

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

---

Faculty Publications from the Center for Plant  
Science Innovation

Plant Science Innovation, Center for

---

November 2006

## United States Patent Application Publication: COUNTER SELECTION STRATEGY FOR GRAM-NEGATIVE BACTERIA

Stephen K. Farrand  
*Seymour, IL*

Paul E. Staswick  
*University of Nebraska - Lincoln, pstaswick1@unl.edu*

Thomas E. Clemente  
*University of Nebraska - Lincoln, tclemente1@unl.edu*

Follow this and additional works at: <https://digitalcommons.unl.edu/plantscifacpub>



Part of the [Plant Sciences Commons](#)

---

Farrand, Stephen K.; Staswick, Paul E.; and Clemente, Thomas E., "United States Patent Application Publication: COUNTER SELECTION STRATEGY FOR GRAM-NEGATIVE BACTERIA" (2006). *Faculty Publications from the Center for Plant Science Innovation*. 15.  
<https://digitalcommons.unl.edu/plantscifacpub/15>

This Article is brought to you for free and open access by the Plant Science Innovation, Center for at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Faculty Publications from the Center for Plant Science Innovation by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.



US 20060265786A1

(19) **United States**  
 (12) **Patent Application Publication** (10) **Pub. No.: US 2006/0265786 A1**  
**Farrand et al.** (43) **Pub. Date: Nov. 23, 2006**

(54) **COUNTER SELECTION STRATEGY FOR GRAM-NEGATIVE BACTERIA**

(76) Inventors: **Stephen K. Farrand**, Seymour, IL (US); **Paul E. Staswick**, Lincoln, NE (US); **Thomas E. Clemente**, Lincoln, NE (US)

Correspondence Address:  
**JONDLE & ASSOCIATES P.C.**  
**858 HAPPY CANYON ROAD SUITE 230**  
**CASTLE ROCK, CO 80108 (US)**

(21) Appl. No.: **11/385,522**

(22) Filed: **Mar. 21, 2006**

**Related U.S. Application Data**

(62) Division of application No. 09/924,101, filed on Aug. 7, 2001, now abandoned.

(60) Provisional application No. 60/223,920, filed on Aug. 9, 2000.

**Publication Classification**

(51) **Int. Cl.**  
*A01H 1/00* (2006.01)  
*C07H 21/04* (2006.01)  
*C12N 1/20* (2006.01)  
*C12N 15/82* (2006.01)  
 (52) **U.S. Cl.** ..... **800/284**; 800/294; 536/23.2;  
 435/252.2

(57) **ABSTRACT**

A Gram-negative bacterium useful for genetically engineering plants is provided. The Gram-negative bacterium contains, as part of genome, an inducible regulatory sequence operatively linked to a nucleotide sequence encoding a levansucrase. Alternatively, the Gram-negative bacterium comprises a recombinant nucleic acid construct containing an inducible regulatory sequence operatively linked to a nucleotide sequence encoding a levansucrase. Also provided are recombinant nucleic acid constructs comprising an inducible regulatory sequence operatively coupled to a nucleotide sequence encoding a levansucrase and a method for transforming plants using the Gram-negative bacterium of the present invention.

## COUNTER SELECTION STRATEGY FOR GRAM-NEGATIVE BACTERIA

[0001] This application is a divisional of U.S. patent application Ser. No. 09/924,101, filed Aug. 7, 2001, which claims the benefit of provisional U.S. Patent Application Ser. No. 60/223,920, filed Aug. 9, 2000.

### BACKGROUND

[0002] *Agrobacterium* is a genus of soil Gram-negative bacteria that are widely used for the introduction of exogenous DNA into plants. The use of *Agrobacterium* species for DNA transfer is based on their natural ability to transfer DNA sequences into the genomes of plants. The most widely used species of *Agrobacterium* is *A. tumefaciens* the causal agent of the neoplastic disease crown gall in plants. A closely related species, *A. rhizogenes*, induces hairy root disease and also has been used for DNA transfer to plant genomes, but to a lesser extent. The ability of these bacteria to transfer DNA into plants depends on the presence of large plasmids (>100 kb) within the cells. These plasmids are referred to as the Ti (Tumor inducing) or Ri (Root inducing) in *A. tumefaciens* and *A. rhizogenes*, respectively. The mechanism for DNA transfer from the bacterium into the plant genome involves the mobilization of specific T-DNA (transfer DNA) molecules from the Ti plasmid into the host cell. The T-DNA region is delineated by 25 bp referred to as the left and right borders. In pathogenic *Agrobacterium* cells, within the T-DNA element reside genes for the over production of auxins and cytokinins which manifest the crown gall symptoms. The T-DNA element of pathogenic *Agrobacterium* strains also contains genes for the production of opines that are utilized by the bacterium as a nitrogen source.

[0003] *Agrobacterium*-mediated DNA transfer to plant cell genomes is usually conducted with "disarmed" (auxin, cytokinin and opine gene sequences removed from the T-DNA element) strains. In transformation studies, sequences of interest are introduced into the T-DNA region of a "disarmed" *Agrobacterium* strain. This chimeric T-DNA element can be carried on a separate, smaller, wide host range plasmid referred to as a binary vector or directly introduced into the resident "disarmed" Ti plasmid. The first step in the basic *Agrobacterium*-mediated transformation protocol, requires the inoculation of plant cells with transconjugants of "disarmed" *Agrobacterium* cells carrying the sequences of interest on the chimeric T-DNA element. The plant cells are subsequently cultured for a period generally ranging from one to seven days in a step of the protocol referred to as co-cultivation. Following the co-cultivation period, the plant cells are subcultured on regeneration medium for whole plant development.

