* gene, or a tissue-specific promoter. Preferably, the second heterologous nucleotide sequence further includes a plastid targeting sequence and/or a reporter gene (*e.g.*, a gene that encodes a fluorescent protein).

The term “transcription factor” refers to any protein that is involved in the initiation of transcription, but which is not itself a part of the polymerase.

Transcription factors interact preferentially with specific nucleotide sequences, *i.e.*, regulatory sequences, and which in appropriate conditions stimulate transcription (“transcriptional activator”) or repress transcription (“transcriptional repressor”).

Preferably, the transcription factor sequence includes an *mGal4* binding domain, which is modified for optimal activity in *Arabidopsis* and a VP16 domain from

Herpes simplex virus. Such a sequence encodes a chimeric transcriptional activator mGal4-VP16.

Any combination of various sequences may be included in the heterologous nucleotide sequence or vector of the invention, such as sequences encoding a promoter, transcriptional leader, terminator element; sequences encoding a selectable and/or screenable marker; reporter genes; functional elements, such as an origin of replication, multiple cloning sites, border sequences; RNA processing signals; rare cutting enzyme sequences; sequences encoding a cofactor, and/or transcription factor, etc.

See, for example, Figure 3, which is a schematic illustration of four different examples of a heterologous nucleotide sequence in a pCas (donor vector) expression cassette. From top to bottom the examples include::

PI-PspI-P_{psbA}-L_{psbA}-MCS-T_{psbA}-PI-PspI;

I-CeuI-P_{rrn}-L_{rbcL}-MCS-Trp_{sl6}-I-CeuI;

I-SceI-P_{rrn}-AAGGAGG-MCS-T_{psbA}-I-SceI; and

I-PpoI-P_{rrn}-L_{psbA}-MCS-T_{psbA}-I-PpoI. Combinations of various promoter, transcriptional leader and terminator elements can be used to produce pCas vectors with different transcription regulation elements, and thus adjustable levels of transgene expression. Different homing endonucleases used to create a number of expression cassettes to be inserted in a single pUniPlast vector.

Method for Producing a Transgenic Bioluminescent Plant

In another aspect, the invention relates to a method for producing a transgenic bioluminescent plant. The method includes transfecting at least one plant cell with a vector, growing the at least one plant cell into a mature plant, and providing a means for regulating expression of the gene. The vector includes a heterologous nucleotide sequence that includes a bacterial LUX operon. The LUX operon includes LUX A, LUX B, LUX C, LUX D, LUX E, and LUX G genes ("the six LUX genes").

The term "transfecting" or "transforming" refers to introducing a nucleotide sequence into a host cell or into plastids of the cell. The nucleotide sequence that is

being introduced to the host cell or plastid of the cell may include a heterologous nucleotide sequence or a vector, as described above.

Transfection of the heterologous nucleotide sequences and plant genomic sequences is achieved by methods known to a skilled artisan. Any method that permits the introduction of a nucleotide sequence into a plant cell or a plastid of a cell is suitable. Examples of such methods include microinjection, electroporation, bombardment with DNA-coated microparticles, permeabilizing a cell with polyethylene glycol, fusion with other DNA-comprising units such as minicells, cells, lysosomes, or liposomes. Preferred methods include, for example, biolistic gene delivery and *Agrobacterium* mediated transformation.

In one embodiment, the heterologous nucleotide sequence that is being transfected is integrated in nuclear DNA of the plant cell. Typically, expression of a foreign gene in a plant is achieved by integration of the foreign gene into nuclear DNA. However, the number of copies of the foreign gene integrated into the transformed plant nuclear genome may be low and the expression levels may also be low. The term "nuclear DNA" refers to the part of the plant genome that is located in the nucleus of a plant cell. Preferably, the heterologous nucleotide sequence that is integrated in nuclear DNA of the plant cell also includes a chloroplast targeting sequence, as described above.

A suitable and well-known method for nuclear transformation or integrating the heterologous nucleotide sequence into nuclear DNA of the plant cell includes the natural transformation system of *Agrobacterium*. The Ti and Ri plasmids of *A. tumefaciens* and *A. rhizogenes*, respectively, carry genes responsible for genetic transformation of the plant. Other bacteria such as *Sinorhizobium*, *Rhizobium*, and *Mesorhizobium* that interact with plants naturally can be modified to mediate gene transfer to a number of diverse plants. These plant-associated symbiotic bacteria can be made competent for transfecting (gene transfer) by acquisition of both a disarmed Ti plasmid and a suitable binary vector. The heterologous nucleotide sequences to be transferred via an *Agrobacterium* -mediated transformation method may include one or more "border" sequences, such as right border (RB) and left border (LB) sequences that usually define the extent of the transferred DNA (T-DNA) containing one or

more genes to be expressed in a plant cell, and may further include an enhancer sequence such as an overdrive sequence or a plurality of overdrive sequences.

In preferred embodiment, heterologous nucleotide sequence that is being transfected is integrated in plastid DNA of the plant cell. "Plastid DNA" refers to that portion of the plant genome located in a plastid of a plant cell. Preferably, the heterologous nucleotide sequence is integrated in plastid DNA by homologous recombination. Once the heterologous nucleotide sequence is integrated in plastid DNA of the plant cell, it is preferable to avoid creating any double stranded breaks or digestion of the heterologous nucleotide sequence in the plastid. In addition, once the heterologous nucleotide sequence is integrated in plastid DNA, there is no need to remove a rare cutting enzyme sequence from the plastid genome.

Any method for stable transformation of a plastid is suitable. Such methods are well known in the art. Examples include use of a particle bombardment using "genegun" and insertion into the plastidic genome by homologous recombination. Other examples include electroporation, direct DNA transfer into plastids of protoplasts using polyethylene glycol. A preferred method is a biolistic transformation method, in which the nucleotide sequence to be transformed is applied to, for example, gold or tungsten particles, and accelerated towards the cell to be transformed. For purposes of the invention, the rare cutting enzymes are not used to transform a plastid.

The method for producing a transgenic bioluminescent plant further includes growing the plant cell into a mature plant. The term "mature plant" refers to a plant in which normal development of all vegetative and reproductive organs has occurred. After the above-mentioned heterologous nucleotide sequence or vector constructs are delivered into plant genomes by the methods described above, the transformed cells are regenerated into fully grown plants using standard hormone and antibiotic-supplemented media by methods known to a skilled artisan. (See Dandekar AM, and Fisk HJ. "Plant transformation: *Agrobacterium*-mediated gene transfer." *Methods Mol Biol.* 286:35-46 (2005)). The plants are then propagated either through seeds or clonal expansion.

The method further includes a means for regulating expression of the heterologous nucleotide sequence. Such means include adding a regulatory sequence or sequences encoding a transcription factor to the heterologous nucleotide sequence that is transfected into the plant cell. Regulatory sequences include those described
5 above, such as promoters, ribosomal binding sites, and termination sequences.

Vector for Transforming a Plastid

In another aspect, the invention relates to a vector for transforming a plastid. The vector includes a first and a second homologous recombination sequence, wherein each of the homologous recombination sequences is flanked by a unique rare
10 cutting enzyme sequence. The vector further includes a multiple cloning site that includes at least one unique rare cutting enzyme sequence.

Suitable homologous recombination sequences typically vary depending on the plant species of the host cell or plastid. Examples of suitable pairs of a first and a second homologous recombination site include *trnI/trnA*; *trnV/rps12*; *trnfM/trnG*,
15 derived from pPRV312L (GeneBank accession number DQ489715), pPRV111A (GeneBank accession number U12812) and pRB94 (GeneBank accession number AJ312392), respectively.

Preferably, each of the homologous recombination sites is flanked by a unique rare cutting enzyme sequence. The term “unique” with reference to a rare cutting
20 enzyme sequence refers to a rare cutting enzyme sequence which occurs only once in a given nucleotide sequence.

In one embodiment, each gene in the heterologous nucleotide sequence or vector is flanked by a unique rare cutting enzyme sequence on both the 5'- and 3'- ends of the homologous recombination sites. For example, a homologous
25 recombination site of *trnI* is preferably flanked by two rare cutting enzyme sequences, *e.g.*, AscI/I-CeuI, and a homologous recombination site of *trnA* is preferably flanked by two rare cutting enzyme sequences, *e.g.*, I-SceI/FseI, which will maintain proper directionality of the heterologous nucleotide sequence when it is transfected in a plant cell. In addition, those skilled in the art can appreciate that flanking the homologous
30 recombination site with a unique rare cutting enzyme sequence allows for efficient deletion and replacement of the homologous recombination site with another

homologous recombination site without affecting the remaining heterologous nucleotide sequence. This allows rapid adaptation of such plastid transformation vector (e.g. pUniPlast) for the use with any plant species of choice.

The term “multiple cloning site” refers to a nucleotide sequence that contains one or a number of different enzyme recognition sites to permit insertion of a nucleotide sequence at a defined locus (*i.e.*, the enzyme recognition site) on a vector. Preferably, the multiple cloning site is flanked by a unique rare cutting enzyme sequence. In one embodiment, the multiple cloning site includes more than one unique rare cutting enzyme sequence. See Figure 6.

The multiple cloning site may permit the introduction of a heterologous nucleotide sequence containing a gene into a vector. In one embodiment, the multiple cloning site includes a full length LUX operon. In a further embodiment, the multiple cloning site includes at least the LUX A, LUX B, LUX C, LUX D, LUX E, and LUX G genes. In another embodiment, the multiple cloning site also includes at least one gene encoding a cofactor for enhancing bioluminescence as described above.

Method for Producing a Vector for Transforming a Plastid

In another aspect, the invention relates to a method for producing a vector for transforming a plastid. The method includes providing an acceptor vector, which includes (i) a first and a second homologous recombination site, wherein each of said homologous recombination site is flanked by a unique rare cutting enzyme sequence; and (ii) a multiple cloning site comprising at least one unique rare cutting enzyme sequence. The method further includes providing at least one donor vector, wherein the donor vector comprises a heterologous nucleotide sequence flanked by a rare cutting enzyme sequence that is sufficiently identical to the at least one unique rare cutting enzyme sequence of the multiple cloning site of the acceptor vector. The method also includes sequentially subcloning the heterologous nucleotide sequence from the donor vector into the acceptor vector, whereby a vector for transforming a plastid is produced.

The method allows for multiple rounds of gene transfer by sequentially subcloning a gene from multiple, different donor vectors into an acceptor vector. The multiple donor vectors will be used in different sequential rounds of cloning to allow

sequential insertion of genes or heterologous nucleotide sequences into the acceptor vector.

The term “acceptor vector” refers to a vector that is a recipient of a heterologous nucleotide sequence from a donor vector. The acceptor vector is characterized by having a first and a second homologous recombination site, as described above. Accordingly, an acceptor vector is capable of integrating into the genome of a plant or plastid. In addition, the acceptor vector includes a multiple cloning site, as described above. Preferably, the acceptor vector includes unique rare cutting enzyme sequences throughout the vector, including in the multiple cloning site, so that integrity of a heterologous nucleotide sequence or the acceptor vector is maintained during genetic manipulation protocols.

In a preferred embodiment, the acceptor vector is derived from the pUniPlast vector system, which is a plastid transformation vector system developed by BioGlow Inc. In another preferred embodiment, the acceptor vector is a binary pPZP-RCS vector, described above (Goderis, 2002 and Tzfira, 2005). Preferably, the acceptor vector is introduced or transfected into a host cell or into a plastid *in vivo*.

The term “donor vector” refers to a vector containing a heterologous nucleotide sequence that is transferred to an acceptor vector through gene recombination. Preferably, the donor vector is not introduced into a host cell or into a plastid *in vivo*. The heterologous nucleotide sequence from the donor vector is preferably transferred to an acceptor vector *in vitro*. In a preferred embodiment, the donor vector is a pCas vector, developed by BioGlow Inc, and pSAT vectors (Tzfira, 2005).

The heterologous nucleotide sequence of the donor vector is preferably flanked by a rare cutting enzyme sequence that is sufficiently identical to the at least one unique rare cutting enzyme sequence of the multiple cloning site of the acceptor vector. A first rare cutting enzyme sequence is sufficiently identical to a second rare cutting enzyme sequence if one or the same rare cutting enzyme is capable of recognizing and cleaving both the first and the second rare cutting enzyme sequences.

The term “sufficiently identical” as used herein refers to a first nucleotide sequence that contains a sufficient or minimum number of identical or equivalent

nucleotides to a second nucleotide sequence, such that the first and second nucleotide sequences share common structural domains or motifs and/or a common functional activity. For example, nucleotide sequences that share common structural domains having at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%,
5 or more identity across the sequences, and share a common functional activity are defined herein as sufficiently identical.

To determine percent identity of two nucleic acid sequences, the sequences are aligned for optimal comparison purposes (*e.g.*, gaps can be introduced in one or both of a first and second nucleotide sequence for optimal alignment). For example, when
10 aligning a first rare cutting enzyme sequence to a second rare cutting enzyme sequence having 10 nucleotides, at least 70%, preferably at least 80%, more preferably at least 90% of the 10 nucleotides between the first and second sequences are aligned. When a position in the first sequence is occupied by the same nucleotide as the corresponding position in the second sequence, then the molecules are identical
15 at that position. The percent identity between the two sequences is a function of the number of identical positions shared by the sequences, taking into account the number of gaps, the length of the sequences, and the length of each gap that need to be introduced for optimal alignment of the two sequences. An algorithm known in the art may be used to determine percent identity between two sequences.

20 Conventional cloning or subcloning techniques, reagents, conditions, enzymes may be used to introduce, clone, or subclone the heterologous nucleotide sequence into the multiple cloning site of the acceptor or donor vector. One or more genes may be cloned or subcloned as a group into the acceptor or donor vector.

The term “subcloning” refers to a transfer of a nucleotide sequence from an
25 initial vector to a more specialized vector. For example, subcloning refers to a transfer of a nucleotide sequence from a donor vector to an acceptor vector. Typically, subcloning includes removing the nucleotide sequence from the initial vector with suitable nucleases and then inserting the removed nucleotide sequence into another vector. Preferably, rare cutting enzymes are used in the subcloning.

30 The term “sequentially” subcloning refers to transferring a nucleotide sequence from a donor vector to an acceptor vector in a stepwise manner.

Sequentially subcloning allows for repeated cycles of subcloning until all of the heterologous nucleotide sequences from a donor vector are delivered into the acceptor vector, while maintaining integrity of the acceptor vector and/or the heterologous nucleotide sequences inserted in the acceptor vector from a previous cycle of subcloning.

In a further embodiment, the method includes a second donor vector that includes a heterologous nucleotide sequence. Preferably, the heterologous nucleotide sequence of the second donor vector is flanked by a rare cutting enzyme sequence that is unique from a rare cutting enzyme sequence of the first donor vector.

For example, at least two donor vectors may be used to express multiple genes in plastids. See Figure 5, for example. A first donor vector may be used to subclone a LUX operon into an acceptor vector and a second donor vector may be used to subclone a RIB operon into the same acceptor vector. Accordingly, the resulting acceptor vector can include more than one operon. The resulting acceptor vector may then be transfected into a plant or plastid for simultaneous transformation of the LUX and RIB operons.

In yet another embodiment, the method for producing a vector for transforming a plastid further includes subjecting a host cell containing the acceptor or donor vector to drug selection or other form of screening/selection.

Other possible modifications and variations can be made to the method, which may include the number of donor vectors; the number and/or arrangements of the rare cutting enzyme sequences; the number and/or arrangements of the heterologous nucleotide sequences, etc. For example, three or more donor vectors with different heterologous nucleotide sequences can be sequentially subcloned into the acceptor vector.

Vector System Kit

In another aspect of the invention, a vector system kit is provided. The kit includes an acceptor vector comprising (i) a first and a second homologous recombination site, wherein each of said homologous recombination site is flanked by a unique rare cutting enzyme sequence; and (ii) a multiple cloning site comprising at

least one unique rare cutting enzyme sequence. The kit further includes at least one donor vector, wherein the donor vector comprises a heterologous nucleotide sequence flanked by a rare cutting enzyme sequence that is sufficiently identical to the at least one unique rare cutting enzyme sequence of the multiple cloning site of the acceptor vector.

The kit can further include reagents, buffers, and materials related to any of the nucleotide sequences and proteins described above. In addition, the kit can include a plant or plant cell produced by the invention.

Variants

The present invention further relates to variants of the nucleotide sequences described herein. Variants may occur naturally, such as a natural allelic variant. Other variants include those produced by nucleotide substitutions, deletions, or additions. The substitutions, deletions, or additions may involve one or more nucleotides. These variants may be altered in coding regions, non-coding regions, or both. Alterations in the coding regions may produce conservative or non-conservative amino acid substitutions, deletions, or additions. Preferably, the variant is a silent substitution, addition, or deletion, which does not alter the properties and activities of the peptide encoded by the nucleotide sequence described herein. Conservative substitutions are also preferred.

Further embodiments of the invention include variant nucleotide sequences comprising a sequence having at least 90% identical, and more preferably at least 95%, 96%, 97%, 98%, or 99% identical to a nucleotide sequence described herein. The nucleotide sequences described herein are the "reference" sequences.

For example, a variant nucleotide sequence that is at least 95% identical to a reference nucleotide sequence (*e.g.*, the LUX operon) described herein is identical to sequence described herein except that the variant nucleotide sequence may include up to five point mutations per each 100 nucleotides of the reference nucleotide sequence (*e.g.*, the LUX operon) sequence described herein.

In other words, to obtain a variant nucleotide sequence that is at least 95% identical to a reference nucleotide sequence described herein, up to 5% of the

nucleotides in the reference sequence may be deleted or substituted with another nucleotide, or a number of nucleotides up to 5% of the total nucleotides in the reference sequence may be inserted into the reference sequence.

These mutations of the reference sequence may occur at the 5' or 3' terminal positions of the reference nucleotide sequence or anywhere between those terminal positions, interspersed either individually among nucleotides in the reference sequence or in one or more contiguous groups within the reference sequence.

Uses

Environmental safety:

In some embodiments of the present invention, one or both of the luciferase subunits, for example subunit A, can be controlled by a promoter sensitive to heavy metals, such as cadmium, a known carcinogen heavily used in industry. Specific promoter that may be used is one controlling the expression of tobacco cdiGRP gene in tobacco. (Ueki et al., "Identification of an interactor of cadmium ion-induced glycine-rich protein involved in regulation of callose levels in plant vasculature," *PNAS* 102(34):12089-94 (2005)). This type of transgenic crop starts glowing if exposed to heavy metals. Such embodiment is highly beneficial for monitoring of industrial waste, by simply planting these crops in the area where pollution may occur, and it will become literarily visible within a few hours.

Ornamental Applications:

The glowing plants can be used in ornamental applications. In 1992, the global flower trade was estimated at one hundred billion U.S. dollars. In recent years, the floral industry has grown six percent annually, while the global trade volume in 2003 was 101.84 billion U.S. dollars. In the U.S. alone, floral industry at 2006 was estimated at approximately twenty-one billion U.S. dollars, representing a considerable market (http://www.aboutflowers.com/press_b1.html). Although it is hard to estimate market share of a product which has not existed before, taking to the consideration a novel and appealing appearance of this metabolically engineered plants, one may reasonably suggest that initial market share can be somewhere

between 5-10%, representing one to two billion dollars in the U.S. and five to ten billion dollars worldwide.

In some embodiments of the present invention, plants are engineered in such way that the luminescence color is modified by shifting wavelength of the emitted light either by luciferase mutagenesis or simply co-expressing with other fluorescent proteins. In some embodiments of the present invention, plants are engineered in such way that organs of the plant where the luminescence is produced is directed by expressing system components under organ-specific promoters. These embodiments allow production of multiple product lines, enabling market share increase.

Basic research:

The present invention is a new step in reporter gene technology, as it allows detection of gene expression without need of any type of additional equipment or reagents, just by observation. For example, it will simplify the procedure of identification of new promoters: leaving one of the genes in the system promoter-less (for instance, luciferase A subunit) thus allowing cloning of a genomic library upstream to it. Simple observation of growing plants allows determination of where cloned promoters are expressed and at which plant life stages.

INCORPORATION OF SEQUENCE LISTING

Incorporated herein by reference in its entirety is the Sequence Listing for the application. The Sequence Listing is disclosed on a computer-readable ASCII text file titled, "sequence_listing.txt", created on July 31, 2008. The sequence_listing.txt file is 205 kb in size.

EXAMPLES

Example 1: Cloning pUniPlast Vector

A donor vector, pSat1 vector (vector used for genomic plant DNA transformation (Tzfira, et al. 2005), is digested with AscI restriction enzyme. Blunt ends are then created with Mung Bean nuclease (New England Biolabs). A pUniPlast

(acceptor vector) multiple cloning site is blunt cloned with annealed primers BG-1 and BG-2. The resulting construct is then sequenced with primers BG-3 and 4.

The *trnI* (approx. 1,240bp) homologues recombination targeting sequence is amplified by polymerase chain reaction (PCR) using *Pfu* polymerase (Stratagene), tobacco plastid DNA as a template and primers BG-5 and 6. The *trnI* is then cloned with *AscI*/*I-CeuI* to maintain insert directionality (double digest in NEB buffer 4).

The *trnA* (approx. 830bp) homologues recombination targeting sequence is amplified by polymerase chain reaction (PCR) using *Pfu* polymerase, tobacco plastid DNA as a template and primers BG-7 and 8. The *trnA* sequence is then cloned with *FseI*/*I-SceI* to maintain insert directionality (double digest in NEB buffer 4).

Sequences for BG primers can be found in Table 1 below.

TABLE 1: BG Primers List

<i>Primer #</i>	<i>Sequence</i>
BG-1	GGCGCGCCTAACTATAACGGTCCTAAGGTAGCGATGGCAAACAGCTA TTATGGGTATTATGGGTTTAATTAATAGGGATAACAGGGTAATGGCC GGCC (SEQ ID NO: 58)
BG-2	GGCCGGCCATTACCCTGTTATCCCTATTAATTAACCCATAATACCCA TAATAGCTGTTTGCCATCGCTACCTTAGGACCGTTATAGTTAGGCGCG CC (SEQ ID NO: 59)
BG-3-seq	CTTCGCTATTACGCCAGCTGG (SEQ ID NO: 60)
BG-4-seq	GTTAGCTCACTCATTAGGCAC (SEQ ID NO: 61)
BG-5- <i>AscI</i> - <i>trnI</i>	CTGGCGCGCCGCTTGTACACACCGCCCGTCAC (SEQ ID NO: 62)
BG-6- <i>CeuI</i> - <i>trnI</i>	TCGCTACCTTAGGACCGTTATAGTTAGCAGCTGGGCCATCCTGGACTT G (SEQ ID NO: 63)
BG-7- <i>SceI</i> - <i>trnA</i>	AATAGGGATAACAGGGTAATCTCCACTTGGCTCGGGGGGATATAG (SEQ ID NO: 64)
BG-8- <i>FseI</i> - <i>trnA</i>	TATGGCCGGCCAGCTTTGTATCGGCTAAGTTCACG (SEQ ID NO: 65)
BG-9- <i>AgeI</i> - PpsbA	TCACCGGTGGGCAACCCACTAGCATATCG (SEQ ID NO: 66)

BG-10-NcoI-PpsbA	AACCATGGTAAAATCTTGGTTTATTTAATC (SEQ ID NO: 67)
BG-11-XbaI-TpsbA	CCTCTAGAGATCCTGGCCTAGTCTATAGG (SEQ ID NO: 68)
BG-12-NotI-TpsbA	TTATGCGGCCGCTCGAATATAGCTCTTCTTTCTTATTTC (SEQ ID NO: 69)
BG-13-aadPsbA	AACCATGGGGGAAGCGGTGATCGCCGAAG (SEQ ID NO: 70)
BG-14-BglII-aadA	TGGAGATCTTTATTTGCCGACTACCTTGGTGATC (SEQ ID NO: 71)
BG-15-AgeI-Prrn	TCACCGGTCGCCGTCGTTCAATGAGAATGG (SEQ ID NO: 72)
BG-16-NcoI-Prrn	AACCATGGATATTCGCCCGGAGTTCGCTC (SEQ ID NO: 73)
BG-17-XbaI-Trps16	CCTCTAGAACCGAAATTCAATTAAGGAAAT (SEQ ID NO: 74)
BG-18-EcoRIrem	TTATGCGGCCGCAGAACACGCATTTCAATGGAAGC (SEQ ID NO: 75)
BG-19-XhoGFPfw	CCACTCGAGAGGAGGGATTTATGGTGAGCAAGGGCGAGGAGC (SEQ ID NO: 76)
BG-20-EcoRI-GFP	GGGAATTCTTACTTGTACAGCTCGTCCATGCCGAG (SEQ ID NO: 77)
BG-21-NcoI-GFP	TTCCATGGAGTTGTAGGGAGGGATTTATGGTGAGCAAGGGCGAGGAGC (SEQ ID NO: 78)
BG-22-fwLUX	ACAGAATTCCCAAAGGAGATTACATGATTAAG (SEQ ID NO: 79)
BG-23-revLUX	TTGGAATTCTTACGTATAGCTAAATGCATCAG (SEQ ID NO: 80)
BG-24-GUSXhoI	CCACTCGAGAGGAGGGATTTATGTTACGTCCTGTAGAAACC (SEQ ID NO: 81)
BG-25-rrnaadA	AACCATGGAGTTGTAGGGAGGGATTTATGGGGGAAGCGGTGATCGCC (SEQ ID NO: 82)
BG-26-aadXho	TTGCTCGAGTTATTTGCCGACTACCTTGGTGATC (SEQ ID NO: 83)
BG-27-GUSEcoRI	GGGAATTCTCATTGTTTGCCCTCCCTGCTGC (SEQ ID NO: 84)

**Example 2: Preparing a pCas-1 vector containing I-CeuI derived
PpsbA/LpsbA/TpsbA expression cassette**

A *PpsbA* promoter (approx. 230bp) is amplified by PCR using *Pfu*
5 polymerase, using tobacco plastid DNA as a template and primers BG-9 and 10. The
PCR product is then digested with AgeI/NcoI enzymes.

The pSat5 (vector used for genomic plant DNA transformation, see Tzfira et al
(2005) (donor vector) is then digested with AgeI/NcoI to remove a 35S promoter.
The backbone of the vector and the *PpsbA* promoter is then ligated using T4 DNA
10 ligase (New England Biolabs).

A *TpsbA* terminator (approx. 400bp) is amplified by PCR using *Pfu*
polymerase, tobacco plastid DNA as a template, and primers BG-11 and 12. The PCR
product is then digested with NotI/XbaI.

The pSat5- *PpsbA* promoter is digested with NotI/XbaI to remove a 35S
15 terminator (note: may need to use INV110 to prevent XbaI site methylation). The
TpsbA terminator is then cloned into the pSAT5-*PpsbA* promoter vector. From now
on this vector is referred as pCas-1 (donor vector).

A *aadA* gene (spectinomycin selection marker; approx. 800bp) is PCR
amplified using *Pfu* polymerase, binary plasmid pPZP-RCS1 as a template (vector
20 used for genomic plant DNA transformation, see Tzfira et al (2005), and primers BG-
13 and BG-26 primer. The *aadA* PCR product is then digested using NcoI/XhoI.

The pCas-1 vector is digested with NcoI/XhoI. The *aadA* product is ligated
into the pCas-1 using T4 DNA ligase. The resulting vector is further referred as pCas-
1-*aadA*.

25 pCas vectors contain pre-cloned expression cassettes, including promoter and
terminator sequences, as well as selection and fluorescent/enzymatic markers, if
required. These cassettes to be excised from the pCas vectors using rare-cutter
enzymes and introduced into plastid transformation pUniPlast vector, containing
homologues recombination plastid targeting sequences for creation of transgenic

plants. Specifically, if transgene expression is not satisfactory, it can be re-cloned in a single cloning step into another pCas vector with desired regulatory elements (i.e. strong or weak promoter, etc) and subcloned in a second step into an existing pUniPlast vector, creating desired expression vector in two cloning steps.

Certain pCas vectors have an array of pre-arranged autofluorescent and enzymatic tags, such as green fluorescent protein (GFP) and beta-glucuronidase (GUS), which can either be fused or co-expressed on the same polycistronic mRNA with the transgene of interest. The transgene of interest can be rapidly reshuffled between different tags in a single cloning step and further be assembled into a single vector pUniPlast vector, allowing rapid autofluorescent or enzymatic tagging.

Similarly, to exchange fluorescent and enzymatic tags, different pCas-1 vectors contain different selection markers, such as spectinomycin resistance encoded by the *aadA* gene or kanamycin resistance encoded by the *neo* gene, or fluorescent protein markers, such as *GFP*. Expression cassettes containing these resistance markers can be easily reshuffled within the pUniPlast vector, without affecting the other sequences cloned, and allowing for plant recovery on different selective media. Extended MCS allows expression of the transgenes on the same polycistronic mRNA with the selection marker. Furthermore, selected pCas vectors bearing selection markers contain pre-engineered homology-based excision elements, such as *Cre-lox* and *attB/attP* recognition sequences, allowing removal of the antibiotic resistance genes using phage site-specific recombinases.

See Figure 8 (A) Plasmids pSat4 and pSat5 have been digested with *AgeI/NcoI* to remove the 35S promoter (approx. 900bp). Figure 8(B) *psbA* (approx. 230bp) and *rrn* (approx. 100bp) promoters were PCR amplified from plastid DNA of *Nicotiana tabacum*. The PCR fragments to be digested with *AgeI/NcoI* and ligated into the same sites of pSat4 and pSat5 to replace 35S promoter and convert the pSat plasmids to pCas vectors, suitable for use in plastids. Figure 8 (C) *rrn* promoter cloned into pCas vector. Lane 1: undigested pCas-3 vector; Lane 2: pCas-3 vector digested with *AgeI/NcoI*, showing presence of the *Prrn*.

Example 3: Preparing pCas-2 vector containing PI-PspI derived *PpsbA/LpsbA/TpsbA* expression cassette

A pCas-2 vector (donor vector) is made following the same steps above as for pCas-1 vector. Plasmid pSat6 (vector used for genomic plant DNA transformation, see Tzfira et al (2005)) (donor vector) is used as a basis vector to create pCas-2.

Several different expression cassettes may be required to express multiple genes in plastids, and thus several different pCas vectors (donor vectors) must be made. For example, pCas-1 will be employed to subclone LUX operon (NCBI accession: M63594) into pUniPlast (acceptor vector) for generation of bioluminescent transgenic plants.

However, if additional genes need to be introduced into the same pUniPlast (acceptor) vector for simultaneous transformation into plastid genome, an additional pCas vector (donor vector), such as pCas-2 or pCas-3, might be required.

For example, RIB operon genes (GeneBank accession AF364106) encoding for the synthesis of riboflavin, one of the substrates derivatives required for the luminescent reaction, might need to be introduced into plant simultaneously with LUX operon to increase bioluminescence level. In this case, pCas-2 or pCas-3 vectors can be used to introduce RIB operon into an existing pUniPlast vectors containing LUX operon introduced using pCas-1.

Example 4: Preparing pCas-3 donor vector containing I-SceI derived *Prrn*/*LrbcL*/*Trps16* expression cassette.

The *Prrn* promoter (approx 100bp) is PCR amplified using *Pfu* polymerase, tobacco plastid DNA as a template and primers BG-15 and 16. The *Prrn* PCR product is then digested with AgeI/NcoI.

The pSat4 (vector used for genomic plant DNA transformation, see Tzfira et al)(donor vector) is digested with AgeI/NcoI to remove 35S promoter. The donor vector backbone and the *Prrn* promoter are ligated using T4 DNA ligase.

The *Trps16* terminator is PCR amplified using *Pfu* polymerase, tobacco plastid DNA as a template and primers BG-17 and primer BG-18. The PCR product (approx. 170bp) is digested with NotI/XbaI.

The pSat4-*Prrn* promoter is digested with NotI/XbaI to remove 35S terminator (note: may need to use INV110 to prevent XbaI site methylation). The *Trps16* terminator is then ligated into pSat4-*Prrn* promoter using T4 DNA ligase. From now on this vector is referred as pCas-3 (donor vector).

5 The *aadA* gene (spectinomycin selection marker; approx. 800bp) is PCR amplified using *Pfu* polymerase, binary plasmid pPZP-RCS1 as a template (vector used for genomic plant DNA transformation, see Tzfira et al (2005)), and primers BG-14 and BG-25. The *aadA* PCR product is digested using NcoI/BglII. The pCas-3 vector is then digested with NcoI/BglII. The *aadA* product is ligated into pCas-3
10 using T4 DNA ligase. The resulting vector is further referred as pCas-3-*aadA* (donor vector).

See Figure 13. pCas-3 donor vector was prepared by BioGlow Inc. pCas-3 vector, containing *rrn* promoter (*Prrn*), multiple cloning site (MCS) and *rps16* terminator (*Trps16*) has been prepared by BioGlow Inc. Transgenes of interest is
15 cloned into pCas-3 multiple cloning site and the cassette comprised of *Prrn*/Transgene of interest/*Trps16* is excised using I-SceI homing endonuclease and subcloned into pUniPlast acceptor vector for generation of transgenic plants.

In Figure 13, pCas-3 enzymatic digest displays presence of *Prrn* (AgeI/NcoI fragment), *Trps16* (NotI/XbaI fragment) and the whole *Prrn*/MCS/*Trps16* cassette (I-SceI fragment), showing that a complete pCas-3 has been constructed (BioGlow Inc,
20 unpublished).

Cloning of Pre-Arranged Fluorescent and Enzymatic Marker Into pCAS Cassettes

25 **Example 5: Cloning of Green Fluorescent Protein (GFP) into pre-arranged pCas vector for expression in a single operon with the *aadA* selection marker.**

Green fluorescent protein (GFP) is PCR amplified using *Pfu* polymerase, pSat6-EGFP (Tzfira et al, 2005) as a template, and primers BG-19 and primer BG-20. The PCR product (approx. 700bp) is digested with XhoI/EcoRI. The PCR product is

then cloned into the same sites (XhoI/EcoRI) in pCas-1-*aadA* and pCas-3-*aadA*. The resulting vectors are further referred as pCas-1-*aadA-GFP* and pCas-3-*aadA-GFP* (donor vectors).

The *aadA-GFP* cassette is then subcloned into I-CeuI site of pUniPlast (acceptor vector). The resulting pUniPlast-GFP operon construct GFP expressing transgenic plants (see below for plant transformation protocol) are then used.

Example 6: Cloning of Beta-Glucuronidase (GUS) into pre-arranged pCas vector for expression in a single operon with the *aadA* selection marker.

Beta-Glucuronidase (GUS) (approx. 1,800bp) is PCR amplified using *Pfu* polymerase, pSat6-GUS (Tzfira et al, 2005) as a template, and primers BG-24 and BG-27. The PCR product is digested with XhoI/EcoRI. The PCR product is then cloned into the same sites (XhoI/EcoRI) in pCas-1-*aadA*. The resulting vectors are further referred as pCas-1-*aadA-GUS* (donor vector).

Example 7: Cloning of the LUX Operon from *Photobacterium leiognathi* into pUNIPLAST Vectors To Create Bioluminescent Plants

The full length LUX operon is PCR amplified from *P leiognathi* using *Pfu* polymerase, *P leiognathi* genomic DNA as a template, and primers BG-22 and BG-23. EcoRI is then used to digest the LUX operon PCR fragment. The fragment is then cloned into the same sites (EcoRI) in pCas-1-*aadA*.

The *aadA-LUX* operon cassettes from pCas1-*aadA-LUX* operon are excised using I-CeuI homing endonuclease. The *aadA-LUX* operon is then subcloned into I-CeuI site of pUniPlast. The resulting pUniPlast-LUX operon construct is then used to produce bioluminescent transgenic plants.

To create bioluminescent plants, LUX operon from a variety of luminescent bacteria can be used, for example, the LUX operon from *Photobacterium leiognathi*., American Tissue Culture Collection strain #25521; NCBI accession: M63594.

Example 8: Assessment of pUNIPLAST System Workability in Plants Using Transient Gene Expression Assay

A transient gene expression assay in chloroplasts is used to assess effectiveness of transgene expression from pCas cassettes. pCas-1-*aadA*-GFP and pCas-1-*aadA*-GUS will be employed for transient expression of GFP and GUS, respectively, in tobacco leaves by the means of biolistic gene delivery. Shortly, 25 µg DNA of each of the constructs expressing the *aadA*-GFP or *aadA*-GUS transcriptional fusions is to be adsorbed onto 10 mg of 1-µm gold particles (Bio-Rad, CA) and bombarded at 150–200 psi into the leaf epidermis of greenhouse-grown *Nicotiniana glauca* plants using a Helios gene gun (PDS-1000/He, Bio-Rad). After incubation for 24 h at 22–24°C, the bombarded tissues are to be viewed under a Zeiss LSM 5 Pascal confocal laser scanning microscope (for GFP) or stained with GUS substrate X-Gluc (for GUS activity).

Example 9: Creation of Bioluminescent Transgenic Plants

The pUniPlast LUX operon is used to produce transgenic bioluminescent plants. *Nicotiniana glauca* plants will be used as a model organism. The following is a preferred protocol for stable expression of LUX operon in the transgenic plants:

Tobacco leaves are harvested at five- to seven-leaf stage of plant growth. An autoclaved Whatman 70-mm circle filter disk is placed on RMOP medium in a deep Petri dish. The leaf on filter disk is placed with its adaxial side facing the medium. The expression vectors are then biolistically delivered into plant tissue (see above). Parafilm is then used to seal the Petri dish.

After 2 days in dark at 24–26°C, 5-mm pieces of bombarded leaves are cut and placed on RMOP selection medium (bombarded side in contact with medium) for the first round of selection. Parafilm is then used to seal the Petri dish. Spectinomycin (500 µg/ml) is used for effective selection of transformants in tobacco. In a culture room, the Petri dishes are sealed with parafilm under white fluorescent lamps (1,900 lux) with 16 h light/8 h dark cycle at 24–26°C. After 4–8 weeks, putative transgenic shoots appear. The putative transplastomic shoots are then screened for transgene integration by PCR; LUX operon primers might be used (BG 22, 23).

For the second round of selection, 2-mm pieces of leaves are cut from the first round plants and are then placed on RMOP selection medium. The leaf sections are then grown in a culture room under white fluorescent lamps (1,900 lux) with 16 h

light/8 h dark cycle at 24-26⁰C. These leaf sections produce transgenic shoots in 3–4 weeks.

The regenerated shoots are then excised and transferred to MSO medium containing the appropriate antibiotic. The shoots are grown in culture room under white fluorescent lamps (1,900 lux) with 16 h light/8 h dark cycle at 24-26⁰C. This step is termed the third round of selection, where rooting occurs in 3–4 weeks. Southern blot analysis can be carried out to confirm integration and determine, as known in the art.