[0004] A critical step in this process is the elimination of the *Agrobacterium* cells during plant development. Currently, *Agrobacterium*-mediated DNA transfer protocols incorporate antibiotics into the regeneration medium as a strategy to counter select *Agrobacterium* cells. Although successful, this approach adds significant cost to the transformation process. Moreover, there have been reports demonstrating a negative impact on plant tissue using medium supplemented with the antibiotics commonly used to eliminate *Agrobacterium* cells such as ticarcillin, cefotaxime, carbenicillin or vancomycin.

[0005] The *sacB* gene from *Bacillus subtilis* encodes for the enzyme levansucrase, which hydrolyzes sucrose to produce the polysaccharide levan, the presence of which causes the lysis of several Gram-negative bacteria, and in particular *Agrobacterium* (Gay et al. (1985) *J. Bacteriol.* 164:918-921). Expression of *sacB* is controlled by its regulatory sequence *sacR*. The present inventors have discovered that the introduction of a sequence encoding a levansucrase and in particular the *sacB* open reading frame (ORF), under strict control of an inducible regulatory sequence allows for the selection of *Agrobacterium* cells without the use of antibiotics. Thus, the present invention overcomes the problems of cost and negative effects on plant culture associated with the use of antibiotics for counter selection of the *Agrobacterium* cells.

### SUMMARY

[0006] Among the several aspects of the invention is provided a Gram-negative bacterium comprising an inducible regulatory sequence operatively linked to a nucleotide sequence encoding a levansucrase contained within the genome of said Gram-negative bacterium.

[0007] Another aspect provides, a Gram-negative bacterium comprising a recombinant nucleotide sequence containing an inducible regulatory sequence other than *sacR*, operatively linked to a nucleotide sequence encoding a levansucrase.

[0008] Yet another aspect provides, a recombinant nucleic acid construct comprising an inducible regulatory sequence other than *sacR*, operatively linked to a nucleotide sequence encoding a levansucrase.

[0009] Still another aspect provides, a method for transforming a plant cell comprising obtaining an *Agrobacterium* strain whose genome contains an inducible regulatory sequence operatively linked to a nucleotide sequence encoding a levansucrase; introducing a DNA construct into a T-DNA element of the *Agrobacterium*; and inoculating at least one plant cell with the *Agrobacterium* containing the construct for a time sufficient for mobilization of the T-DNA element from the *Agrobacterium* to the plant genome.

[0010] Another aspect provides, a method for transforming a plant cell comprising obtaining an *Agrobacterium* strain comprising a first recombinant nucleic acid construct containing an inducible regulatory sequence other than *sacR*, operatively linked to a nucleotide sequence encoding a levansucrase; introducing a second DNA construct into a T-DNA element of the *Agrobacterium*; and inoculating at least one plant cell with the *Agrobacterium* containing the first and second constructs for a time sufficient for mobilization of the T-DNA element from the *Agrobacterium* to the plant genome.

[0011] A further aspect provides, a method for counter selecting against a Gram-negative bacterium whose genome contains an inducible regulatory sequence operatively linked to a nucleotide sequence encoding a levansucrase comprising, introducing, in the presence of sucrose, a suitable inducer to cause the production of levansucrase by the bacterium resulting in the lysis of the bacterium.

[0012] Still a further aspect provides, a method for counter selecting against a Gram-negative bacterium containing a recombinant nucleic acid construct that includes an induc-

ible regulatory sequence other than sacR, operatively linked to a nucleotide sequence encoding a levansucrase, comprising introducing, in the presence of sucrose, a suitable inducer to cause the production of levansucrase by the bacterium resulting in lysis of said bacterium.

[0013] Another aspect provides, a vector comprising a recombinant nucleic acid construct containing an inducible regulatory sequence other than sacR, operatively linked to a nucleotide sequence encoding a levansucrase.

#### Definitions

[0014] As used herein, “regulatory sequence” means a sequence of DNA concerned with controlling expression of a gene; e.g. promoters, operators and attenuators. A regulatory sequence, may, potentially operate in conjunction with the biosynthetic apparatus of a cell.

[0015] As used herein, “polynucleotide” and “oligonucleotide” are used interchangeably and mean a polymer of at least two nucleotides joined together by a phosphodiester bond and may consist of either ribonucleotides or deoxynucleotides.

[0016] As used herein, “sequence” means the linear order in which monomers in a polymer, for example, the order of amino acids in a polypeptide or the order of nucleotides in a polynucleotide.

[0017] As used herein, “peptide”, and “protein” are used interchangeably and mean a compound that consist of two or more amino acids that are linked by means of peptide bonds.

[0018] As used herein, “levansucrase” means a protein, a protein fragment or peptide that has the property of synthesizing a carbohydrate polymer consisting of repeating fructose residues, using sucrose as a substrate. The repeating fructose residues may be linked by  $\beta$ -2-1 linkage or a  $\beta$ -2-6 linkage or any combination of the two linkage types. The polymer of repeating fructose units may contain one terminal glucose residue, derived from a sucrose molecule, and at least two fructose residues.

[0019] As used herein, “inducer” means a substance that interacts with a regulatory sequence, either directly or indirectly, to increase the rate of transcription of the nucleotide sequence controlled by the regulatory sequence.

[0020] LB means 10 g tryptone, 5 g yeast extract, 5 g NaCl and 1 ml 1N NaOH per liter H<sub>2</sub>O.

#### DETAILED DESCRIPTION

[0021] The following detailed description is provided to aid those skilled in the art in practicing the present invention. Even so, this detailed description should not be construed to unduly limit the present invention as modifications and variations in the embodiments discussed herein can be made by those of ordinary skill in the art without departing from the spirit or scope of the inventive discovery.

[0022] All publications, patents, patent applications and other references cited in this application are herein incorporated by reference in their entirety as if each publication, patent, patent application or other reference were specifically and individually indicated to be incorporated by reference.