RMOP medium

<u>Material</u>	<u>Amount per 1 liter</u>
MS (Murashige and Skoog salts)	according to manufacturer's instructions
myo-inositol	100 mg
thiamine HCl	1 mg
BAP (6-Benzylamino purine)	1 mg
NAA (1-Naphthaleneacetic acid)	0.1 mg
Sucrose	30 g
BactoAgar	6 g
Adjust pH to 5.8	

MSO medium

<u>Material</u>	<u>Amount per 1 liter</u>
MS (Murashige and Skoog salts)	according to manufacturer's instructions
Sucrose	30 g
BactoAgar	6 g
Adjust pH to 5.8	

Example 10: Vector for nuclear expression of LUX genes

Endonucleases, homing endonucleases or gateway technology is used to clone LUX genes into either plant cellular or chloroplast expression vectors. The LUX genes are then delivered into nuclear or chloroplast genomes by methods known to a skilled artisan and expression is governed by corresponding regulatory sequences. In

particular, the LUX operon is transferred into a plant, which can be preferably achieved by the use of recently developed pSAT vector system (Tzfira et al., (2005) “pSAT vectors: a modular series of plasmids for autofluorescent protein tagging and expression of multiple genes in plants.” *Plant Mol. Biol.*, 57(4):503-16 (2005)),
5 allowing simultaneous transfer of 12-15 (or more) genes into the same organism, or by direct delivery of the operon into chloroplast genome.

For example, LUX genes are PCR amplified and subcloned into MCS of pSAT vectors (one gene per one vector), under plant-specific regulatory promoter and terminator sequences, such as 35S promoter and terminator. Different pSAT vectors
10 can contain different promoters (such as 35S, nopaline synthase (NOS), manopine synthase (MAS), etc., to prevent gene silencing when several consecutive genes on the chromosome are expressed from the same promoter. Further, the cassettes containing promoter, lux gene and a terminator are subcloned from the SAT vectors into PZP-RCS binary vector using homing endonucleases, such as PI-PspI or I-CeuI.
15 The resulting single binary PZP-RCS-Lux plasmid is used for plant transformation.

The following procedure is used to clone individual LUX genes into expression cassettes of vecors, e.g., pSAT.

A restriction map of the sequence to be cloned (i.e. specific LUX gene) is generated using appropriate software, such as “DNA Strider”. The restriction
20 enzymes which do not have restriction sites within a given gene are identified. These enzymes can be used to clone the LUX gene into a cloning vector. For instance, LUX A of *P. leiognathi* is has no restriction sites for enzymes BglII and EcoRI.

Design of the forward (fw) primer. Primer’s gene specific sequence should be 20-25bp long, in order to anneal to the target sequence. A restriction enzyme
25 polynucleotide sequence is to be added to the gene specific sequence. Gene specific forward sequence used for the LUX A gene is: ATGAAAATTAGTAATATCTG (SEQ ID NO: 85). Since BglII restriction site is located closer to promoter than EcoRI in the pCAS/pSAT multiple cloning site, this enzyme is used with the forward primer to keep the directionality of the gene correct, i.e. 5’ of the gene directly
30 following the promoter sequence. NOTE: several nucleotides (underlined) must be added at the 5’ of the primer for efficient digest of the PCR fragment. For instance,

BglII requires 3 nt to be added. Thus, final 5' primer for cloning of LUX A into a cloning vector using BglII will be: GAAGATCTATGAAAATTAGTAATATCTG (SEQ ID NO: 86) (BglII restriction site in bold).

Design of the reverse (rev) primer. Primer's gene specific sequence should be 20-25bp long, in order to anneal to the target gene. A restriction enzyme polynucleotide sequence is to be added to the gene specific sequence. Gene specific reverse primer sequence used for the LUX A gene is:

ACTTAAAAGATCCTCAGTAA (SEQ ID NO: 87). Since EcoRI site is located closer to terminator than BglII within the pSAT/pCAS multiple cloning sites, we will use EcoRI enzyme with the LUX A reverse primer to keep the directionality of the gene correct, i.e. 3' of the LUX A gene directly preceding the terminator. NOTE: several nucleotides (underlined) must be added at the 5' end of the primer for efficient digest of the PCR fragment. For instance, EcoRI requires 1-2 nt to be added. Thus, reverse primer for cloning of LUX A into a cloning vector using EcoRI will be:

ACTTAAAAGATCCTCAGTAAGAATTCCG (SEQ ID NO: 88). Finally, the reverse primer sequence is reversed in order to constitute a compatible pair with the forward primer (i.e. both sequences must allow DNA replication in opposite direction to one another), and thus the final LUX A reverse primer is:

CGGAATTCTTACTGAGGATCTTTTAAGT (SEQ ID NO: 89).

LUX genes A, B, D, E and G is cloned with BglII at the forward primer (5' of the gene, promoter's end) and with EcoRI at the reverse primer (gene's 3' end and terminator's beginning).

LUX C will be cloned with XhoI at the forward primer (5' of the gene, promoter's end), since there is a BglII restriction site within native LUX C sequence, and with EcoRI at the reverse primer (gene's 3' end and terminator's beginning). Using same restriction sites with different primer pairs intended to facilitate cloning into standartized pSAT vectors and to simplify overall cloning strategy.

The aforementioned primers are to be used in a polymerase chain reaction to amplify the LUX genes to be cloned, using *Photobacterium leiognathi* genomic DNA as a template. PCR conditions, including melting temperature (T_m , dependent on primer's length and AT/GC content) and elongation time (2 min/kbp for Pfu

polymerase) are specific for each given pair of primers. Proofreading Pfu polymerase must be used to decrease probability of point mutations.

LUX A (Length: 1065bp; SwissProt # P29238)

5 **Forward primer:**

ATGAAAATTAGTAATATCTG (SEQ ID NO: 90)

Length – 20; GC - 4; AT – 16; Tm=53

Addition of BglIII restriction site (in bold):

GAAGATCTATGAAAATTAGTAATATCTG (SEQ ID NO: 91)

10

Reverse primer:

ACTTAAAAGATCCTCAGTAA (SEQ ID NO: 92)

Length – 20; GC – 6; AT – 14; Tm=56

Addition of EcoRI restriction site (in bold):

15 ACTTAAAAGATCCTCAGTAAGA**ATTC**CG (SEQ ID NO: 93)

Reverse sequence to be used as final primer:

CGGAATTCTTACTGAGGATCTTTTAAGT (SEQ ID NO: 94)

PCR parameters:

20 Tm for PCR=54 C.

Elongation time using *Pfu*: 2.5 mins.

LUX B (Length: 981bp; SwissProt #: P29239)

Forward primer:

25 ATGAATTTTCGGGTTATTTTCC (SEQ ID NO: 95)

Length – 22; GC - 7; AT – 15; Tm=61.

Addition of BglIII restriction site (in bold):

GAAGATCTATGAATTTTCGGGTTATTTTCC (SEQ ID NO: 96)

30 **Reverse primer:**

CAAAGATAACCTTATTAAATAA (SEQ ID NO: 97)

Length – 22; GC – 4; AT – 18; Tm=56.

Addition of EcoRI restriction site (in bold):

CAAAGATAACCTTATTAAATAAGAATTCCG (SEQ ID NO: 98)

Reverse sequence to be used as final primer:

CGGAATTCTTATTTAATAAGGTTATCTTTG (SEQ ID NO: 99)

5 PCR parameters:

T_m for PCR=58 C.

Elongation time using *Pfu*: 2 mins.

LUX C (Length: 1437bp; SwissProt #: P29236)

10 Forward primer:

ATGATTAAGAAGATCCCAATGA (SEQ ID NO: 100)

Length – 22; GC - 7; AT – 15; Av. T_m=61.

Addition of XhoI restriction site (in bold):

CCGCTCGAGATGATTAAGAAGATCCCAATGA (SEQ ID NO: 101)

15

Reverse primer:

GTTCTCGTATTTGTACCGTAG (SEQ ID NO: 102)

Length – 22; GC – 10; AT – 12; T_m=66.

Addition of EcoRI restriction site (in bold):

20 GTTCTCGTATTTGTACCGTAGGAATTCCG (SEQ ID NO: 103)

Reverse sequence to be used as final primer:

CGGAATTCTACGGTACAAATACGAGGAAC (SEQ ID NO: 104)

PCR parameters:

25 T_m for PCR=63 C.

Elongation time using *Pfu*: 3 mins.

LUX D (Length: 948bp; SwissProt #: P21309)

Forward primer:

30 ATGGAAAATACACAACATTC (SEQ ID NO: 105)

Length – 20; GC - 6; AT – 14; T_m=56.

Addition of BglII restriction site (in bold):

GAAGATCTATGGAAAATACACAACATTC (SEQ ID NO: 106)

Reverse primer:

GCGTATTGGCTAGTGTCTAA (SEQ ID NO: 107)

Length – 20; GC – 9; AT – 11; Tm=61.

Addition of EcoRI restriction site (in bold):

5 GCGTATTGGCTAGTGTCTA**AGAATTCCG** (SEQ ID NO: 108)

Reverse sequence to be used as final primer:

CGGAATTCTTAGACACTAGCCAATACGC (SEQ ID NO: 109)PCR parameters:

10 Tm for PCR=58 C.

Elongation time using *Pfu*: 2 mins.**LUX E (length: 1122bp; SwissProt #: P29334)**Forward primer:

15 ATGTCAACATTATTAAATATAG (SEQ ID NO: 110)

Length – 22; GC - 4; AT – 18; Tm=56.

Addition of BglII restriction site (in bold):

GAAGATCTATGTCAACATTATTAAATATAG (SEQ ID NO: 111)20 Reverse primer:

AATGGCCACGAGCCTGAAATAA (SEQ ID NO: 112)

Length – 22; GC – 10; AT – 12; Tm=66.

Addition of EcoRI restriction site (in bold):

AATGGCCACGAGCCTGAAATA**AGAATTCCG** (SEQ ID NO: 113)

25 Reverse sequence to be used as final primer:

CGGAATTCTTATTTTCAGGCTCGTGGCCATT (SEQ ID NO: 114)PCR parameters:

Tm for PCR=61 C.

30 Elongation time using *Pfu*: 2.5 mins.**LUX G (length: 705 bp; SwissProt #: P29237)**Forward primer:

ATGATTTTAAATTGCAAGGTAA (SEQ ID NO: 115)

Length – 23; GC - 5; AT – 18; Tm=59.

Addition of BglII restriction site (in bold):

GAAGATCTATGATTTTAAATTGCAAGGTAA (SEQ ID NO: 116)

5 Reverse primer:

GATGCATTTAGCTATACGTAA (SEQ ID NO: 117)

Length – 21; GC – 7; AT – 15; Tm=59.

Addition of EcoRI restriction site (in bold):

GATGCATTTAGCTATACGTAAGAATT**CCG** (SEQ ID NO: 118)

10 Reverse sequence to be used as final primer:

CGGAATTCTTACGTATAGCTAAATGCATC (SEQ ID NO: 119)

PCR parameters:

Tm for PCR=59 C.

15 Elongation time using *Pfu*: 1.5 mins.

Primers' Summary:

<i>Primer's name</i>	<i>Sequence</i>
LUX A fw	<u>GAAGATCT</u> ATGAAAATTAGTAATATCTG (SEQ ID NO: 91)
LUX A rev	<u>CGGAATTCT</u> TACTGAGGATCTTTTAAGT (SEQ ID NO: 94)
LUX B fw	<u>GAAGATCT</u> ATGAATTCGGGTATTTTTCC (SEQ ID NO: 96)
LUX B rev	<u>CGGAATTCT</u> TATTTAATAAGGTATCTTTG (SEQ ID NO: 99)
LUX C fw	<u>CCGCTCG</u> AGATGATTAAGAAGATCCCAATGA (SEQ ID NO: 101)
LUX C rev	<u>CGGAATTC</u> TACGGTACAAATACGAGGAAC (SEQ ID NO: 104)
LUX D fw	<u>GAAGATCT</u> ATGGAAAATACACAACATTC (SEQ ID NO: 107)

	106)
LUX D rev	<u>CGGAATTCT</u> TAGACACTAGCCAATACGC (SEQ ID NO: 109)
LUX E fw	<u>GAAGATCT</u> ATGTCAACATTATTAAATATAG (SEQ ID NO: 111)
LUX E rev	<u>CGGAATTCT</u> TATTTTCAGGCTCGTGGCCATT (SEQ ID NO: 114)
LUX G fw	<u>GAAGATCT</u> ATGATTTTAAATTGCAAGGTAA (SEQ ID NO: 116)
LUX G rev	<u>CGGAATTCT</u> TACGTATAGCTAAATGCATC (SEQ ID NO: 119)

Example 11: Generation of transgenic plant lines

Preparation of transgenic LUX genes expressing plants involve the following steps: (A) Preparation of gold particles coated with the plant transformation vectors, for either nuclear or chloroplast transformation, as shown in Figures 9 and 10, respectively; (B) Biolistic delivery of the DNA into plant tissue; (C) Placing transformed plant tissue on regeneration medium, containing appropriate antibiotics and hormones, and (D) Regenerated transgenic plants appear within several weeks. (adopted from: <http://www.ag.usask.ca>)

Example 12: Generation of transgenic tobacco plants

Transgenic tobacco plants expressing free DsRed and GFP have been generated using tobacco leaf disk transformation method and selection on kanamycin supplemented media. The use of pPZP-RCS2 binary vector with subcloned GFP, DsRed and ntpII cassettes, derived from pSAT vectors (Tzfira et al, 2005), resulted in high expression levels of both markers in certain progeny plants selected on kanamycin supplemented media (Krichevsky et al, unpublished). See figure 11.

Example 13: Remote activation of an inducible promoter in-planta.

Gal4 DNA binding domain (mGal4) fused to transcriptional activator from the Herpes Simplex Virus (VP16) was cloned under constitutive 35S promoter into

transient plant expression vector pRTL2 and biolistically delivered into leaves of an Arabidopsis plant carrying UAS driven, genome-integrated GUS reporter gene. Transient expression of mGal4-VP16 resulted in activation of GUS expression and was detected by histochemical staining. See Figure 12 (A) and (C). Specificity of UAS promoter activation has been confirmed when similar biolistic delivery has been made into a wild-type Arabidopsis plants, carrying no UAS driven GUS gene. See Figure 12 C.

Example 14: LUX operons alignment.

Alignment of LUX operons from three major luminescent bacteria families, namely *Photobacterium*, *Vibrio* and *Xenorhabdus* (aka *Photorhabdus*). A sequence comparison was produced using ClustalW software (EBI) of three luminescent bacteria species: *Photobacterium leiognathi* (abbreviated as Pl), *Vibrio harveyi* (abbreviated as Vh) and *Xenorhabdus* (aka *Photorhabdus*) *luminescens* (abbreviated as Xl). See Figure 14. These 3 species have been selected because they are representative example of the largest and most prominent luminescent bacteria families, namely *Photobacterium* *Vibrio* and *Xenorhabdus* (aka *Photorhabdus*).

Conserved nucleotides shown by asterisks. Accordingly, variant nucleotide sequences of the LUX operon preferably do not include a variation of the conserved nucleotides shown in asterisks. Overall conservation between all three families approximates 60%, showing high degree of homology between luminescent bacteria from the three major luminescent bacteria families and suggesting strong conservation of the LUX operon amongst various luminescent bacteria species.

Primers for cloning LUX operons of these 3 bacterial species were designed. Sequence-specific part of the forward primer for individual gene normally starts from gene's first codon, namely ATG. In case of the whole operon cloning, the forward primer will start approx. 15-20 bp downstream of the first ATG of the first gene in the operon in order to include ribosome binding site allowing transcription of the first gene in the operon (each consecutive gene in the operon has its own ribosome binding site preceding its ATG). Further, sequence-specific part of the reverse primer for individual gene normally ends with this particular gene's stop codon (TAA, TAG, etc). Sequence specific primer aimed to amplify the whole operon ends with the last

stop codon of the last gene in the operon, allowing to amplify by PCR and further subclone whole operon as a single unit into an expression vector (i.e. pCas/pUniPlast).

The following are examples of primers designed to clone the LUX operon.

5 **Full LUX operon sequence for *Photobacterium leiognathi*, GeneBank # M63594:**

Primers and cloning strategy: This LUX operon will be cloned as EcoRI PCR fragment (restriction site underlined) directly into the same sites of a pCas vector and further subcloned into pUniPlast for making of transgenic plants.

10

Pl forward primer:

ACAGAATTCCCAAAGGAGATTACATGATTAAG (SEQ ID NO: 120)

Pl reverse primer:

15 **CTGATGCATTTAGCTATACGTAAGAATTCCAA** (SEQ ID NO: 121)

Final reverse primer:

TTGGAATTCTTACGTATAGCTAAATGCATCAG (SEQ ID NO: 122)

PCR conditions:

20 Primers T_m = over 60.

Expected fragment: approx. 6,550 bp.

25

Full LUX operon sequence for *Vibrio harveyi*, GeneBank # EU192082:

Primers and cloning strategy: This LUX operon will be cloned as XhoI PCR fragment directly into the same sites of a pCas vector and further subcloned into pUniPlast for making of transgenic plants.

30

Vh forward primer:

AACTCGAGCAAAAGAGAAGCTCTTGATATGG (SEQ ID NO: 123)

Vh reverse primer:

GTGTGAAGTGAGTTGGTCTTAGCTCGAGAA (SEQ ID NO: 124)

Final reverse primer:

TTCTCGAGCTAAGACCAACTCACTTCACAC (SEQ ID NO: 125)

5

PCR conditions:Primers T_m = over 60 C.

Expected fragment: approx. 7,130 bp.

10

Full LUX operon sequence for *Xenorhabdus* (aka *Photorhabdus*) *luminescens*,**GeneBank # AF403784:****Primers and cloning strategy:** This LUX operon will be cloned as BamHI PCR

15 fragment (restriction site underlined) directly into the same sites of a pCas vector and further subcloned into pUniPlast for making of transgenic plants.

XI forward primer:

ATGGATCCCATTTAAATGGATGGCTAATATGAC (SEQ ID NO: 126)

20

XI reverse primer:

CTTAAACCAAGCATTTAATAGTTGAGGATCCAT (SEQ ID NO: 127)

Final reverse primer:

ATGGATCCTCAACTATTAAATGCTTGTTTAAAG (SEQ ID NO: 128)

25

PCR conditions:Primers T_m = over 60 C.

Expected fragment: approx. 5,820 bp.

30

PRIMERS' SUMMARY

<u>Primer's name</u>	<u>Sequence</u>
Pl forward primer	ACAGAA <u>TTCC</u> CAAAGGAGATTACATGATTAAG (SEQ ID NO: 120)

Pl reverse primer	TTGGAATTCTTACGTATAGCTAAATGCATCAG (SEQ ID NO: 122)
Vh forward primer	AACTCGAGCAAAAGAGAAGCTCTTGATATGG (SEQ ID NO: 123)
Vh reverse primer	TTCTCGAGCTAAGACCAACTCACTTCACAC (SEQ ID NO: 125)
Xl forward primer	ATGGATCCCATTAATGGATGGCTAATATGAC (SEQ ID NO: 126)
Xl reverse primer	ATGGATCCTCAACTATTAAATGCTTGGTTTAAG (SEQ ID NO: 128)

What is claimed is:

1. A transgenic bioluminescent plant comprising:

 an expressible heterologous nucleotide sequence comprising a bacterial LUX operon, comprising LUX A, LUX B, LUX C, LUX D, LUX E, and LUX G genes,

 wherein the heterologous nucleotide sequence is expressed to render the plant bioluminescent.
2. The plant of claim 1, wherein the heterologous nucleotide sequence is placed in a single vector.
3. The plant of claim 2, wherein the vector comprises a rare cutting enzyme sequence.
4. The plant of claim 3, wherein the vector is a chloroplast transformation vector.
5. The plant of claim 3, wherein the vector is a binary vector.
6. The plant of claim 1, wherein the heterologous nucleotide sequence further comprises a LUX H gene.
7. The plant of claim 1, wherein the heterologous nucleotide sequence further comprises a LUX F gene.
8. The plant of claim 1, wherein the heterologous nucleotide sequence further comprises a LUX Y gene.
9. The plant of claim 1, wherein the heterologous nucleotide sequence further comprises a plastid targeting sequence.
10. The plant of claim 1, wherein the heterologous nucleotide sequence further comprises at least one gene encoding a cofactor for enhancing bioluminescence.
11. The plant of claim 10, wherein the cofactor comprises a polypeptide encoded by a RIB operon.
12. The plant of claim 10, wherein the cofactor comprises a bacterial or plant acyl carrier protein.
13. The plant of claim 10, wherein the cofactor comprises an FRE flavin reductase enzyme.
14. The plant of claim 10, wherein the heterologous nucleotide sequence further comprises an inducible promoter.
15. The plant of claim 14, further comprising a second heterologous nucleotide sequence comprising a gene encoding a tissue-specific promoter and a gene encoding a transcription factor.

16. The plant of claim 15, wherein the second heterologous nucleotide sequence further comprises a plastid targeting sequence.
17. The plant of claim 15, wherein the second heterologous nucleotide sequence further comprises a gene encoding a fluorescent protein.
18. A method for making a transgenic bioluminescent plant comprising: transfecting at least one plant cell with at least one vector comprising a gene encoding for a bacterial lux operon; growing said at least one plant cell into a mature plant; and providing a means for regulation of expression of said gene.
19. The method of claim 18, wherein the means for regulating expression comprises a promoter.
20. A method for producing a transgenic bioluminescent plant comprising transfecting a plant cell with a vector comprising an expressible heterologous nucleotide sequence comprising a bacterial LUX operon, comprising LUX A, LUX B, LUX C, LUX D, LUX E, and LUX G genes.
21. The method of claim 20, wherein the vector comprises rare cutting enzyme sequences.
22. The method of claim 21, wherein the vector comprises a plastid transformation vector.
23. The method of claim 21, wherein the vector comprises a binary vector.
24. The method of claim 20, wherein the heterologous nucleotide sequence further comprises a LUX H gene.
25. The method of claim 20, wherein the heterologous nucleotide sequence further comprises a LUX F gene.
26. The method of claim 20, wherein the heterologous nucleotide sequence further comprises a LUX Y gene.
27. The method of claim 20, wherein the heterologous nucleotide sequence is integrated in nuclear DNA of the plant cell.
28. The method of claim 27, wherein the heterologous nucleotide sequence further comprises a chloroplast targeting sequence.
29. The method of claim 20, wherein the heterologous nucleotide sequence is integrated in plastid DNA of the plant cell.
30. The method of claim 20, wherein the heterologous nucleotide sequence further comprises at least one gene encoding a cofactor for enhancing bioluminescence.
31. The method of claim 30, wherein the cofactor comprises a polypeptide encoded by a RIB operon.

32. The method of claim 30, wherein cofactor comprises bacterial or plant acyl carrier protein.

33. A vector for transforming a plastid, comprising:

- a. a first and a second homologous recombination site, wherein each of said homologous recombination site is flanked by a unique rare cutting enzyme sequence; and
- b. a multiple cloning site comprising at least one unique rare cutting enzyme sequences.

34. The vector of claim 33, wherein each of said homologous recombination site is flanked by two unique rare cutting enzyme sequences.

35. The vector of claim 33, wherein the multiple cloning site comprises a lux operon.

36. The vector of claim 35, wherein the multiple cloning site further comprises at least one gene encoding a cofactor for enhancing bioluminescence.

37. A method for producing a vector for transforming a plastid comprising:

- a. providing an acceptor vector comprising (i) a first and a second homologous recombination site, wherein each of said homologous recombination site is flanked by a unique rare cutting enzyme sequence; and (ii) a multiple cloning site comprising at least one unique rare cutting enzyme sequence;
- b. providing at least one donor vector, wherein the donor vector comprises a heterologous nucleotide sequence flanked by a rare cutting enzyme sequence that is sufficiently identical to the at least one unique rare cutting enzyme sequence of the multiple cloning site of the acceptor vector; and
- c. sequentially subcloning the heterologous nucleotide sequence from the donor vector into the acceptor vector,

wherein a vector for transforming a plastid is produced.

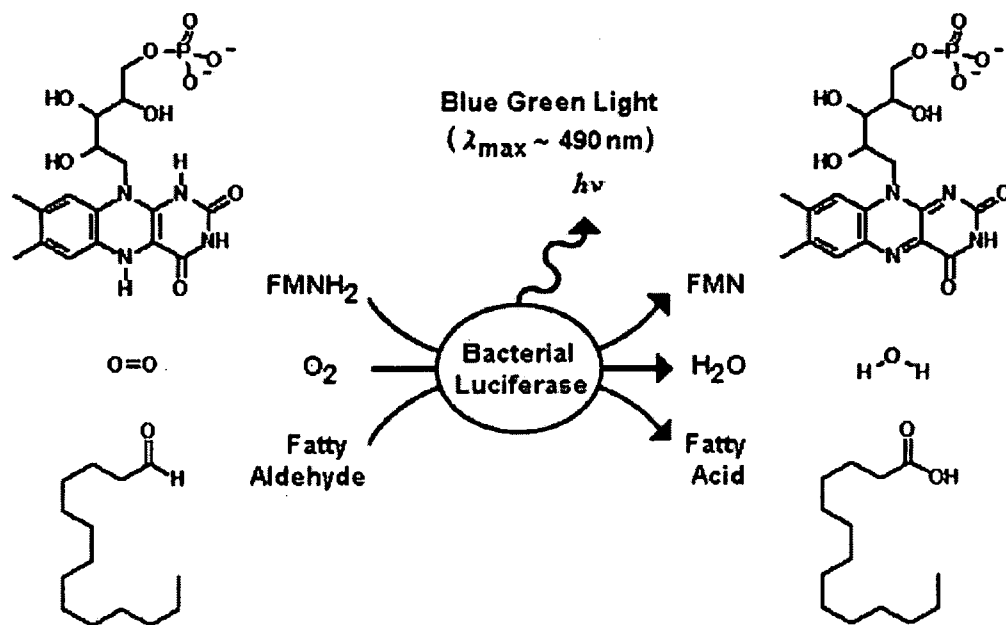
38. A vector system kit comprising:

- a. an acceptor vector comprising (i) a first and a second homologous recombination site, wherein each of said homologous recombination site is flanked by a unique rare cutting enzyme sequence; and (ii) a multiple cloning site comprising at least one unique rare cutting enzyme sequence; and

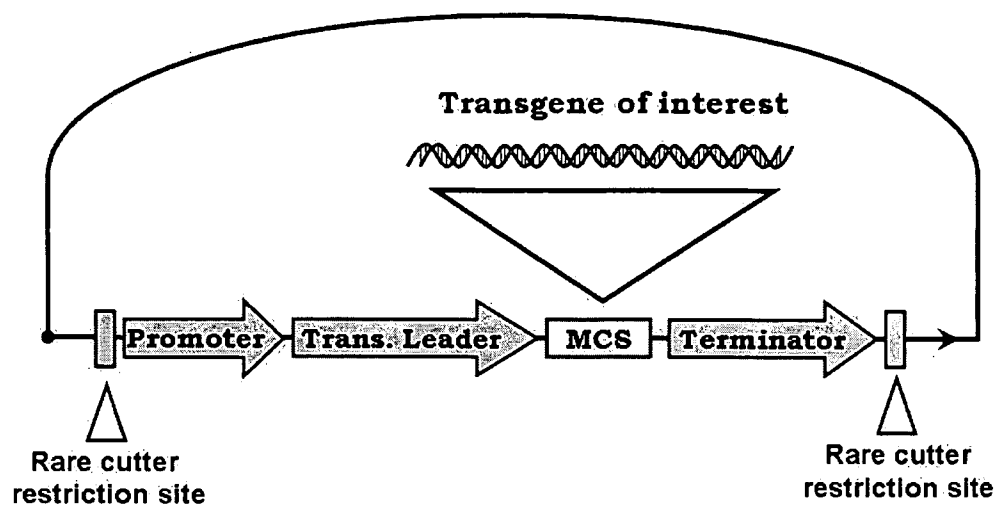
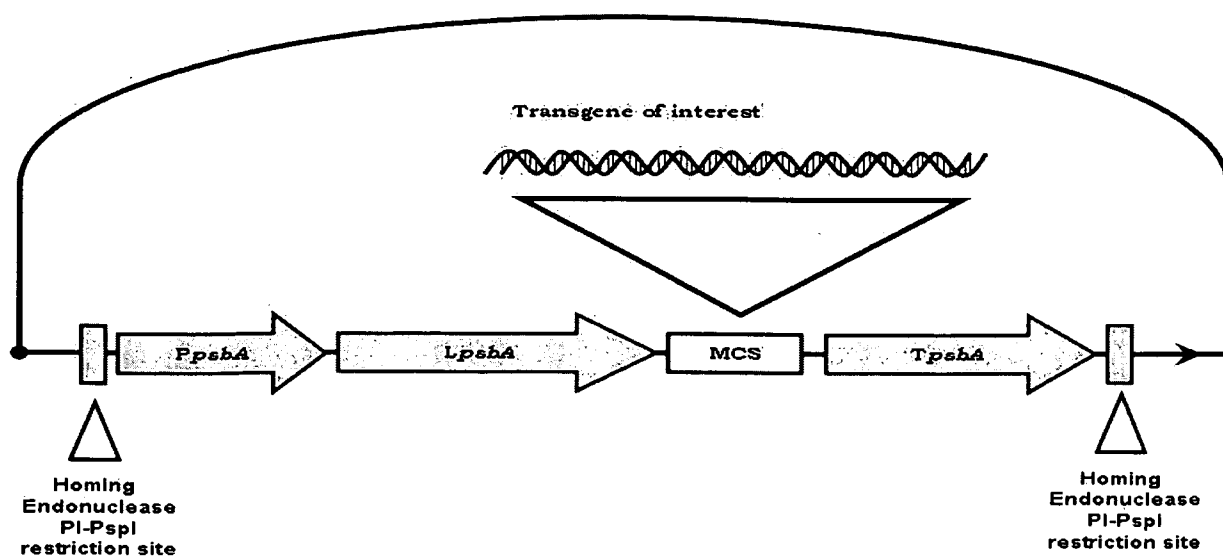
- b. at least one donor vector, wherein the donor vector comprises a heterologous nucleotide sequence flanked by a rare cutting enzyme sequence that is sufficiently identical to the at least one unique rare cutting enzyme sequence of the multiple cloning site of the acceptor vector.
- 39. The vector system kit of claim 38, wherein the heterologous nucleotide sequence further comprises a regulatory sequence.
- 40. The vector system kit of claim 38, wherein the heterologous nucleotide sequence further comprises a selection marker sequence.