[0023] The present invention provides a method for transforming plant cells and constructs and bacteria useful in said method. The invention involves the insertion of a sequence encoding a levansucrase, and in particular the sacB gene ORF, under strict control of an inducible regulatory sequence into a Gram-negative bacterium and in particular *Agrobacterium tumefaciens*. The sacB gene encodes levansucrase (sucrose 2,6, $\beta$ -D-fructan 6- $\beta$ -D-glucosyltransferase; (IC 2.3.1.10)), a 50 kD enzyme secreted by *B. subtilis* after induction by sucrose. Levansucrase catalyzes transfructosylation from sucrose to various acceptors. The two main physiological reactions resulting are, 1) levan synthesis and 2) sucrose hydrolysis. Since sucrose is the primary carbon source used in most plant tissue culture medium formulations, the present invention allows for the efficient counter selection of the bacterium without the use of antibiotic supplements.

[0024] One aspect provides a Gram-negative bacterium useful for the transfer of heterologous polynucleotide sequences into a host cell. The bacterium has as part of its genome a recombinant nucleic acid sequence comprising an inducible regulatory sequence operatively linked to a nucleotide sequence encoding the enzyme levansucrase. Alternatively, the bacterium comprises a recombinant nucleic acid sequence comprising an inducible regulatory sequence other than sacR, operatively linked to a nucleotide sequence encoding the enzyme levansucrase. A nucleic acid sequence is operatively linked when it is placed into a functional relationship with another nucleic acid sequence. For example, DNA for a presequence or secretory leader is operatively linked to a DNA element encoding for a polypeptide to be expressed as a preprotein which participates in the secretion of the polypeptide; a promoter or regulatory sequence is operatively linked to a coding sequence if it affects the transcription of the coding sequence; or a ribosome binding site is operatively linked to a coding sequence if it is positioned so as to facilitate translation. Any nucleic acid sequence encoding a protein, polypeptide or protein fragment that is functional as a levansucrase can be used. Numerous sequences encoding levansucrase are known in the art and can be found in publicly available databases such as those maintained by the National Center for Biotechnology Information. Representative examples include, without limitation, sequences encoding levansucrases obtained from *Acetobacter xylinus* (GenBank AB034152), *Gluconacetobacter diazotrophicus* (GenBank L41732), *Zymomonas mobilis* (GenBank L33402), *Paenibacillus polymyza* (GenBank AJ133737), *Rahnella aquatilis* (GenBank U91484), *Pseudomonas syringae* (GenBank AF052289), *Bacillus stearothermophilus* (GenBank U34874), *Bacillus subtilis* (GenBank X02730) and *Bacillus amyloliquefaciens* (GenBank X52988). Copies of the GenBank records for the listed accession numbers are attached and are to be considered part of this application. In one embodiment, the polynucleotide sequence encoding levansucrase enzyme is the coding region of the sacB gene of *Bacillus subtilis*. As used herein, the term “coding region” refers to the nucleotide sequence of a gene that is translatable into a polypeptide. Methods for producing artificial nucleotide sequences such as by cloning or nucleotide synthesis are well known in the art. Such sequences are included within the scope of the invention as long as they encode a biological equivalent to a levansucrase.

[0025] The regulatory sequences can consist of an inducible promoter, in combination with an operator sequence. As used herein, the term “operator” or “operator sequence” refers to a polynucleotide sequence to which a repressor protein can bind, thereby regulating the expression of a gene. Any inducible promoter that is functional within Gram-negative bacterium can be used. It is preferred that the promoter or combination of the promoter and operator be strictly inducible so that there is little or no production of levansucrase in the absence of the inducing agent. In one preferred embodiment, the regulatory sequence is one that is functional in members of the genus *Agrobacterium*, and in particular *A. tumefaciens*. Examples of suitable regulatory sequences include, but are not limited to, the Plac promoter and operator of *E. coli* (SEQ ID NO: 1), the *nocR* gene (SEQ ID NO: 3), which encodes for the transcriptional activator of *Pi2* (*noc*) (SEQ ID NO: 2), and in the presence of the *nocI* operon (SEQ ID NO: 2) which encodes for the nopaline transport system of *A. tumefaciens* (Von Lintig et al. (1991) *Molec. Plant Microbe Interaction*, 4:370-378) and the  $P_{BAD}$  promoter (SEQ ID NO: 5) and *araC* operator (SEQ ID NO: 4) of *E. coli* (Gallegos et al. (1997) *Microbiol. Mol. Biol. Rev.* 61:393-410).

[0026] Another aspect provides a recombinant nucleic acid construct comprising an inducible regulatory sequence operatively linked to a nucleotide sequence encoding a levansucrase. Any previously mentioned regulatory sequence or levansucrase encoding sequences can be used, although other suitable sequences will be apparent to those of ordinary skill in the art and are considered within the scope of the present invention. As used herein, “recombinant construct” is defined either by its method of production or its structure. In reference to its method of production, e.g., a product made by a process, the process is use of recombinant nucleic acid techniques involving human intervention in the nucleotide sequence. Alternatively, in terms of structure, it can be a sequence comprising fusion of two or more nucleic acid sequences which are not naturally contiguous or operatively linked to each other.

[0027] The recombinant constructs of the present invention are produced using methods well known to those of ordinary skill in the art which can be found, for example, in standard texts such as Sambrook et al. *Molecular Cloning*, 2<sup>nd</sup> ed., Cold Spring Harbor Laboratory Press, 1989 and Ausubel, et al. *Short Protocols in Molecular Biology*, 3<sup>rd</sup> ed., Wiley & Sons, 1995. In general, recombinant constructs are produced by a series of restriction enzyme digestions and ligation reactions which result in the sequences being assembled in the desired configuration. If suitable restriction sites are not available, alternative strategies, for example, the use of synthetic oligonucleotide linkers and adaptors, which are well known to those skilled in the art and described in the references cited above, can be employed to assemble the desired recombinant constructs.