1 / 22

FIG. 1

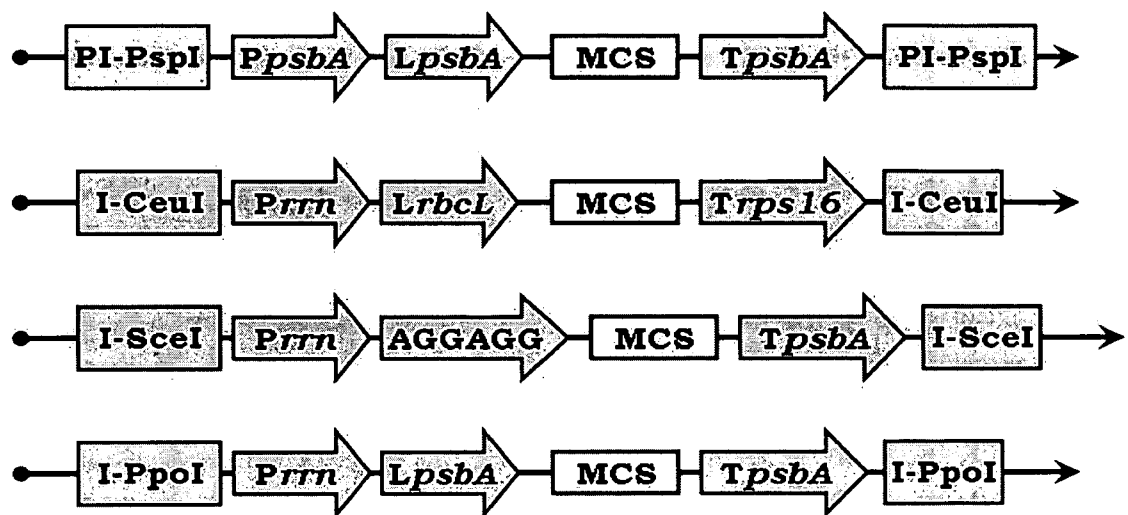


2 / 22

FIG. 2
A**FIG. 2 B**

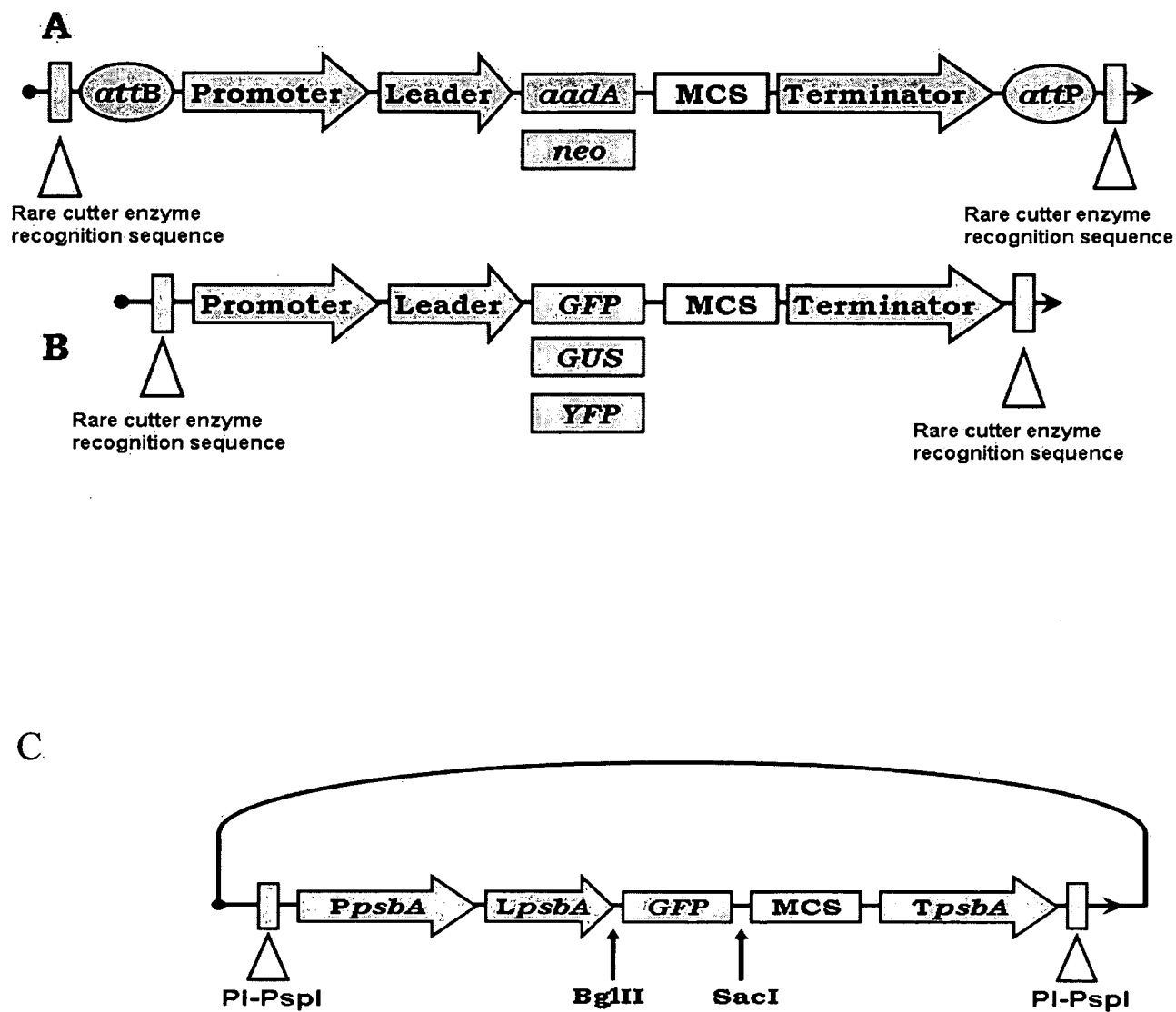
3 / 22

FIG. 3



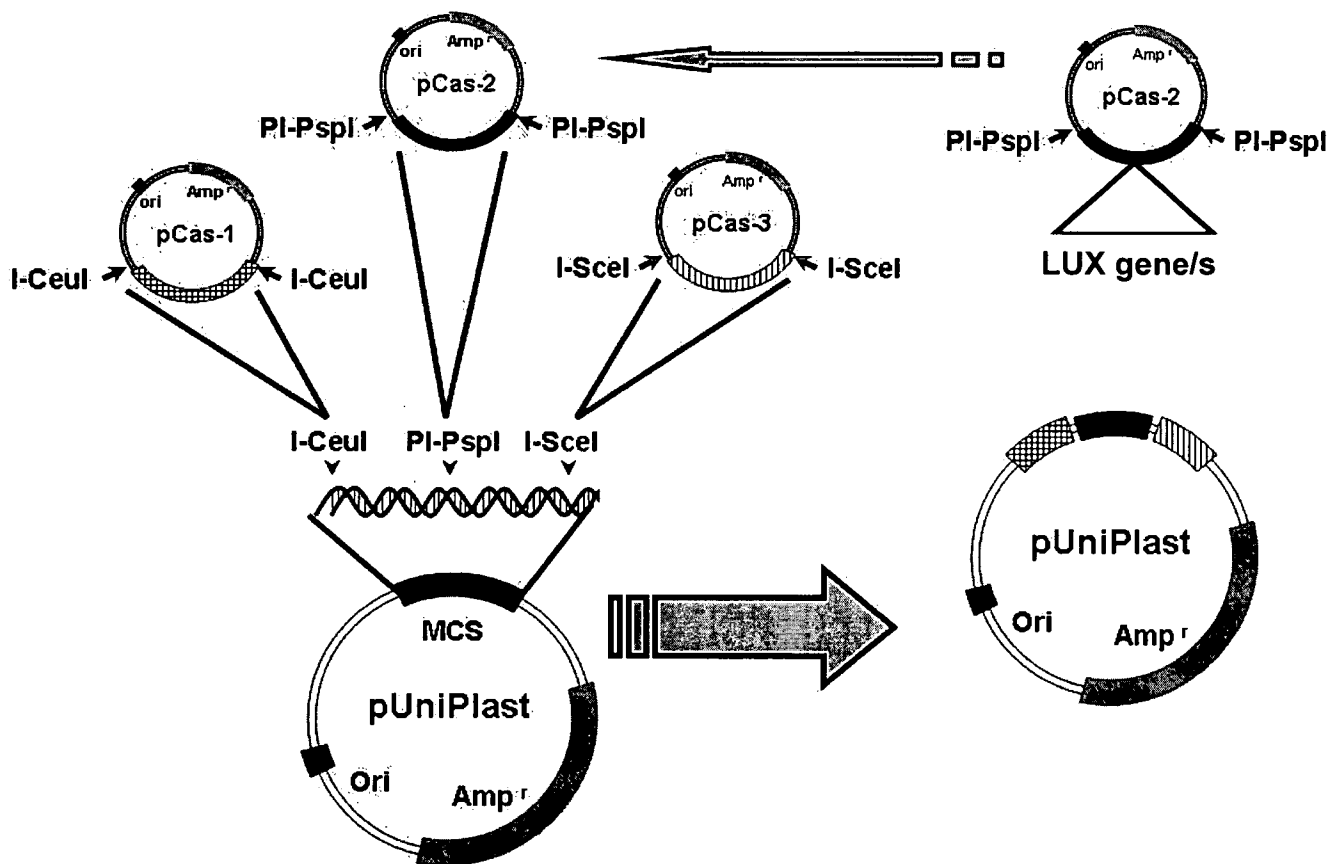
4 / 22

FIG. 4



5 / 22

FIG. 5



6 /22

FIG. 6

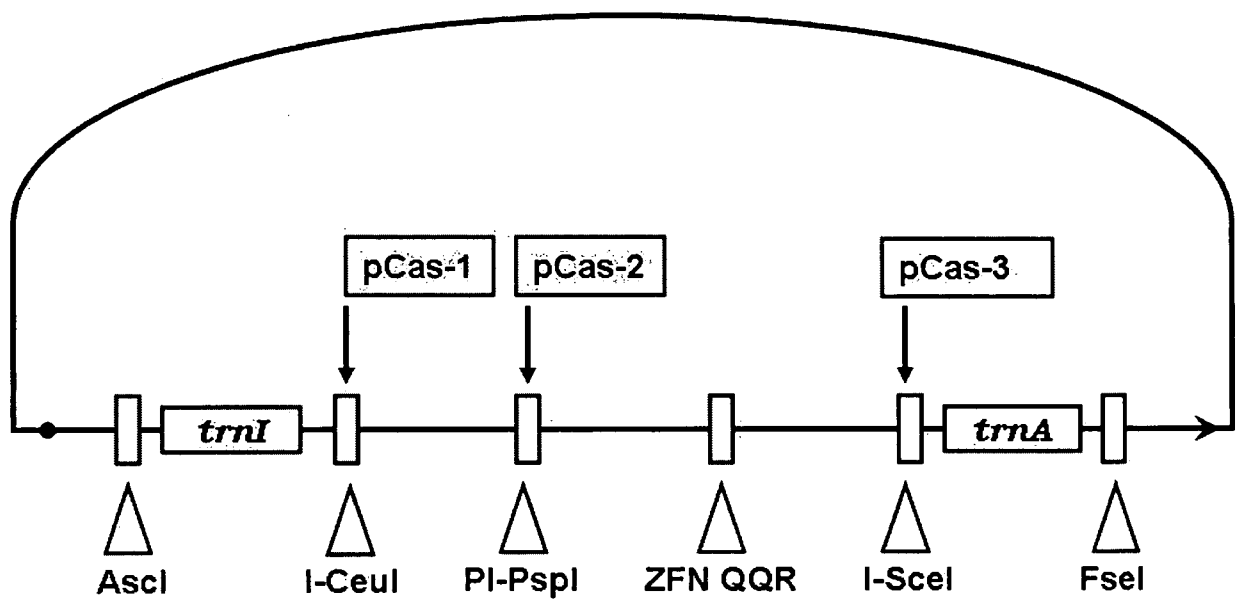


FIG. 7

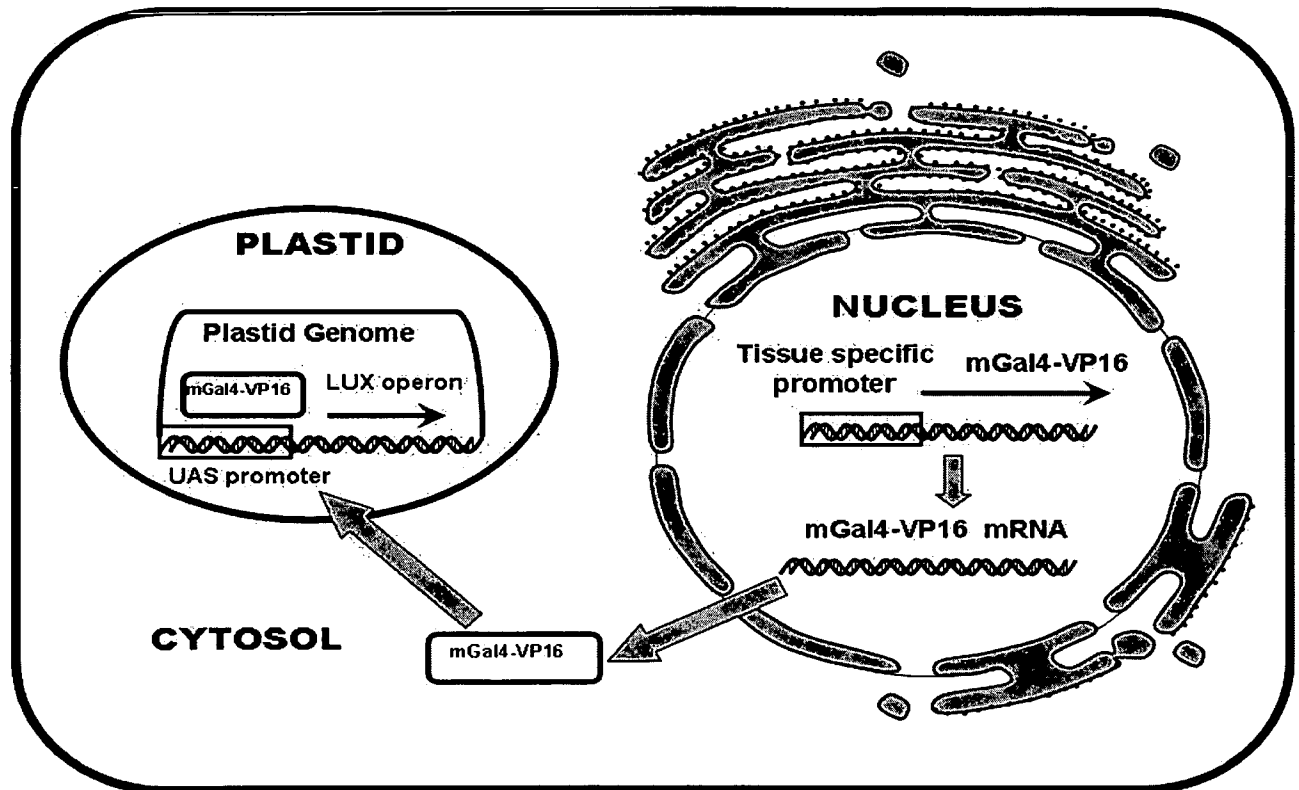
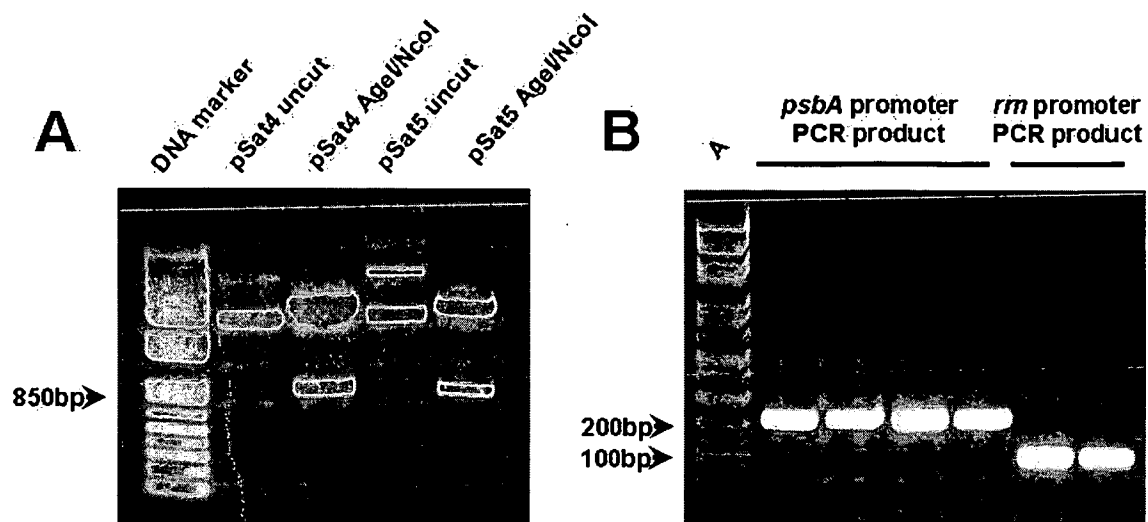
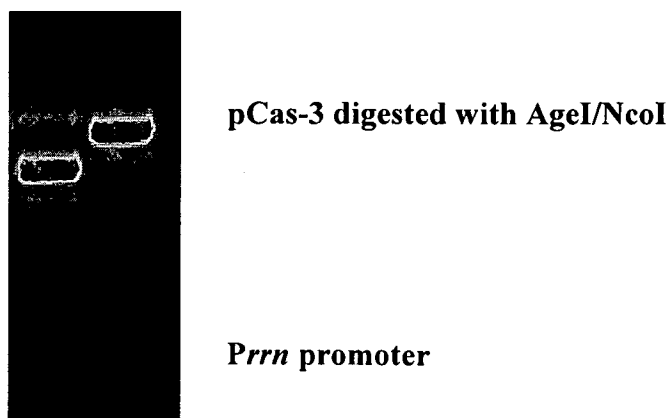


FIG. 8



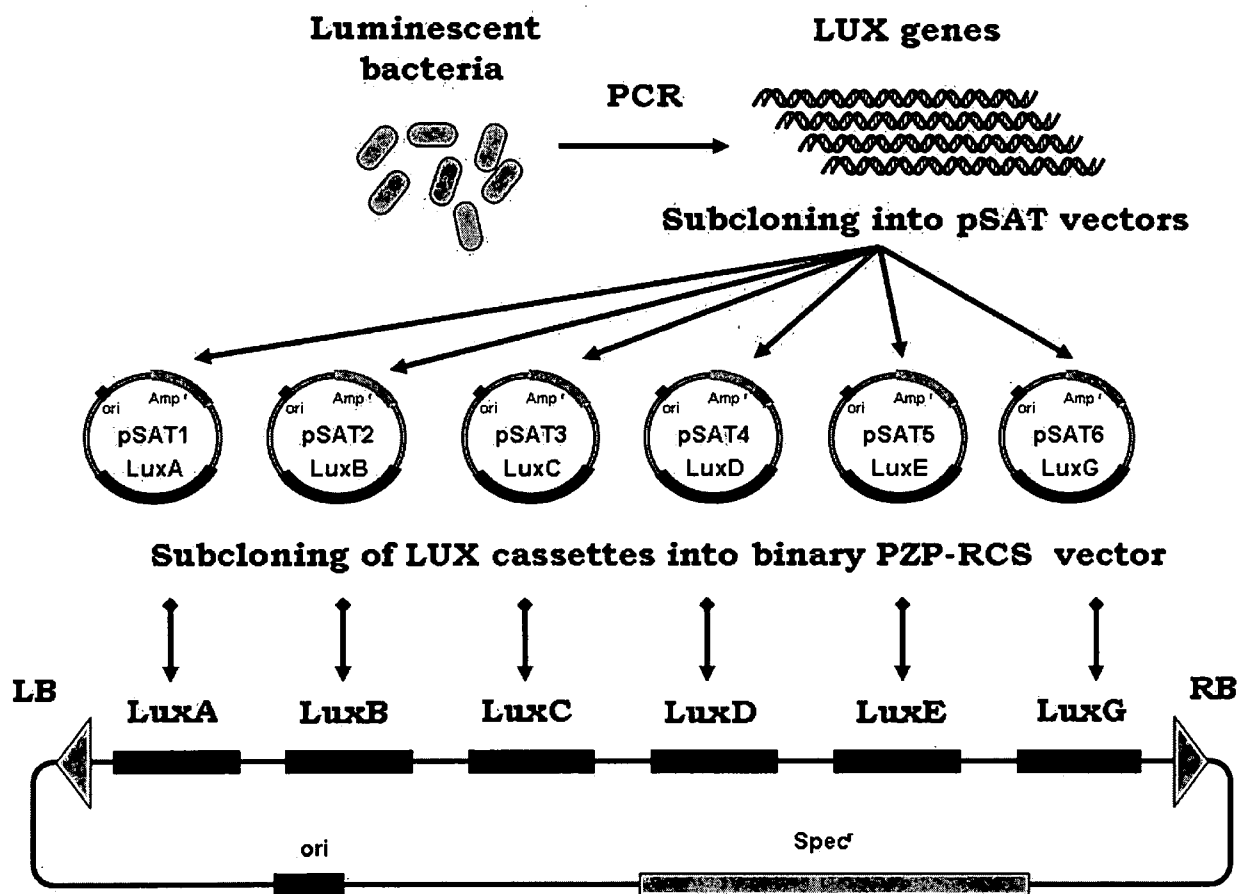
1 2

FIG. 8 C



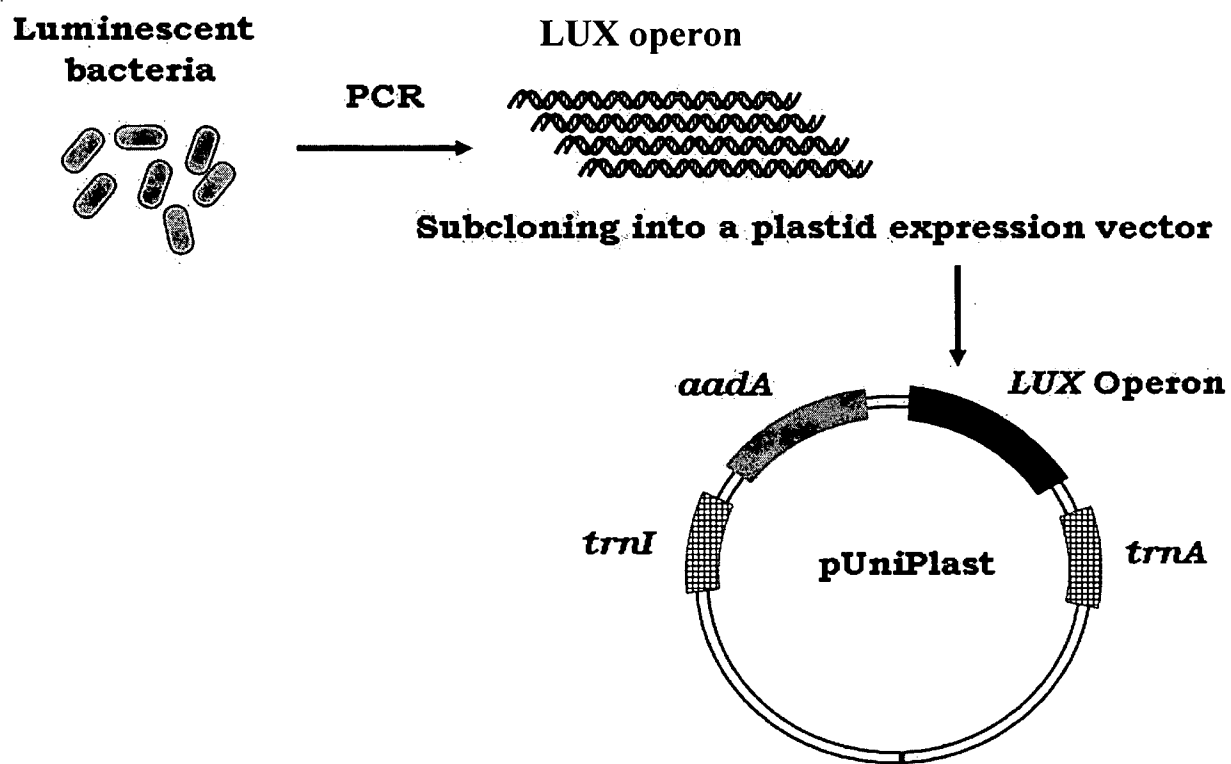
9 / 22

FIG. 9



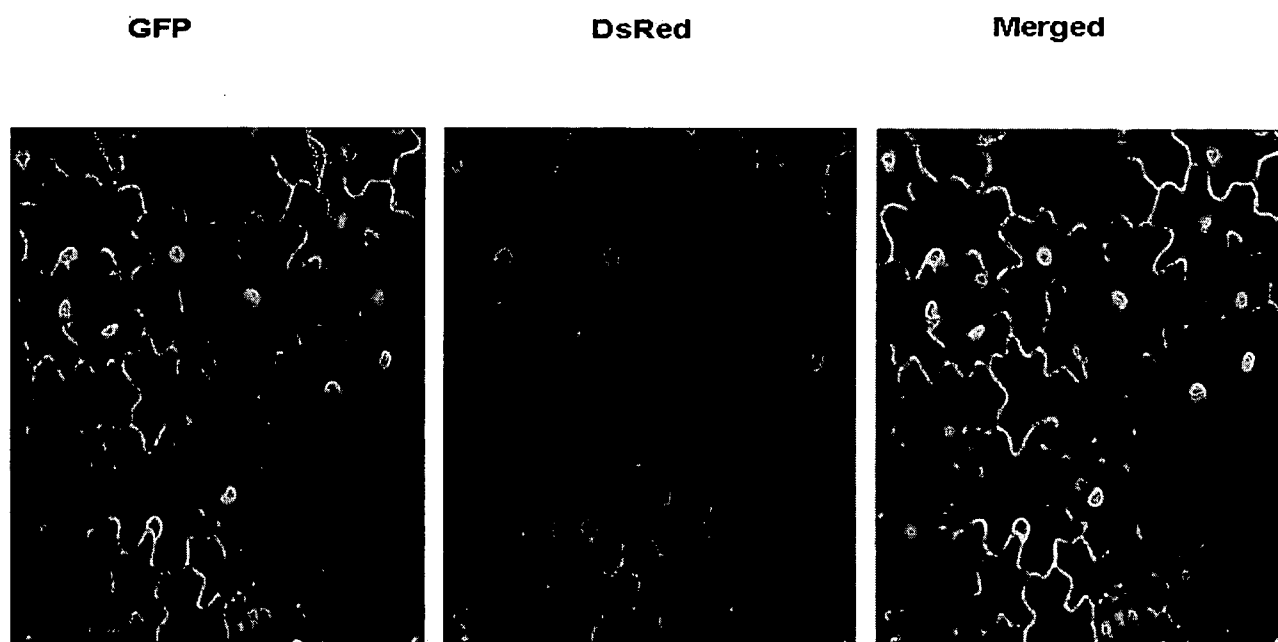
10 /22

FIG. 10



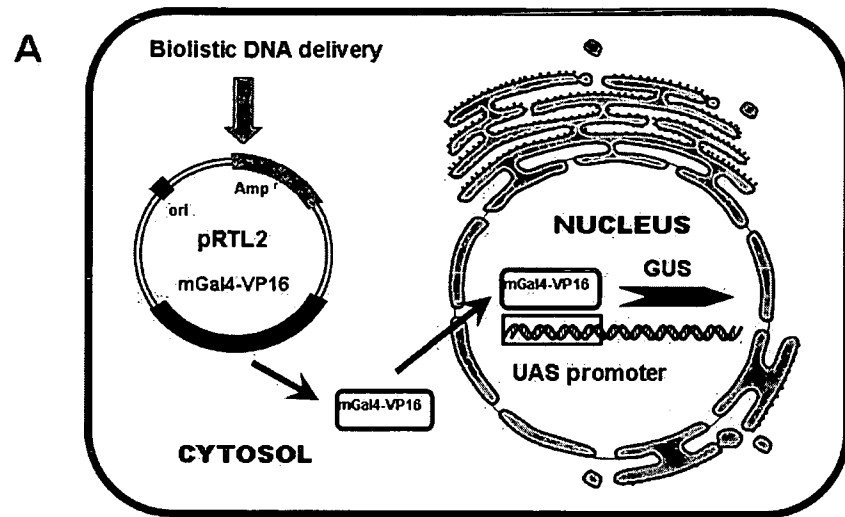
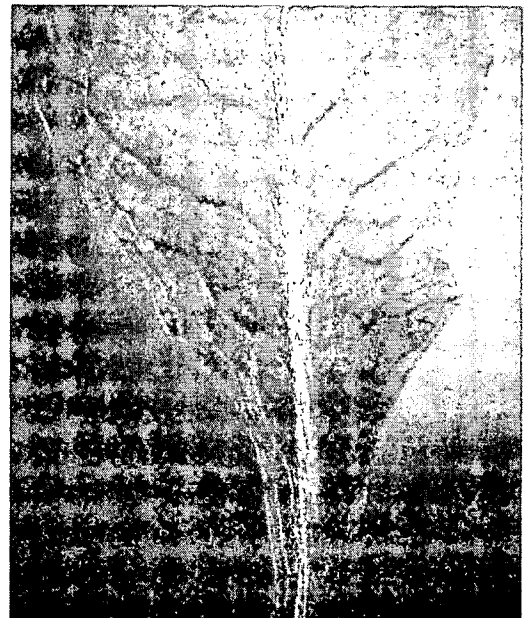
11 /22