[0028] It will be apparent to those of ordinary skill in the art, that the precise restriction enzymes, linkers and/or adaptors required as well as the precise reaction conditions will vary with the sequences and cloning strategies used. The assembly of recombinant constructs, however, is routine in the art and can be readily accomplished by the skilled technician without undue experimentation. Non-limiting

illustrations of the assembly of recombinant constructs useful in the present invention can be found in the examples that follow.

[0029] Once made, the recombinant constructs can be inserted into the genome of Gram-negative bacterium or introduced separately on a self replicating plasmid of a Gram-negative bacterium used in transforming host cells. In one embodiment, the Gram-negative bacterium is *Agrobacterium tumefaciens*. The inventors have found that insertion of the recombinant construct into the genome of the bacterial vector is superior to the method of placing the construct in *A. tumefaciens* via a self replicating plasmid. As shown in Example 1, it was discovered that a self replicating plasmid carrying the *sacB* coding region was frequently lost resulting in overgrowth of the bacterial cells during plant regeneration.

[0030] Any method capable of introducing the construct into the genome of the bacterial vector can be used. In one embodiment, the construct is inserted by the use of homologous recombination in particular the method of Ruvkun and Ausubel ((1981) *Nature*, 289:85-88). In this method, a mutation, in the form of the recombinant construct of the present invention, is directed to a specific locus on the chromosome by homologous exchange recombination. Any locus which allows the inducible expression of levansucrase and does not impede with the DNA transfer machinery can be used. In one embodiment, the construct is inserted at the *tetR/tetA* (SEQ ID NO: 9) loci of *Agrobacterium* (Luo and Farrand (1999) *J. Bacteriol* 181:618-626).

[0031] Another aspect provides method for transforming a plant cell using an *Agrobacterium tumefaciens* bacterium of the present invention as a vector. In general the method involves obtaining an *A. tumefaciens* strain whose genome includes a nucleotide sequence encoding a levansucrase operatively linked to an inducible regulatory sequence as described above or an *A. tumefaciens* strain that contains a recombinant nucleotide sequence encoding a levansucrase operatively linked to an inducible regulatory sequence as described above. The nucleotide sequence(s) of interest that are to be transferred to the plant cell can be inserted within the T-DNA element and introduced either directly to the resident Ti plasmid or separately using a binary plasmid strategy. Methods for the introduction of exogenous nucleotide sequences into the T-DNA element and the use of the derived *Agrobacterium* transconjugant to transform plant cells are well known in the art (see, Maliga et al. *Methods in Plant Molecular Biology*, Cold Spring Harbor Laboratory Press, 1995). Once the sequence to be delivered to the plant cell is assembled into the T-DNA element and introduced either directly to the resident Ti plasmid or via the binary vector strategy to the *Agrobacterium* of the present invention, the bacterium is subsequently used to inoculate plant cells either by direct injection or by co-cultivating the bacterium of the present invention with individual plant cells or pieces of plants such as leaf discs. Co-cultivation is carried out in medium supplemented with a carbon source, preferably glucose, for a sufficient amount of time to allow the T-DNA element to be mobilized from the bacterium to the plant cell genome. The co-cultivation period is determined empirically, but generally ranges from one to seven days. Co-cultivation periods may vary for a particular plant species, but determinations are routine in the art and can be made by one of ordinary skill in the art without undue

experimentation. Following co-cultivation, the transforming bacteria are counter selected prior to the regeneration of the plant cells to whole plants. Typically, the transforming bacteria are removed by using antibiotic supplements to the regeneration medium. In the present invention, however, an inducing agent that activates the promoter linked to the levansucrase coding region is added to sucrose amended regeneration medium. Activation of the inducible promoter results in the production of levan which causes the lysis of the *Agrobacterium* cells, thus providing efficient counter selection strategy. As a result, the regeneration of whole plants from the inoculated plant cells can be carried out, in the absence of antibiotics typically used to counter select *Agrobacterium* cells, following standard protocols (see, Maliga et al. supra).

[0032] An alternative embodiment provides an *Agrobacterium tumefaciens* vector in which the nucleotide sequence encoding a levansucrase operatively linked to an inducible regulatory sequence is contained in the *Agrobacterium* as part of a binary vector system. Binary vector systems and their constructions are well known in the art and are described, for example, in Maliga et al. supra and *The Encyclopedia of Molecular Biology*, J. Kendrew, ed., Blackwell Science, 1994). The resultant *Agrobacterium* strain can subsequently be used to genetically engineer plant cells as described above.

## EXAMPLES

### Example 1

Counter Selection Using sacB in a Binary Vector System

[0033] The nptI-sacB-sacR loci from the vector pUM24 (Reid and Collmer (1987) *Gene* 57:239-246) was subcloned as a BamHI fragment into the binary vector pZP112 (Hajdukiewicz et al. (1994) *Plant Molec. Biol.* 25:989-994), at the BclI site just outside the left border region. The resultant counter selection (suicide) vector was referred to as pPTN114. The binary vector was mobilized into *A. tumefaciens* strains C58C1 (Koncz and Schell (1986) *Mol. Gen. Genet.* 204:383-396) and EHA105 (Hood et al (1993) *Transgenic Research* 2:208-218) via tri-parental mating (Ditta et al (1980) *Proc. Natl. Acad. Sci USA* 77:7347-7351). The C58C1 transconjugant was used for subsequent evaluation in tobacco transformations, while the EHA105 transconjugant was tested in soybean transformations.

[0034] Culturing of the C58C1 and EHA105 carrying the binary vector pPTN114 on LB medium supplemented with 1% to 3% sucrose was lethal, while LB medium supplemented with 1.5% glucose was conducive to cell growth. Tobacco and soybean transformations were initiated to evaluate the efficacy of incorporating the sacB locus in a binary vector as a counter selection strategy for *A. tumefaciens* in the absence of antibiotics. In both plant species, *Agrobacterium* growth was impeded for approximately 10 days, at which point bacterial growth was observed on the plant tissue culture medium containing 3% sucrose. The bacteria from the tissue culture failed to grow when transferred onto LB medium supplemented with kanamycin. These results suggested that the *Agrobacterium* cells had lost the binary vector.

### Example 2

Counter Selection by Genomic Incorporation of sacB

[0035] A construct was assembled that specifically targeted the sacB locus to the tetR/tetA (SEQ ID NO: 9) loci in *A. tumefaciens*. An EcoRI fragment from pSWE8.5 bearing the tetR/tetA loci (SEQ ID NO: 9) was subcloned into pGEM T-Easy. The resulting plasmid was digested with Hind III and a 0.5 kb fragment at the tetR locus was replaced with the 3.8 kb BamHI insert from pUM24 containing the nptI-sacB-sacR region (Reid and Collmer (1987) *Gene* 57:239-246). This step was accomplished by annealing after adding homopolymeric G and C tails to the vector and insert, respectively. The resulting construct was electroporated into *Agrobacterium* strain NT1/pEHA105 and transformants were selected on kanamycin (50 mg/L) LB plates. Since the pGEM backbone was not expected to replicate in *A. tumefaciens*, kanamycin resistant transformants were presumed to be due to chromosomal integration of the nptI-sacB-sacR cassette at the tetR/tetA loci (SEQ ID NO: 9). Individual colonies were picked and replica plated to kanamycin supplemented LB medium with and without sucrose (3% w/v). The resulting bacterial patches that showed little or no growth on sucrose, but vigorous growth on kanamycin alone, were recovered, diluted and spread on kanamycin plates to isolate individual colonies. These were again tested as before until an isolate that consistently gave no growth on sucrose was recovered.

### Example 3

Use of the *E. coli* Plac/Repressor System to Control sacB Expression

[0036] The *E. coli* lactose operon (SEQ ID NO: 1) is tightly regulated by the presence of a 21 bp operator that resides immediately down stream of the Plac promoter (SEQ ID NO: 1). In the absence of  $\beta$ -galactose sugar, the lac repressor will bind to the cis operator and prevent RNA polymerase initiation. In the presence of a  $\beta$ -galactose sugar, the lac repressor cannot bind to the operator and thus RNA polymerization proceeds. This system may be exploited as a strategy for tight regulation of the sacB expression in *Agrobacterium tumefaciens* cells.

[0037] The sacB open reading frame (ORF) can be subcloned downstream of the Plac promoter element (SEQ ID NO: 1) coupled with the 21 bp operator sequence. The lacI cassette (SEQ ID NO: 6) coding for the lac repressor may be ligated to the derived sacB cassette. The genetic element carrying the Plac-sacB and lac repressor cassettes can be introduced to the chromosome of *Agrobacterium tumefaciens* via homologous recombination. The preferred site for the recombination event would be the tetR/tetA loci (SEQ ID NO: 9) of *Agrobacterium* recently described by Luo and Farrand ((1999) *J. Bacteriol.* 181:618-626).

[0038] This strategy will be useful for the genetic engineering of both monocotyledonous and dicotyledonous plant species. Various steps are followed in the *Agrobacterium*-mediated transformation of plant species. Generally the first step involves the inoculation of the explant (plant cells or tissue segments) with *Agrobacterium tumefaciens* cells. The explant can be leaf segment, cotyledon, stem, root, flower part or cells thereof. After a period of one to seven days, generally termed the co-cultivation period, the explant is

transferred to plant regeneration medium supplemented with sucrose as the carbon source. The *sacB* system in this example may be induced by the addition of isopropyl- $\beta$ -D-thiogalactoside (IPTG) at levels ranging from 0.1  $\mu$ M up to 1 mM and/or lactose at levels from 0.1  $\mu$ M to 1 mM.

#### Example 4

Use of the Nopaline-Inducible Marker System to Regulate *sacB* Expression

[0039] The Pi2(noc) promoter (SEQ ID NO: 2) (Von Lintig et al. (1991) *Molec. Plant Microbe Interaction* 4:370-378) from *Agrobacterium tumefaciens* is induced in the presence of nopaline. This regulatory sequence may be employed as a strategy to regulate *sacB* expression in *Agrobacterium tumefaciens* cells. In this example, the *sacB* ORF will be subcloned down stream of the Pi 2(noc) promoter (SEQ ID NO: 2). The resultant cassettes preferably will be introduced into the chromosome of *Agrobacterium tumefaciens* via homologous recombination. The preferred site for recombination is the tetR/tetA loci (SEQ ID NO: 9). The *nocR* gene (SEQ ID NO: 3), which encodes for the transcriptional activator of pi2(noc) (SEQ ID NO: 2), and the *nocI* operon (SEQ ID NO: 2) which encodes for the nopaline transport system can be supplied by the *vir* region of the resident Ti plasmid, or by cloning these two loci onto a self-replicating plasmid. In this example the *sacB* counter selection (suicide) system may be induced upon the addition of nopaline at levels ranging from 50  $\mu$ g/L up to 200  $\mu$ g/L.