FIG. 11



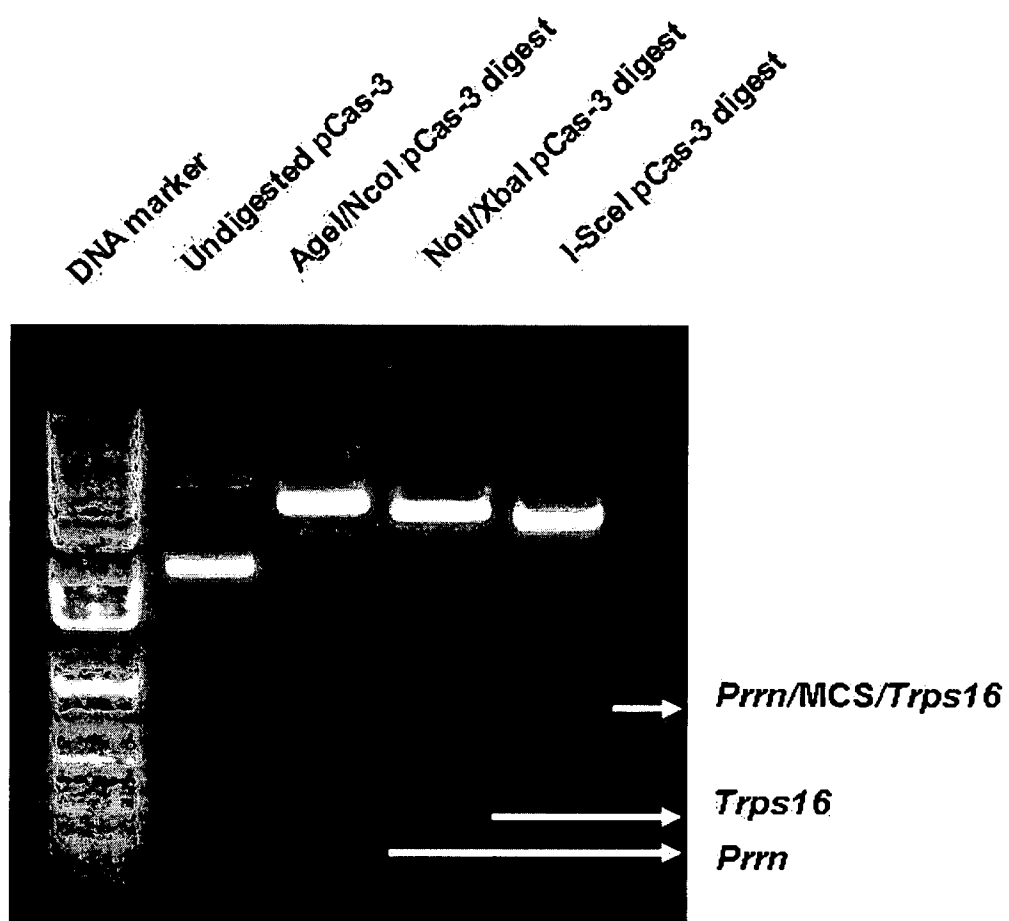
12 / 22

FIG. 12

**B****C**

13 / 22

FIG. 13



14 / 22

FIG. 14

```

Vh      CAAAAGAGAAGCTCTTGATATGGAAAAACACTTACCTTTAATAATAAATGGGCAAAT--- 57
Xl      CATTAAATGGATGGCTAATATGACTAAAAAAATTCATTCAATTATTAACGGCCAGGTGA 60
Pl      --CCAAAGGAG---ATTACATGATTAAGAAGATCCCAATGATTATGGGGGTGTAGTTCA 55
          * *      * * * *      * * * *      * * * *      * * * *      * * * *

Vh      AGTTTCTACTAAAGAAAATCGATTGAG-----ATCAGTTTGAAGAAAAAAAAGTTAA 111
Xl      AATTTTCCCGAAAGTGATGATTAGTGCAATCCATTAAATTTTGGTGATAATAGTGTTA 120
Pl      AAACACGCTCTGGATATGGCATGCGTGAAC-----TAACGCTCAACAATAATAAAGTGAA 109
          *      *      *      *      *      *      *      *      *      *      *

Vh      GATTGATTCCTTTAATCATTTTCATTTAATCCAGATGGTGACTCATGATTATTTAAATGA 171
Xl      CCTGCCAATATTGAATAATTTCTCATGTAAAAAACATTATTGATTATAATGAAAATAATAA 180
Pl      TATCCCTATCATCACCCAAAGTGATGTTGAAGCTATTCAATCACTAAATATAGAAAACAA 169
          *      *      *      *      *      *      *      *      *      *      *

Vh      TCTAAATATTAATAACATCATCAATTTTCTTTATACAACGGGGCAGCGTTGGAAGAGCGA 231
Xl      ATTACGGTTGCATAATATTGTCAATTTTCTCTATACGGTAGGGCAAAGATGGAAAAATGA 240
Pl      ATTGACTATAAATCAGATAGTTAATTTCTTATATACAGTGGGACAAAAATGGAAGAGCGA 229
          *      *      *      *      *      *      *      *      *      *      *

Vh      AGAATATTCAAGAAGAAGGGCATATATTCGGTCTCTTATTACTTATTTGGGGTATTCACC 291
Xl      AGAATATTCAAGACGCAGGACATACATTTCGTGATTTAAAAAATATATGGGATATTCAGA 300
Pl      AACTTACAGCCGACGACTCACTTATATTCGAGATCTTATTAAGTTCTCTCGGTTACTCACA 289
          *      *      *      *      *      *      *      *      *      *      *

Vh      ACAAATGGCGAAACTAGAAGCAAATTGGATTGCAATGATCCTTTGCTCTAAGAGTGCCT 351
Xl      AGCAATGGCCAGTTAGAGGCCAACTGGATATCTATGATTTTATGTTCTAAAGGTGGCCT 360
Pl      AGAGATGGCAAAACTTGAAGCTAACTGGATCTCAATGATTCTGTGTAGCAAAAGTGCCTT 349
          *      *      *      *      *      *      *      *      *      *      *

Vh      CTACGACATTATTGATACCGAGCTTGGCTCAACGCATATACAAGATGAATGGCTACCGCA 411
Xl      TTATGATGTTGTAGAAAATGAACCTTGGTTCTCGCCATATCATGGATGAATGGCTACCTCA 420
Pl      GTACGATATTGTTGAGAATGATCTTAGCTCACGGCATATTATTGATGAGTGGATCCCCCA 409
          *      *      *      *      *      *      *      *      *      *      *

Vh      GGGTGAGTGTTATGTGAGGGCTTTTCTTAAAGGACGCACGATGCATTTGCTTGCGGGGAA 471
Xl      GGATGAAAGTTATATTAAAGGCTTTTCCGAAAGGTAAGTCTATACATCTGTTGGCAGGTAA 480
Pl      AGGTGAATGTTATGTCAAAGCGCTCCCAAAAGGAAAATCTGTACACCTATTAGCTGGTAA 469
          *      *      *      *      *      *      *      *      *      *      *

Vh      TGTTCTCTCTCCGGTGTGACCTCAATACTACGAGGCATACTGACGAGAAATCAATGTAT 531
Xl      TGTTCCATTATCTGTGATCATGTCTATATTACGCGCAATTTTAACCAAGAATCAGTGTAT 540
Pl      CGTACCACTATCTGGTGTGACTTCTATTCTTCGTGCGATTTTGACCAAAAACGAGTGCAT 529
          *      *      *      *      *      *      *      *      *      *      *

Vh      TGTGAGAATGTCAGCATCGGATCCTTTTACTGCCACGCGCTAGCGATGAGCTTTATTGA 591
Xl      TATAAAAACATCGTCAACCGATCCCTTTACCGCTAATGCATTAGCGTTAAGCTTTATCGA 600
Pl      CATAAAAACGTCATCAGCTGATCCTTTTACAGCTACTGCGCTAGTTAATAGTTTATCGA 589
          *      *      *      *      *      *      *      *      *      *      *

Vh      CGTCGATCCGAATCATCCAATTTCTCGTTCTATCTCCGTATTGTATTGGCCTCATGCATC 651
Xl      TGTAGACCCTAATCATCCGATAACGCGCTCTTTGTCTGTTGTATATTGGCCACACCAAGG 660
Pl      TGTAGATGCAGAACACCCGATCACAGTTCAATCTCAGTTATGTATTGGTCACATAGCGA 649
          *      *      *      *      *      *      *      *      *      *      *

Vh      GGATACGACACTCGCTGAAGAGTTACTCAGTCATATGGATGCAGTGGTTGCTTGGGGGGG 711
Xl      TGATACATCACTCGCAAAAGAAATTATGCAACATATGGATGTTATTGTCGCTTGGGGAGG 720
Pl      GGATCTTGCTATTTCAAAACAAATAATGAGCTGTGCTGATGTGGTTATTGCATGGGGTGG 709
          *      *      *      *      *      *      *      *      *      *      *

Vh      GGCGGGATGCCATTGATTGGGCGGTTAAGCATTCTCCTTCACATATCGATGTTTGAAGT 771
Xl      GG-AAAGATGCGATTAATTGGGCTGTAGAACATGCACCACCTATGCTGACGTGATTAAAT 779

```

15 / 22

Pl TG-ATGATGCAATTAAATGGGCTACAGAACATGCACCATCACACGCAGATATTCTAAAAT 768
* * * * *

Vh TTGGTCCAAAGAAGAGTTTTACCGTGTAGACCATCCAGCCGATCTAGAAGAAGCCGCT 831
Xl TTGGCTCTAAAAGAGTTTTGTCATTATTGATAATCCAGTTGATTTAACGTGAGCAGCTA 839
Pl TTGGTCCCAAAAAGAGTATATCCATTGTTGACAACCCAACAGATATTAAGGCTGCTGCTA 828
* * * * *

Vh CGGGTGTTGCCCATGATATTGCTTTTATGACCAAAATGCCTGCTTTTCTACTCAGAATA 891
Xl CCGGTGCGGCTCATGATATTGTTTTTACGATCAGCGCGCTTGTTTTTCTGCCCAAAACA 899
Pl TCGGTGTAGCACATGATATCTGTTTTTACGATCAGCAAGCATGTTTCTCCACCCAAGATA 888
* * * * *

Vh TTTACTTTTCTGAGATAAGTATGAAGAATTTAAATTTAAACTTGTGAAAACTGAATC 951
Xl TATATTACATGGGAAATCAGTATGAGGAATTTAAGTTAGCGTTGATAGAAAACTTAATC 959
Pl TTTATTATATTGGCGATAGCATAGACATATTTTGTGATGAATTAGCTCAGCAATTAATA 948
* * * * *

Vh TCTATCAAGAAGTTTTACCAAAATCAAAACAAAGTTTTGATGACGAAGCTTTATTTTCTA 1011
Xl TATATGCGCATATATTACCAACGCCAAAAAGATTTTGATGAAAAGCGGCCATTCTT 1019
Pl AATATAAAGACATATTGCCTAAAGGTGAGCGGAATTTTGATGAAAAGCAGCTTTTCTT 1008
* * * * *

Vh TGACTCGTCTTGAGTGTCAATTTTCTGGGTTGAAAGTTATATCAGAACCGGAAAATAACT 1071
Xl TAGTCCAAAAAGAGAGCTTATTTGCTGGATTAAAAGTAGAGGTGGATGTTTCATCAACGTT 1079
Pl TAACGGAAAGAGAATGTTTGTGTTGCCAAATATAAAGTTCAAAAAGGTGAAAGCCAATCTT 1068
* * * * *

Vh GGATGATCATCGAGTCAGAGCCCGGGTTGAATATAACCATCCATTAAGTCGTTGCGTTT 1131
Xl GGATGATTATTGAGTCAAATGCGGGTGTGGAATTTAATCAACCACTTGGCAGATGTGTGT 1139
Pl GGTATTAAACGCAATCACCTGCGGGATCATTTGGTAATCAGCCGTTATCAGCTCGGCTT 1128
* * * * *

Vh ATGTCCACAAAATAAATAAGGTTGATGATGTTGTTGAATATATAGAAAAACATCAAACAC 1191
Xl ATCTTCATCAGTCGATAATATTGAGCAAGTATTGCCTTATGTTCAAAAAATAAGACAC 1199
Pl ATATTCAATCAAGTAAATGACATTTCAGAAGTCATTCCATTGTCATAAGGCGGTAACGC 1188
* * * * *

Vh AAACGATTTCTTTTATCCATGGGAATCTTCCAAGAAATATCGAGATGCATTGCGCCGCA 1251
Xl AAACCATATCTATTTTCTTGGGAATCCGCATTTAAGTATCGAGATGCGTTGGCATTAA 1259
Pl AAACCGTCGCAATAGCGCGTGGGAGTCGTCTTCAAATATAGAGATATATTAGCAGAAC 1248
* * * * *

Vh AAGGGGTAGAAAGATCGTTGAATCTGGGATGAATAATATATTAGAGCTGGTGGCGCAC 1311
Xl GAGGTGCGGAAAGGATTGTAGAAGCAGGAATGAATAATATATTTCGAGTTGGTGGATCTC 1319
Pl ATGGTGCAGAACGAATTATAGAAGCGGAATGAATAATATATTTCGAGTAGGTGGCGCCC 1308
* * * * *

Vh ATGATGCAATGCGCCCACTTCAACGTTAGTTCGATTGTTTTCTCATGAAAGACCATATA 1371
Xl ATGACGGAATGAGGCGTTACAACGATTAGTGACATATATTTCTCATGAGAGGCATCTC 1379
Pl ATGATGGGATGCGTCCCCCTTCAACGCTTGTAACTATATATCATGAAAGGCGGTCAT 1368
* * * * *

Vh ACTTCACCACTAAGGATGTATCTGTGCAAAATAGAGCAAACCCGCTTTCTTGAAGAAGATA 1431
Xl ATTATACTGCTAAGGATGTTGCGGTTGAAATAGAACAGACTCGATTCTTGAAGAAGATA 1439
Pl CATATACCACTAAGATGTCTCGGTGAAAATCGAACAGACTCGTTATCTTGAGGAAGATA 1428
* * * * *

Vh AATTCTTGTTTTCTGTCCTTAAATAAGGAAA-ATTACAATGAATAA--TCAATGCAAG 1487
Xl AGTTCCTTGATTTGTCCCGTAATAGGTAAAA-AGTATGGAATAAATCCAAATATAAA 1498
Pl AGTTCCTGATTTGTACCGTAGAAAGAGATATATCATGGAATAACACAACATTCATTA 1488
* * * * *

Vh ACTATTGCACACGTGTTACGCGTGAACAATGGTCAGGAACCTCACGTCTGGGAAACGCCC 1547
Xl ACCATCGACCATGTTCTTTGTGTTGAAGGAAATAAAAAAATTCATGTTTGGGAAACGCTG 1558
Pl CCTATTGATCACGTAATTGATATTGGTGATAACCGTTATATTGAGTATGGGAAACCAAG 1548
* * * * *

Vh CCAAAAGAAAATGTGCCTTCTAAAAATAACACCATTTTAATTGCTTCTGGTTTTGCCAGA 1607
Xl CCAGAAGAAACAGCCCAAGAGAAAGAAATCCATTATTATTGCGTGGGTTTTGCCCGA 1618
Pl CCGAAAAATAAGAAACCAAGCGTAATAATACCATCGTTATCGCCTCAGGCTTTGCTCGA 1608
* * * * *