#### Example 5

Use of the *E. coli* *araC* Regulator to Control *sacB* Expression

[0040] The *E. coli* P<sub>BAD</sub> promoter (SEQ ID NO: 5) is highly induced in the presence of L-arabinose. The induction of the system is controlled by the presence of a *cis* acting element upstream of the P<sub>BAD</sub> promoter (SEQ ID NO: 5), *araC* (SEQ ID NO: 4) (Gallegos et al. (1997) *Microbiol. Molec. Biol. Rev.* 61:393-410). This system can be utilized to regulate the expression of the *sacB* in *Agrobacterium tumefaciens* cells by placing the *araC* *cis* element (SEQ ID NO: 4) just 5' to the P<sub>BAD</sub> promoter (SEQ ID NO: 5) (Luo and Farrand (1999) *J. Bacteriol.* 181:618-626); Newman and Fuqua (1999) *Gene* 227:197-203) and subsequently subcloning the *sacB* open reading frame downstream of the assembled DNA elements. In this example, counter selection of the bacterial cells can be induced following the co-cultivation period in plant transformation protocols by supplementing the regeneration medium with levels of L-arabinose ranging from 5 g/L up to 20 g/L.

#### Example 6

Use of the traCDG Promoter to Control *sacB* Expression

[0041] The *sacB* gene can be placed under the control of the traCDG promoter (SEQ ID NO: 7) (Farrand et al., (1996) *Bacteriol.* 178:4233-4247; Oger et al., (1998) *Mol. Microbiol.* 27:277-288; Luo and Farrand, (1999) *Proc. Natl. Acad. Sci. USA* 96:9009-9014). Initiation of this promoter is absolutely dependent upon activated TraR (SEQ ID NO: 7). Second, the TraR (SEQ ID NO: 7) will be placed under the direct control of an opine-responsive promoter system. We will fuse traR (SEQ ID NO: 7) directly to a fragment of DNA containing the *occ* promoter (SEQ ID NO: 8) from the

octopine-type Ti plasmid pTiR10. This promoter is activated by OccR (SEQ ID NO: 8), a *lysR*-like activator in response to the opine, octopine (Habeeb, et al., 1991). The *occR* gene (SEQ ID NO: 8) is located directly adjacent to the *occ* promoter (SEQ ID NO: 8) and will be included in the recombinant construct. One can provide a copy of traM (SEQ ID NO: 7), which encodes for the antiactivator, in the traR-pOcc-*occ* construct. When the two constructs are combined in an *Agrobacterium* host, expression of *sacB* should be strongly suppressed in the absence of the opine. However, addition of octopine (which will activate expression of traR (SEQ ID NO: 7) leading to accumulation of the activator to levels that overcome the antiactivator, TraM (SEQ ID NO: 7)) should strongly induce *sacB*.

[0042] The genetic elements described above can be introduced to a neutral site in the chromosome of *Agrobacterium tumefaciens*, with respect to plant transformation effects, via homologous recombination. The preferred site for the recombination event would be the tetR/tetA loci (SEQ ID NO: 9) of *Agrobacterium*. However, with the imminent availability of the genome sequence of *Agrobacterium* strain C58 one should be able to identify alternative sites within the chromosome.

#### Example 7

Use of a Second Copy of the *sacB* Gene

[0043] To circumvent potential mutational inactivation of *sacB* a second copy of the gene can be provided. To prevent recombination between the two copies one can use a *sacB* gene from another bacterium. Possibilities of alternative *sacB* sources include *B. stearothermophilus*, *B. amyloliquefaciens*, or *Streptococcus mutans*. The sequence for each of these is available in the data bases. The alternative *sacB* gene can be fused to a second copy of the TraR-dependent traCDG promoter (SEQ ID NO: 7).

#### CONCLUSION

[0044] In light of the detailed description of the invention and the examples presented above, it can be appreciated that several aspects of the invention are achieved.

[0045] It is to be understood that the present invention has been described in detail by way of illustration and example in order to acquaint others skilled in the art with the invention, its principles, and practical application. Particular formulations and processes of the present invention are not limited to the descriptions of the specific embodiments presented, but rather the descriptions and examples should be viewed in terms of the claims that follow and their equivalents. While some of the examples and descriptions above include some conclusions about the way the invention may function, the inventors do not intend to be bound by those conclusions and functions, but put them forth only as possible explanations.

[0046] It is further understood that the specific embodiments of the present invention as set forth are not intended as being exhaustive or limiting of the invention, and that many alternatives, modifications, and variations will be apparent to those of ordinary skill in the art in light of the foregoing examples and detailed description. Accordingly, this invention is intended to embrace all such alternatives, modifications, and variations that fall within the spirit and scope of the following claims.

## SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 9

<210> SEQ ID NO 1

<211> LENGTH: 7477

<212> TYPE: DNA

<213> ORGANISM: Escherichia coli

<400> SEQUENCE: 1

```
gacaccatcg aatggcgcaa aacctttcgc ggtatggcat gatagcgccc ggaagagagt    60
caattcaggg tggtagaatgt gaaaccagta acggtatacg atgtcgcaga gtatgccgggt    120
gtctcttata agaccgtttc ccgctgtggtg aaccaggcca gccacgtttc tgcgaaaacg    180
cgggaaaaag tggaagcggc gatggcggag ctgaattaca ttcccaaccg cgtggcacia    240
caactggcgg gcaaacagtc gttgctgatt ggcggtgcca cctccagtct ggcctgcac    300
gcgcccgcgc aaattgtcgc ggcgattaaa tctcgcgcgc atcaactggg tgcagcgtg    360
gtggtgtcga tggtagaacg aagcggcgtc gaagcctgta aagcggcggg gcacaatctt    420
ctcgcgcaac gcgctcagtg gctgatcatt aactatccgc tggatgacca ggatgccatt    480
gctgtggaag ctgcctgcac taatgttccg gcgttatctt ttgatgtctc tgaccagaca    540
cccatcaaca gtattatctt ctcccatgaa gacggtagcgc gactgggctg ggagcatctg    600
gtcgcattgg gtcaccagca aatcgcgctg ttagcgggcc cattaagtct tgtctcggcg    660
cgtctcgcgc tggctggctg gcataaata tctcactcgc atcaaattca gccgatagcg    720
gaacgggaag gcgactggag tgccatgtcc ggttttcaac aaaccatgca aatgctgaat    780
gagggcatcg ttcccactgc gatgctggtt gccaacgata agatggcgtc ggcgcgaatg    840
cgcgccatta ccgagtcagg gctgcgcggtt ggtgcggata tctcggtagt gggatacgc    900
gataccgaag acagctcatg ttatatcccc ccgtcaacca ccatcaaaaca ggattttcgc    960
ctgctggggc aaaccagcgt ggaccgcttg ctgcaactct ctcagggcca ggcggtgaa    1020
ggcaatcagc tgttgcccgt ctactggtg aaaagaaaaa ccaccctggc gcccaatacg    1080
caaaccgcct ctccccgcgc gttggccgat tcattaatgc agctggcacg acaggtttcc    1140
cgactggaaa gcgggcagtg agcgcacgc aattaatgtg agttagctca ctcattaggc    1200
accccaggct ttacacttta tgcttccggc tcgtatgttg tgtggaattg tgagcgata    1260
acaatttcac acaggaaca gctatgacca tgattacgga ttcaactggc gtcggtttac    1320
aacgctgtga ctgggaaaac cctggcgcta cccaacttaa tcgccttgca gcacatcccc    1380
ctttcgccag ctggcgtaat agcgaagagg cccgcaccga tcgccttcc caacagttgc    1440
gcagcctgaa tggcgaatgg cgctttgcct ggtttccggc accagaagcg gtgocgaaa    1500
gtggtctgga gtgcgatctt cctgagggcg atactgtcgt cgtcccctca aactggcaga    1560
tgcacggtta cgatgcgccc atctacacca acgtaaccta tcccattagc gtcaatccgc    1620
cgtttgttcc cacggagaat ccgacgggtt gttactcgtc cacatttaat gttgatgaaa    1680
gtggtctaca ggaaggccag acgcgaatta tttttgatgg cgtaactcg gcgtttcatc    1740
tgtggtgcaa cgggcgctg gtcggttacg gccaggacag tcgtttgccg tctgaatttg    1800
acctgagcgc atttttacgc gccggagaaa accgcctcgc ggtgatggtg ctgcgttggg    1860
gtgacggcag ttatctgtaa gatcaggata tgtggcggat gagcggcatt ttccgtgacg    1920
```

## -continued

---

tctcgttgct	gcataaaccc	actacacaaa	tcagcgattt	ccatgttgcc	actcgcctta	1980
atgatgattt	cagccgcgct	gtactggagg	ctgaagtcca	gatgtgcggc	gagttgcgtg	2040
actacctacg	ggtaacagtt	tctttatggc	agggtgaaac	gcaggtcgcc	agcggcaccg	2100
cgcccttcgg	cggtgaaatt	atcgatgagc	gtggtggtta	tgccgatcgc	gtcacactac	2160
gtctgaacgt	cgaaaacccc	aaactgtgga	gcgccgaaat	cccgaatctc	tatcgtgcgg	2220
tggttgaact	gcacaccgcc	gacggcacgc	tgattgaagc	agaagcctgc	gatgtcggtt	2280
tccgcgaggt	gcggattgaa	aatggtctgc	tgctgctgaa	cggcaagccg	ttgctgattc	2340
gaggcgtaa	ccgtcacgag	catcatcctc	tgcatggtca	ggtcatggtg	gagcagacga	2400
tgggtcagga	tatcctgctg	atgaagcaga	acaactttaa	cgccgtgcgc	tgcttcgcatt	2460
atccgaacca	tccgctgtgg	tacacgctgt	gcgaccgcta	cggcctgtat	gtggtggatg	2520
aagccaatat	tgaaacccac	ggcatggtgc	caatgaatcg	tctgaccgat	gatccgcgct	2580
ggctaccggc	gatgagcgaa	cgcgtaacgc	gaatggtgca	gcgcgatcgt	aatcacccca	2640
gtgtgatcat	ctggtcgtcg	gggaatgaat	cagggccacg	cgctaatac	gacgcgctgt	2700
atcgtctgat	caaatctgtc	gacccctccc	gcccggtgca	gtatgaaggc	ggcggagccg	2760
acaccacggc	caccgatatt	atctgcccga	tgtacgcgcg	cgtggatgaa	gaccagccct	2820
tcccggctgt	gccgaaatgg	tccatcaaaa	aatggctttc	gctacctgga	gagacgcgcc	2880
cgctgatcct	ttgcgaatac	gcccacgcga	tgggtaacag	tcttggcggg	ttcgctaaat	2940
actggcaggc	gtttcgtcag	tatccccggt	tacagggcgg	cttcgctctg	gactgggtgg	3000
atcagtcgct	gattaaatat	gatgaaaacg	gcaaccctgt	gtcggcttac	ggcggtgatt	3060
ttggcgatac	gccgaacgat	cgccagttct	gtatgaacgg	tctggtcttt	gccgaccgca	3120
cgccgcatcc	agcgtcgacg	gaagcaaac	accagcagca	gtttttccag	ttccgtttat	3180
ccgggcaaac	catcgaagtg	accagcgaat	acctgttccg	tcatagcgat	aacgagctcc	3240
tgcaactgat	ggtggcgtcg	gatggtaagc	cgctggcaag	cgggtaagtg	cctctggatg	3300
tcgctccaca	aggtaaacag	ttgattgaac	tgctgaaact	accgcagccg	gagagcgcgg	3360
ggcaactctg	gctcacagta	cgcgtagtgc	aaccgaacgc	gaccgcatgg	tcagaagccg	3420
ggcacatcag	cgctctggag	cagtggcgtc	tggcggaaaa	cctcagtggt	acgctccccg	3480
ccgcgtccca	cgccatcccc	catctgacca	ccagcgaat	ggatttttgc	atcgagctgg	3540
gtaataagcg	ttggcaattt	aaccgccagt	caggctttct	ttcacagatg	tggattggcg	3600
ataaaaaaca	actgctgacg	ccgctgcgcg	atcagttcac	ccgtgcaccg	ctggataacg	3660
acattggcgt	aagtgaagcg	accgcattg	accctaacgc	ctgggtcgaa	cgctggaagg	3720
cgccgggcca	ttaccaggcc	gaagcagcgt	tgttgcagtg	cacggcagat	acaactgtctg	3780
atgcggtgct	gattacgacc	gctcacgcgt	ggcagcatca	ggggaaaacc	ttatttatca	3840
gccggaaaac	ctaccggatt	gatggtagtg	gtcaaatggc	gattaccggt	gatggtgaag	3900
tggcgagcga	tacaccgcat	ccggcgcgga	ttggcctgaa	ctgccagctg	gcgcaggtag	3960
cagagcgggt	aaactggctc	ggattagggc	cgcaagaaaa	ctatcccggc	cgcttactg	4020
ccgcctgttt	tgaccgctgg	gatctgccat	tgctcagacat	gtataccccg	tacgtcttcc	4080
cgagcgaaaa	cggctctgcg	tgcgggacgc	gcgaattgaa	ttatggccca	caccagtggc	4140
gcggcgactt	ccagttcaac	atcagccgct	acagtcaaca	gcaactgatg	gaaaccagcc	4200