Vh	AGAATGGATCACTTTTGCTGGGTTGGCCGAGTATCTATCTGAAAATGGTTTTCCATGTGTTC	1667
Xl	AGGATGGATCATTTTTGCTGGTTTAGCGGAATATTTATCGCGGAATGGGTTTCATGTGATT	1678
P1	CGCATGGATCATTTTGCTGGTCTTGCCGAATATTTAGCAAATAATGGTTTTCGTGTTATT *****	1668
Vh	CGTTATGACTCTCTTCATCATGTGGGGCTCAGCTCGGGTTCGATTGATGAATTCACCATG	1727
Xl	CGCTATGATTACCTTACCCAGCTTGGGTTGAGTTCAGGGACAATTGATGAATTTACAATG	1738
P1	CGTTATGATTCTGTTAAATCATGTCGGTCTTAGTCAGCGAGAGATCAAACAGTTCGCATG ** *****	1728
Vh	ACGACGGGCCAAAAATAGTTTGTGCACTGTTTATCACTGGCTGCAGACCAAGGGCACACAA	1787
Xl	TCTATAGGAAAAACAGAGCCTATTAGCCGTGGTTGATTGGTTAAATACAGAAAAATAAAT	1798
P1	TCAGTAGGTAAACACAGTTTGCTAACTGTTATTGATTGGCTAAAAGAACGAAATATTAAC * ** *	1788
Vh	AACATTGGCTTGATTGCGGCAAGTCTCTCAGCTCGTGTCGCCTATGAAGTTATCTCTGAT	1847
Xl	AACCGTGGTATTTTGGCTTCAAGCTTATCTGCACGGATAGTTTATGCAAGTCTATCTGAA	1858
P1	AATATAGGTCTTATTGCTTCGAGTCTTCTGCTCGAATTGCTTATGAAGTGGCAGCAGAA ** * *	1848
Vh	CTGGAGCTGTCTTTTCTGATTACTGCGGTTGGTGTGGTGAAGTTGCGTGACACACTAGAA	1907
Xl	ATTAATGTTTTCATTTTTTAATCACC GCAGTCGGTGTTGTTAACTTAAGATATACGCTTGAA	1918
P1	ATTGATTTGT CATTTTTAATTACGCCGTCGGTGTTGTCAATTTAAGAAGTACCGTAGAA * * * *	1908
Vh	AAAGCGCTTG GTTTTTGATTACCTCAGTTTG CCTATCGATGAGCTACCAAACGATCTTGAT	1967
Xl	AGAGCTTAGAATTTGATTATCTCAGTTTACCCATTAAATGAATTGCCGAATAATTTGAAT	1978
P1	AAAGCACTGAAATATGATTATCTACAAATGGAAGTAAATACTATTCTCGAAGATTTAAGT * *** *	1968
Vh	TTTGAAGGTCATAAGCTTG GTTCTGAAGTG TCGTTGCGACTGCTTCGAGCATCACTGG	2027
Xl	TTTGAAGGCCATAAA TTGGGTGCTGAAGTCTTTGCGAGAGATTGCCTTGATT TTGGCTGG	2038
P1	TTTGAAGGACACAATCTAGGTTCAA AAGTCTTTGTGACAGATTGTTT TGAAAATAATTGG *****	2028
Vh	GATACCTTAGATTCTACTCTCGACAAAGTAGCCAATACCTCGGTTCC TTTAATCGCCTTT	2087
Xl	GAAGATTTAAC TTTCTACAATCAATAGCATGATGTATCTTGATATACGTTTTATTGCTTTT	2098
P1	GACTCATTAGATT CGACAATAAATAAAATTTGTGAAGT AGATATTC CATTATTGCTTTC ** *** *	2088
Vh	ACCGCTAACACGATGATTGGGTTAAGCAAGAAGAAGTCTATGACAT - GTTAGCGCATAT	2146
Xl	ACTGC AAAATAACGACAATTGGGTAAAGCAAGATGAAGT- TATCACATTGTTATCAAATAT	2157
P1	ACTTCAGATGGCGATGATTGGGTTTGCCAACATGAAGTAAACATTTAGTCAGTA - ACGT ** * *	2147
Vh	CCGCACTGGGCATTGCAAGCTCTACTCC TTGCTTGGTAGCTCTCATGACTTGGGCGAAAA	2206
Xl	TCGTAGTAATCGATGCAAGATACATTC TTTGTTAGGAAGTTGCGATGACTTGTGTGTTTTT	2217
P1	TAAATCTGACAAAAAGAAAATTTACTC ACTCGTTGGCTCATCTCATGATTTGGGCGAAAA * ** *	2207
Vh	CTTG TGTGCTGT TACGTAATTTTACCAATCCGTCACCAAAGCGCCATCGCAATGGATGG	2266
Xl	CTTAGTGGTCTGCGCAATTTTTATCAATCGGTTACGAAGGCTGCTATCGCGATGGATAA	2277
P1	CCTAGTGGTGCTTCGTAAC TTTCTATCAATCAATGACGAAAGCTGCTGTGAGCTTAGATCG * * * *	2267
Vh	AGGCAGCTTAGAAATCGACGTCGACTTTATCGAGCCTGATTTTGAACAACTCACCATCGC	2326
Xl	TGATCGTCTGGATATTGATGTTGATATTATTGAACCATCATTCGAACATCTAACTATTGC	2337
P1	TCAATTAGTAGAGCTTGTTGATGAAATTATTGAACCAAATTTTGAAGACCTAACAGTTAT * * * *	2327
Vh	GACTGTGAATGAACGTCGCTTGAAAGCGGAAATTGAAAGCCGTACGCCAGAAATGGCTTA	2386
Xl	GACAGTCAATGAACGTCGAATGAA AATTGAGATTGAA AATCAAGCGCATTTTCGCTGTCTTA	2397
P1	TACGCTAAATGAACGCGCTCAAAAATAAAATCGAAATGAAAT - TATTATAGATTA ** * *	2385
Vh	GGTCTCATCGTAATA-----	2401
Xl	AAACCTATTGGGATAGATAATTACCCTATAGATTTCAAGATGGATCGCGACGCAAGGGAG	2457
P1	GCTGATCGCGTATTGGCTAGTGTCTAAATAGTACTTACCTAAGT----- * *	2429
Vh	-----CCAATAAATAA	2411

17 / 22

X1 CGAATCCCGGAGCATAGCAAACCTATGTGACCGGGGTGAGTGAGTGCAGCCAAACAAAGAA 2517
P1 -----ACAGCCAAAAG 2440
** ** *

Vh G-----GAAATGT-----TATG 2424
X1 GCAACTTGAAAGATAACGGGTATAGTTAATTCTATCACTCAAATATAAGGGCTCTCTATG 2577
P1 G-----AAGAAATA-----ATG 2452
* * * *

Vh AAATTTGGAAACTTCCTTCTCACTTATCAGCCACCTGAGCTATCTCAGACCGAAGTGATG 2484
X1 AAATTTGGAAACTTTTTGCTTACATACCAACCCCCCAATTTTCTCAAACAGAAGTAATG 2637
P1 AAAATTAGTAATATCTGTTTCTCATACCAACCACCAGGTGAATCACATCAAGAGGTAATG 2512
*** ** * * * * * * * * * * * * * * * *

Vh AAGCGATTGGTTAATCTGGGCAAAGCGTCTGAAGGTTGTGGTTTCGACACTGTTTGGTTA 2544
X1 AAACGTTTGGTTAAATTAGGTCGTTTCTGAGGAGTGTGGTTTGTGATACTGTATGGTTA 2697
P1 GAGCGCTTTATTCTGTTTAGGCGTTGCATCAGAAGAGCTCAACTTTGATGGTTTCTATACA 2572
* * * * * * * * * * * * * * * *

Vh CTAGAGCACCACCTTCACTGAATTGGGTTGTTAGGGAATCCTTATGTTGCTGCCGCACAT 2604
X1 CTGGAGCATCATTTACGGAGTTTGGTTTGGCTTGGTAACCCCTTATGTCGCTGCTGCATAT 2757
P1 CTTGAACACCATTTCACTGAGTTTGGTATTACAGGTAACCTTTATATTGCCTGTGCCAAT 2632
** * * * * * * * * * * * * * * * *

Vh CTATTAGGTGCGACAGAAAAGCTCAATGTTGGCACCAGCAGCCATTGTATTGCCGACTGCC 2664
X1 TTACTTGGTGCAACCAAAAAATGAATGTAGGGACTGCGGCTATTGTTCTTCCACCGCT 2817
P1 ATTCTTGGTCAACCAAAAGGATCCAAGTCGGTACCATGGGGATAGTGTACCACAGAG 2692
* * * * * * * * * * * * * * * *

Vh CATCCGGTTCGACAAGCAGAAGACGTAACCTACTGGATCAAATGTCGAAAGGACGATTC 2724
X1 CATCCAGTGCGCCAACCTGAAGATGTGAATTTATTGGATCAAATGTCGAAAGGACGATTC 2877
P1 CACCCAGCACGACATGTGAAAAGTCTTCTCGTTTATAGATCAACTGTCTAAAGGGCGCTTT 2752
** * * * * * * * * * * * * * * * *

Vh CGTTTTGGGATTTGTGCGGGTTTGTACGACAAAGATTTCCGTGTCTTTGGTACAGACATG 2784
X1 CGGTTTGGTATTTGTGCGGGGCTTTACAATAAAGACTTTGCGGTATTTGGCAGCGATATG 2937
P1 AACTACGGTACTGTTGCGGACTCTACCATAAAGATTTTCGTGTTTGGTACATCACAG 2812
* * * * * * * * * * * * * * * *

Vh GATAACAGCCGAGCCTTAATGGACTGTTGGTATGACTTGATGAAAGAAGGCTTCAATGAA 2844
X1 AATAACAGTCGCGCTTTAACGGAGTGCTGGTACGGGTTGATAAAAAATGGCATGACAGAG 2997
P1 GAAGATTCTCGTAAGACCGCAGAAAATTCTACTCTATGATCTTGGATG-CATCAAAAAC 2871
* * * * * * * * * * * * * * * *

Vh GGCTATATCGCG-GCGGATAACGAACATATTAAGTTCCCGAAAAATCCAAGTGAATCCATC 2903
X1 GGATATATGGAA-GCTGATAATGAACATATCAAGTTCCATAAGGTAAAAGTAAACCGAC 3056
P1 AGGTGTGCTACATACTGACGGTGAAGTAGTAGAGTTCCAGATGTCAATGTTTATCCAGA 2931
* * * * * * * * * * * * * * * *

Vh GGCTTACACACAAGGTGGCGCTCTGTTTATGTGTCGCGGAGTCAGCATCAACGACAGA 2963
X1 AGCATATAGTAAAGGTGGAGCCCTGTTTATGTGGTTGTGAATCAGCCTCGACAACCTGA 3116
P1 AGCTTACAGCAAAAAACAA--CCACCTGCATGACAGCCGAATCATCCGAGACCATCAC 2988
** * * * * * * * * * * * * * * * *

Vh ATGGGCTGCAGAGCGTGGCCTACCAATGATTCTAAGCTGGATCATCAATACTCACGAGAA 3023
X1 ATGGGCGCTCAATTTGGTTTACCGATGATATTAAGTTGGATTATAAATACTAACGAAAA 3176
P1 TTATTAGCTGAACGTGGTTTACCAATGGTGTAAAGTTGGATTATTCGGTCAAGTGAGAA 3048
* * * * * * * * * * * * * * * *

Vh GAAAGCGCAGCTTGATCTTTACAATGAAGTCGCGACTGAACATGGCTACGATGTGACTAA 3083
X1 GAAAGCACAGCTTGAGCTTTATAACGAGGTGGCTCAAGAATATGGGCACGATATTCTATA 3236
P1 AGTCTCACAAATGGAATTGTACAATGAAGTTGCGGCAGAGCATGGTCATGACATTAAACAA 3108
* * * * * * * * * * * * * * * *

Vh GATTGACCACTGTTTGTCTTACATCACCTCCGTCGATCATGACTCAAATAAAGCCAAAGA 3143
X1 TATCGACCATTGCTTATCATATATAACATCTGTAAATTATGACTCAAATAAAGCGAAAGA 3296
P1 CATTGAACATATCTTAACCTTTATTGCTCTGTAAATGAAGACGGTGAAAAAGCAGACAG 3168
* * * * * * * * * * * * * * * *

Vh TATTTGCCGCAACTTCTTGGGTCAATTGGTACGACTCATACGTGAATGCCACCAAGATTTT 3203
X1 GATTGTGCGGAAATTTCTAGGCGATTGGTATGATTCTTATGTGAATGCCACGACCATTTT 3356
P1 CGTATGCCGTAATTTCTAGAAAATTGGTACGACTCTTACAAAAATGAACCAACATCTT 3228

Vh	GTGAGACTCTGACCAAAACAAAAGGCTACGACTTCAATAAAGGGCCAATGGCGCGATTTTGT	3263
Xl	TGATGATTTCAGACAAAACAAGAGGTTATGATTTCATAAAGGGCAGTGGCGTGACTTTGT	3416
Pf	CAACGACAGTAACCAAACCTCGTGGCTACGATTACCTCAAAGCTCAGTGGCGTGAGTGGGT * * * * * * * * * * * * * * * * * * * * * * * *	3288
Vh	GTTGAAAGGCCACAAAGATACCAACCGCCGAATTGATTACAGCTACGAAATCAACCCAGT	3323
Xl	ATTAAAGGGACATAGAGATACTAATCGCCGCATTGATTACAGTTACGAAATCAATCCCGT	3476
Pf	AATGAAGGGTTGGCTGATCTCGTCGCCGACTTGATTACAGTAACGAATTAACCCCTGT * * * * * * * * * * * * * * * * * * * * * * * * *	3348
Vh	AGGGACGCTGAAGAGTGTATCGCGATTATCCAGCAAGATATCGATGCAACGGGTATTGA	3383
Xl	GGGAACCCCGCAGGAATGCATTGACATAATTCAAAAAGACATTGATGCCACGGGAATATC	3536
Pf	CGGCACGCCAGAACGATGCATTGAGATCATTCAAAGTAATATTGATGCCACTGGAATTAA * * * * * * * * * * * * * * * * * * * * * * * * *	3408
Vh	CAATATTGTTGTGGTTTTGAAGCAAACGGTTCTGAAGAAGAAATTATCGCATCTATGAA	3443
Xl	AAATATCTGTTGTGGGTTTGAAGCGAATGGAACAGTAGACGAAATTATTGCTTCCATGAA	3596
Pf	GCATATTACTGTTGGATTTTGAAGCGAATGGTTCTGAACAAGAAATTCGTGAATCCATGGA * * * * * * * * * * * * * * * * * * * * * *	3468
Vh	GCTATTCCAGTCCGATGTGATGCCATATCTCAAAGAAAAACAGTAATTAATATTTTCTAA	3503
Xl	GCTCTTCCAGTCTGATGTCATGCCGTTTCTTAAAGAAAAACAACGTTCCGTATTATAGTA	3656
Pf	GCTATTTATGGA AAAAGTAGCGCCACACTTAAAGATCCTCAGTAAGCTGTTCTTTTAA * * * * * * * * * * * * * * * * * * * *	3528
Vh	-----AAGGAAAGAGACATGAAATTTGGATTATTCTTCTCTCAATT	3543
Xl	GCT-----AAGGAAAAAGAAATGAAATTTGGATTGTTCTTCTCTTAAT	3699
Pf	ACTATTCAATATCAAGGCATAAGGAATAAAATATGAATTTCCGGTTATTTTTCCTAAAT * * * * * * * * * * * * * * * * * * * * *	3588
Vh	TTATGAATTCAAAGCGTTCTTCTGATCAAGTCATCGAAGAAATGTTAGATACCGCACATT	3603
Xl	TCATCAATTCAAACACTGTTCAAGAACAAAGTATAGTTCGCATGCAGGAAATAACGGAGT	3759
Pf	TCCAGCTGAAGGTATGACTTCAGAAATGGTTTGAACAACATGGTAGACTACTGTCGCAT * * * * * * * * * *	3648
Vh	ACGTAGATCA-GTTGA- ---AGTTTGACAGGTTGGCTGTTTACGAAAACCATTTCTCGA	3657
Xl	ATGTTGATAA-GTTGA- ---ATTTTGAACAGATTTTGGTGATGAAAATCATTTTTCAG	3813
Pf	TAGTGGATAAAGATGATTACCACTTTAAAGAGTGCTCGTCAGCGAGCATCATTTTCTA * * * * * * * * * * * * * * * * * * *	3708
Vh	ACAATGGTGTGGTTGGTGCCCCATTAAACAGTGGCTGGTTTTTACTTGGTATGACAAAGA	3717
Xl	GTAATGGTGTGTGCGTGCTCCTCTGACTGTTTCTGGTTTTTGTCTCGGTTTAAACAGAAA	3873
Pf	AAAACGGCATTATCGGAGAACCTTTGACAGCGATTAGCTTCTTACTTGGTTTGACTAAAC * * * * * * * * * * * * * * * * * * * * *	3768
Vh	ACGCCAAAGTGGCTTCGTTGAATCACGTTATTACAACACATCATCCAGTACGTGTGGCGG	3777
Xl	AAATTA AAATTTGGCTCATTGAATCACATCATTACAACATCATCTGTCCGAATAGCGG	3933
Pf	GTATAGAAATTTGGTTCCTTAAATCAAGTGATTACCACCATCATCTGTACGTATCGGAG * * * * * * * * * * * * * * * * * * * * * * * * * * * *	3828
Vh	AAGAAGCGTGTCTACTCGACCAAATGAGTGAAGGCCGTTTTGTCTTTGGCTTTAGTGATT	3837
Xl	AGGAGGCTTGGCTTACTTGGATCAAAATTAAGCGAAGGGAGATTTATTTTAGGGTTTAGTGATT	3993
Pf	AACAAACGGCTTACTTTGATCAATGTCTTACGGTCGTTTTAGGCTTTAGGCTTAAGTGACT * * * * * * * * * * * * * * * * * * * *	3888
Vh	GTGAAAAGAGTGCAGATATGCGCTTCTTTAATCGACCAACGGATTCTCAGTTTCAGTTGT	3897
Xl	GTGAAAAAAAAGATGAAATGCGCTCTTTTAAATCGCCCTGTTGAATATCAACAGCAACTAT	4053
Pf	GTGTCAATGACTTCGAAATGGATTCTTTAAGAGAAAAACGTAGCTCTCAACAGCAACAAT * * * * * * * * * * * * * * * * * * * *	3948
Vh	TCAGTGAGTGTCAAGATCATCAATGATGCATTACTACTGGGTACTGCCATCCAAACA	3957
Xl	TTGAAGAGTGTATGAAATCATTAACGATGCTTTAACAACAGGCTATTGTAATCCCGATA	4113
Pf	TCGAAGCATGTTACGAAATTTTAAATGAAGCGCTGACGACAACTATTGTCAGGCAGATG * * * * * * * * * * * * * * * * * * * * *	4008
Vh	ATGATTTTTATAGTTTTTCTTAAATCTCTGTTAACCACACGCTTACACTGAAGGCGGTC	4017
Xl	ATGATTTTTATAGTTTCCCTAAAATATCGGTA AACCCCACGCTTATACCAAGGCGGGC	4173
Pf	ATGACTTCTTTAACTTCCACGTAATTTCTGTTAACC CGCA---TTGTATT---AGCCAAG * * * * * * * * * * * * * * * * * * * * * * * * * * * * *	4062

19 / 22

Vh CTGCGCAATTTGTGAATGCGACGAGCAAAGAAGTGGTTGAATGGGCGGCTAAGTTAGGGC 4077
X1 CTCGGAGATATATTACAGCAACCAGTCATCATATTTGTTGAATGGGCGGCTAAAAAAGGCA 4233
P1 TAAAAAATATATTTTAGCTTCAAGCATGGGCGTGGTTGAATGGGCGAGCAAGAAAAGGAT 4122
* * * * *

Vh TTCCACTCGTGTTTAAATGGGACGACTCAAACGCGCAAAGAAAAGAATACGCCGGTTTGT 4137
X1 TTCTCTCATCTTTAAGTGGGATGACTCCAATGATGTTAGATATGAATATGCTGAAAGGT 4293
P1 TGCCACTCACTTACCGCTGGAGTGACAGCCTAGCAGAAAAAGAAAAATACTATCAGCGTT 4182
* * * * *

Vh ACCACGAAGTTGCTCAGGCACATAGTGTTCGATGTTAGTCAGGTTCTGA-CACAAGCTGACG 4196
X1 ATAAAGCCGTTGCTGATAAAATATGGCAATTGAC-TTATCAGCGATAGATCATCAGTTAATG 4352
P1 ATCTCGCTGTGCTAAAGAGAATAATATTGAT-GTATCAAATATTGACCACCAATTCCCA 4241
* * * * *

Vh CTGTTGGTCAACCAAAATGTAGATGGTGAAGCAGCAAGGGCAGAAGCACGCCTGTATTTG 4256
X1 GTATTTGGTTAACTATAACGAAGATAGTCACAAAGCTAAACAAGAGACGCGTGCATTTATC 4412
P1 CTGCTCGTTAATATCAATGAAAATCGTCGTATTGCTCGAGATGAAGTAAGGGAGTATATA 4301
* * * * *

Vh GAAGAGTTTGTCCGTGAATCTTACCCAAATA--CCGAATTTGAGCAAAAAATGGCAGAG 4313
X1 CGTGATTATGTTCTTGAATGTATCCTAATG--AAAATCTCGAAAAATAAAGTTGAAGAG 4469
P1 CAAAGTTATGTGAGTGAAGCCTACCCTACTGACCCCAACATTGAGCTAAGAGTAGAAGAG 4361
* * * * *

Vh CTGTTGTGCAAAAAATGCCATCGGTACTTATGAAGAAAGTACTCAGGCAGCGCAGTTGCG 4373
X1 ATAATCACAGAAAACGCTGTGCGAGATTATACGGAATGTATAGCTGCGGCTAAGCTGGCA 4529
P1 CTTATTGAGCAGCATGCTGTGCGCAAAAGTGGATGAGTACTACGACTCAACAATGCACGCA 4421
* * * * *

Vh ATTGAGTGTGTGGTGCCGCGGACCTATTGATGTCTTTTGAAGTTCGATGGAAGATAAAGCG 4433
X1 ATTGAAAAGTGCAGTGCAAAAAGGGTATTATTATCCTTTGAACCAATGAATGACTTGATG 4589
P1 GTAAAAGTTACAGGTTCAAAAAATTTACTCTCTTTTGAATCAATGAAAAATAAAGAC 4481
* * * * *

Vh CAGCAAAGAGCGGTTATCGATGTGGTAAACGCCAACATCGTCAAATACCACTCGTAACGT 4493
X1 CACCAAAAAATGTAATCAATATTGTTGATGATAATTTAAAAAGTACCACATGTAGTA- 4648
P1 GATGTTACCAAGCTTATAAATATGTTTAAATCAAAAAATCAAAGATAACCTTATTAAATAA 4541
* * * * *

Vh TT-----AACTGATGCTGAAGGGGCGAGCA-----TGCCCTTATAT----- 4530
X1 -----AAAGAATAT-----GGCAGCAA-----CGCT----- 4669
P1 TTTAATTACGGATAGATATTTTCGATATATCTAAGTCTTACTACCATTATATAAACTAT 4601
* * * * *

Vh -----CACCATTCTTTTCGCGCGA-----TAGCGCTAACTAATAG--- 4565
X1 -----GCCATATTCTCTA-----AATTATTG--- 4691
P1 TTATACAGATAACGTTTCATTGATTAAGTCAGTAAATAATTGCCATTAAATTAATGGCAG 4661
* * * * *

Vh -----AGGCATTTATATGGACGTACTTTTCAG 4591
X1 -----GAGGGGTAAAACAGGTATGACTTCAT 4717
P1 TGCAGATCCTTACACTGCCATTATAAATTAATAAGGGTTAACATGTCAACATTATTAA 4721
* * * * *

Vh CGGTTAAGCAGGAAAAATATCGCAGCGAGCACAGAAATCGATGACTTGA'TTTTCATGGGAA 4651
X1 ATGTTGATAACAAGAGATCATAGCAAGCTCAGAAATTGATGATTTGATTTTTCAGCG 4777
P1 ATATAGATGCAACTGAAATTAAGGTGAGTACAGAAATAGATGATATTATTTTACATCAT 4781
* * * * *

Vh CTCCTCAGCAATGGTCAATGCGAGGAACAAAAACAGCTGACATCTCGCCTTGTTAAAGGGG 4711
X1 ATCCATTAGCTTGGTCTTACGATGAACAGGAAAAATCAGAAACAAATTGTTCTTGATG 4837
P1 CACCCTAACGTTACTATTTGAAGATCAAGAAAAATACAGAAAGAACTTATTTTGGAGT 4841
* * * * *

Vh CATATCAATACCATTACCACAATAATGATGACTATCGTCAGTTCTGCCAAAGGCTGGGAG 4771
X1 CATTTTCGTAATCACTATAAACATTGTCAAGAATACCGTCACTACTGTCAGGTACACAAAG 4897
P1 CTTTCCATTATCATTACAATCATAATAAGATTATAAGTACTATTGTAATATACAAGCG 4901
* * * * *

Vh TCGGAGAGGAGGTGCAAGATCTCAATGAGATCCCCGTTTCCCTACTTCTATTTTAAAGT 4831
X1 TAGACGACAATATTACGGAAATGATGACATACCTGTATTCCCAACATCAGTTTAAAGT 4957

20 / 22

P1 TAGATGAGAATATACAGTCCATTGACGATATTCTGTTTTCTTACTTCAATGTTCAAGT 4961
* * * * *

Vh TGAAGACCCATTAAACACTTGACGATGAAGAGGTAGAGAATCGCTTTACTAGCAGCGGTA 4891
X1 TTAAGTGGTATTAACCTTCTCAGGAGAACGAGATTGAAAGTTGGTTTACCAGCAGCGGCA 5017
P1 ACTCAAGATTACATACTGCTGATGAATCAAATATTGAAAATTGGTTTACTAGTAGTGGTA 5021
* * * * *

Vh CTAGTGGCATCAAAAGTATTGTGCGACGAGATAGACTCAGTATTGAGCGACTTCTTGGCT 4951
X1 CGAGTGGTTTTAAAAAGTCAGGTGGCGCGTGACAGACTAAGTATTGAGAGACTCTTAGGCT 5077
P1 CAAAGGGAGTCAAAAGTCATATAGCTCGAGATCGGCAGAGTATTGAACGCTTGCTAGGTT 5081
* * * * *

Vh CAGTAAATTTCCGGTATGAATTACGTTGGTGATTGGTTTGACCATCAGATGGAGTTGGTGA 5011
X1 CTGTGAGTTATGGCATGAAATATGTTGGTAGTTGGTTTGATCATCAAATAGAGTTGGTCA 5137
P1 CTGTTAATTACGGCATGAAATCTTGGGTGAATTTACGAGCATCAATTAGAACTAGTGA 5141
* * * * *

Vh ACTTAGGCCCAGATCGCTTTAATGCCAACAAATATTGGTTCAAGTACGTCATGAGCTTAG 5071
X1 ACTTAGGGCCAGATAGATTTAATGCTCATAACATTTGGTTTAAATATGTTATTAGTTGG 5197
P1 ATATGGGGCCAGATCGTTTCAGTGCCTCAAATGTTGGTTTAAATATGTAATGAGCTTAG 5201
* * * * *

Vh TCGAGCTCCTTTATCCGACCGCATTTACTGCCACTGAGGATGAGATCGACTTTGAGGCGA 5131
X1 TAGAATTATTATATCCACGACATTTACCGTAATGGAAGAACGAATAGATTTTGTTAAGA 5257
P1 TTCAATTACTTTACCCAACAACATTTACCGTTGAAAACGATGAAATCGATTTTGAACAAA 5261
* * * * *

Vh CGCTAGCTAATATGAATCG - TATTAAGCAGTCTGGTAAAACCATTTGTCTTATCGGCCCT 5190
X1 CATTGAATAGCTTGAGCG - AATAAAAAATCAAGGGAAGATATTGTCTTATCGGCTCA 5316
P1 CCATC - TTAGCGTTAAAGCAATTCAGCGTAAAGGAAAGGAATTTGTTTAAATTGGCCCT 5320
* * * * *

Vh CCTTATTTTATCTATCTACTGTGCTGTTTCATGCGCGAGCAAGGTCAAACCTTTCAATGGT 5250
X1 CCATACTTTATTTATTTGCTCTGCCAGTATATGAAAGATAAAAACATCTCATTTTATGGG 5376
P1 CCGTATTTTATTTATTTGTTATGCTTACCATGAAAGAGCATAATATCGAATTTAATGCT 5380
* * * * *

Vh GGTGCGGATCTTTACATCATCACTGGCGGCGGCTGGAAAAACATCAGGATCAATCGCTC 5310
X1 GATAAAAACCTTTATATCATAACGGGGCGGCTGGAAAAGTTATGAAAAAGAGTCCCTA 5436
P1 GGTGCACATATGTTTATCATACAGGTGGGGGATGGAAAACCAACAAAAAGAGCGCTA 5440
* * * * *

Vh GATAGAGACGAGTTCAACCAGCTTTTGTGTGAGACTTTTACCTTAGAAAGCGCAGAGCAG 5370
X1 AAACGCGATGATTTCAATCATCTTTTATTCGACACGTTCAACCTCAATAATATTAGTCAA 5496
P1 AACCAGACAAGATTTCAATCAACTATTGATGGAGACTTTAGCCTTTTCCATGAAAGTCAA 5500
* * * * *

Vh ATTCGAGACACATTTAATCCAGTTGAACCTGAACACCTGCTTTTTTGAAGATACAGAACAC 5430
X1 ATCCGCGATATTTAATCAAGTTGAACCTCAACACTTGTTCTTTGAGGATGAAATGCAA 5556
P1 ATTCGAGATATCTTTAACCAAGTAGAGCTAAACACTTGTTTCTTTGAAGACAGCTACAG 5560
* * * * *

Vh AAAAAGCGTGTACCGCCCTGGGTCTTTGCAAGAGCTCTGGATCCTAAAACATTAAAGCCG 5490
X1 CGTAAACGTGTTCCGCCGTGGGTATATGCGCGAGCACTTGATCCTGAAACATTGAAACCT 5616
P1 CGTAAACATGTACCACCGTGGGTATATGCTCGTGCCTTGATCCTGTCACTTTAACGCC 5620
* * * * *

Vh CTTCCGATGGCCAGCCAGGACTGATGAGCTATATGGATGCCTCGGCGGTACAGCTATCCA 5550
X1 GTACCTGATGGAATGCCGGGTTTGATGAGTTATATGGATGCGTCATCAACGAGTTATCCG 5676
P1 GTAGAAGATGGCCAAGAGGGCTTGATGAGTTATATGGATGCCTCATCTACCAGTACCCG 5680
* * * * *

Vh TGTTTTCTAGTGACGGATGACATCGGCATCGTGCG - - - - -AGAAGAAGAAGGCGATCGC 5604
X1 GCATTTATTGTTACCGATGATGTCGGGATAATGAG - - - - -CAGAGAATATGGTCAATAT 5730
P1 ACATTTATTGTTACCGACGATATTGGTATTGTTGCCATCTAAAAGAACCAGATCCATTC 5740
* * * * *

Vh CCGGGAACCACTGTTGAGATCGTTAGAAGAGTTAAGACGCGGGGTATGAAGGGGTGTGCT 5664
X1 CCTGGTGACTTGTTGAGATTTTACGTCGCGTCAATACGAGGGCACAGAAAGGGTGTGCT 5790
P1 CAAGGAACAACGGTTGAAATTTGTTTCGTCGTTTAAATACGCGAGAACAAAAGGATGTTCA 5800
* * * * *

21 /22

Vh CTCAGTATGTCTCAAGCATTTA-CAGCTAAGAATGAAGGAGGCAACTGACATGTTAT-GT 5722
X1 TTAAGCTTAAACCAAGCATTTAATAGTTGA----- 5820
P1 CTCTCAATGG-CCACGAGCCTG--AAATAAAAGCAGGGCTTAATCATGATTTTAAATGTC 5857
* * ** * * * *

Vh TCGATAGAAAAAATTGAGCCGTTAACTAACCTCATATTCGAGTATTGCTCAAGCCAGAT 5782
X1 -----
P1 AAGGTTAAAAAAGTCGAAGCATCTGACAGCCATATTTACAAAGTGTTTATTAAGCCTGAC 5917

Vh CAGCCTTTTGAATTTAGGGCAGGGCAGTACATTAACTGTCAGTTTAAAGCTTTGGTAGTTTA 5842
X1 -----
P1 AAATGCTTTGATTTTAAAGCGGGTCAATATGTAATTGTGTATCTCAATGGAAAAAATTTG 5977

Vh CCGTTTTCTATAGCCTCATGTCTTCTAATGGTGCCTTTTTAGAACTCCATATTGGTGGC 5902
X1 -----
P1 CCGTTTTCTATTGCTAACTGCCAACTTGTAATGAGCTCCTTGAATTACATGTAGGAGGT 6037

Vh TCAGATATCAGCAAGAAAAATACGCTTGTGATGGAAGAACTCACCAATTCATGGGGCTGC 5962
X1 -----
P1 TCGG---TAAAAGAATCCGCCATTTGAAGCTATTTTCGCACTTTATTAATGCATTTATTTAT 6094

Vh GGCAACATGGTTGAAGTCAGTGAGGCGCGAGGTGAGGCTTGGTTGCGTGATGAGAGTGTC 6022
X1 -----
P1 CAAAAAGAATTTACAATCGATGCACCACACGGTGATGCATGGCTGAGAGATGAAAGCCAA 6154

Vh AAACCTTGTATTGTTGTCGAGGCGGGACGGGAATGTCATACACCCTAAGTATTTTGA 6082
X1 -----
P1 TCACCTTTACTACTTTATAGCAGGAGGGACAGGTTTATCATATATCAATAGCATTTTAAAGT 6214

Vh AATAGCTTGGAGCAAGGGTTTACCCAGCCGATTTACGTCTATTGGGGCGCCAAGGATATG 6142
X1 -----
P1 TGTGTATTAGTAAACAGTTATCTCAGCCTATCTATCTTTATTGGGGAGTAAATAACTGT 6274

Vh GATAACCTGTATGTACATGACGAACTGGTGGATATTGCGCTTGAAAACAAAAACGTACGT 6202
X1 -----
P1 AATTTACTCTATGCTGATCAACAATAAAAACACTCGCCGCACAATACAGAAATATAAAT 6334

Vh TACGTGCCAGTCACTGAAATATCAACCTGTCCCCAATACGCTAAGCAAGGAAAGGTGTTG 6262
X1 -----
P1 TATATTCCTGTGGTAGAGAATTTAAATACTGACTGGCAGGGAAAAATTGGTAATGTTATT 6394

Vh GAGTGTGTGATGAGTGATTTCGTAACCTTATCTGAGTTCGATATCTACTTGTGTGGTCCT 6322
X1 -----
P1 GACGCGGTTATTGAAGATTTTTCAGATTTATCTGACTTTGATATCTATGTCTGCGGGCCA 6454

Vh TGCAAAATGGTTGAAGTGGCTCGTGATTGGTTCTGTGACAAAAGAGGGGCAGAACCAGAG 6382
X1 -----
P1 TTTGGTATGAGCCGGACTGCGAAAGATATTCTGATCTCACAGAAAAAGGCGAATATAGGA 6514

Vh CAACTTTACGCGGACGCGTTTCGCTTATTTGTAATCATTATCAAGGAGAAGAACTATGAG 6442
X1 -----
P1 AAAATGTATTCTGATGCATTTAGCTATACGTAA----- 6547

Vh CTCAACGTCACTACTAGATGAGTTTGGCACTCCAGTACAAAGGGTAGAAAGAGCGATTGA 6502
X1 -----
P1 -----

Vh GGCTCTGAAAAATGGCCTTGGTGTCTATTAAATGGATGATGAGGATCGCGAGAACGAAGG 6562

22 /22

X1 -----
P1 -----

Vh CGACCTTATCTTCTCTGCACAGCATCTTACCGAAGCGCAAATGGCACTCATGATTCTGTA 6622
X1 -----
P1 -----

Vh ATGCAGTGGTATCGTGTGTTTGTGCTTAACGGAGGAACGCGCCAATTGGTTAGAGCTTCC 6682
X1 -----
P1 -----

Vh TCCTATGGTGAAAGATAATCGCAGTAAAAACCAGACCGCTTTTACGGTTTCGATTGAAGC 6742
X1 -----
P1 -----

Vh GAAAGAAGGGGTGACGACAGGAGTCTCTGCGAAAGATCGCGTTACAACGGTTAAAACGGC 6802
X1 -----
P1 -----

Vh TACTTATTTTGATGCTCAACCAGAAGATTTAGCAAGACCAGGCCATGTTTTTCCGCTGGT 6862
X1 -----
P1 -----

Vh TCGGAAAACAAATGGCGTGTTGGCCCGTCGAGGTCATACCGAAGGTACGATCGATTGAT 6922
X1 -----
P1 -----

Vh GTATCTAGCAAACCTTAGTCCCATCAGGGATCCTTTGCGAACTGACTAACCGTGATGGAAC 6982
X1 -----
P1 -----

Vh CATGGCGAAACTACCAGAAACCATTGAGTTTGCAAGACGTCATGGAATGCCAGTGCTCAC 7042
X1 -----
P1 -----

Vh TATTGAAGATATCGTCGATTATCGGACGGTAATTGAACTGAGAAATGAATATGAGAGTGG 7102
X1 -----
P1 -----

Vh CTTAGTGTGTGAAGTGAGTTGGTCTTAG 7130 (SEQ ID NO: 4)
X1 ----- (SEQ ID NO: 5)
P1 ----- (SEQ ID NO: 6)

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 08/09310

A. CLASSIFICATION OF SUBJECT MATTER

IPC(8) - C12N 15/87 (2008.04)

USPC - 800/288; 800/282

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC(8): C12N 15/87 (2008.04)

USPC: 800/288; 800/282;

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
USPC:435/189

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

PubWEST(USPT,PGPB,EPAB,JPAB); DialogPRO--Chemical Engineering and Biotechnology Abstracts, INSPEC, NTIS (National Technical Information Service), PASCAL, Current Contents Search, MEDLINE: bioluminescent plant, bacterial LUX operon,LUX A-G, F, vector, recombination

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X ----- Y	US 2005/0081268 A1 (Kay, et al.) 14 Apr 2005 (14.04.2005); Fig 1, para [0012], [0018], [0019], [0031], [0032]- [0040], [0050]-[0052], [0073], [0083], [0089]	33, 34 ----- 1-32 and 35-40
Y	US 2004/0002148 A1 (Gupta, et al.) 01 Jan 2004 (01.01.2004); para [0010]-[0012], [0017]	1-32, 35, 36
Y	US 2004/0126883 A1 (Liu) 01 Jul 2004 (01.07.2004); para [0014], [0030], [0091]	37-40
Y	Swartzman, et al. Delineation of the Transcriptional Boundaries of the lux Operon of Vibrio harveyi Demonstrates the Presence of Two New lux Genes. J Biol Chem 1990, 265(6):3513-3517, abstract	6, 24
Y	MEAGHEN. Bacterial Luminescence: Organization, Regulation, and Application of lux Genes. FASEB J 1993, 7(11):1016-1022, abstract	7, 8, 25 and 26
Y	LEE, et al. Riboflavin Synthesis Genes Are Linked with the lux Operon of Photobacterium phosphoreum. J Bacteriol 1994, 176(7):2100-2104; abstract	10-17, 31 and 36
Y	US 20060253916 A1 (Biesgen) 09 Nov 2006 (09.11.2006); para [0118]	12, 32
Y	DIAZ, et al. Biodegradation of Aromatic Compounds by Escherichia coli. Microbiol Mol Biol Rev 2001, 65(4):523-569; pg 534, col 1, para 1	13
Y	US 2006/0117400 A1 (Herz, et al.) 01 Jun 2006 (01.06.2006); para [0018], [0040]	29, 38-40

☐

Further documents are listed in the continuation of Box C.

☐

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

29 November 2008 (29.11.2008)

Date of mailing of the international search report

11 DEC 2008

Name and mailing address of the ISA/US

Mail Stop PCT, Attn: ISA/US, Commissioner for Patents

P.O. Box 1450, Alexandria, Virginia 22313-1450

Facsimile No. 571-273-3201

Authorized officer:

Lee W. Young

PCT Helpdesk: 571-272-4300

PCT OSP: 571-272-7774