-continued

---

atcgccatct gctgcacgcg gaagaaggca catggctgaa tatcgacggt ttccatatgg	4260
ggattggtgg cgacgactcc tggagcccgt cagtatcggc ggaattccag ctgagcggcg	4320
gtcgctacca ttaccagttg gtctggtgtc aaaaataata ataaccgggc aggccatgtc	4380
tgcccgtatt tcgcgtaagg aaatccatta tgtactattt aaaaaacaca aacttttggg	4440
tgttcggttt attctttttc ttttaacttt ttatcatggg agcctaactc cggtttttcc	4500
cgatttggtc acatgacatc aaccatatca gcaaaagtga tacgggtatt atttttgccc	4560
ctatttctct gttctcgcta ttattccaac cgctgtttgg tctgctttct gacaaaactc	4620
ggctgcgcaa atacctgctg tggattatta cggcatggt agtgatgttt gcgccgttct	4680
ttatttttat ctctgggcca ctgttacaat acaacatttt agtaggatcg attggtgggt	4740
gtatttatct aggcctttgt tttaacgccg gtgcgccagc agtagaggca tttattgaga	4800
aagtcagccg tcgcagtaat ttcgaatttg gtcgcgcgcg gatgtttggc tgtgttggtc	4860
ggcgctgtg tgcctcgatt gtcggcatca tgttcacat caataatcag tttgttttct	4920
ggctgggctc tggctgtgca ctcatcctcg ccgttttact ctttttcgcc aaaacggatg	4980
cgccctcttc tgccacggtt gccaatgccc taggtgccaa ccattcggca tttagcetta	5040
agctggcact ggaactgttc agacagccaa aactgtggtt tttgtcactg tatgttattg	5100
gcgtttcctg cacctacgat gtttttgacc aacagtttgc taatttcttt acttcgttct	5160
ttgctaccgg tgaacagggg acgcggtgat ttggctacgt aacgacaatg ggcgaattac	5220
ttaacgcctc gattatgttc tttgcgccac tgatcattaa tcgcatcggg gggaaaaacg	5280
ccctgctgct ggctggcact attatgtctg tacgtattat tggctcatcg ttcgccacct	5340
cagcgctgga agtggttatt ctgaaaacgc tgcataatgt tgaagtaccg ttcctgctgg	5400
tgggctgctt taaatatatt accagccagt ttgaagtgcg tttttcagcg acgatttattc	5460
tggctgttt ctgcttcttt aagcaactgg cgatgatttt tatgtctgta ctggcgggca	5520
atatgtatga aagcatcggg ttccagggcg cttatctggt gctgggtctg gtggcgtggt	5580
gcttcacctt aatttccgtg ttcacgctta gcggccccg cccgctttcc ctgctgcgtc	5640
gtcaggtgaa tgaagtcgct taagcaatca atgtcggatg cggcgcgacg cttatccgac	5700
caacatatca taacggagtg atcgcattga acatgccaat gaccgaaaga ataagagcag	5760
gcaagctatt taccgatatg tgcgaaggct taccggaaaa aagacttctg gggaaaaact	5820
taatgtatga gtttaatcac tcgcatccat cagaagtga aaaaagagaa agcctgatta	5880
aagaaatggt tgccacgcta ggggaaaacg cctgggtaga accgctgtc tatttctctt	5940
acggttccaa catccatata ggcgcaatt tttatgcaa tttcaattta accattgtcg	6000
atgactacac ggtaacaatc ggtgataacg tactgattgc acccaacgtt actctttccg	6060
ttacgggaca ccctgtacac catgaattga gaaaaaacg cgagatgtac tcttttccga	6120
taacgattgg caataacgct tggatcggaa gtcatgtggt tattaatcca ggcgtcacca	6180
tcggggataa ttctgttatt ggcgcggtta gtatcgtcac aaaagacatt ccaccaaacg	6240
tcgtggcgcg tggcgcttct tgcgggtta ttcgcgaaat aaacgaccgg gataagcact	6300
attatttcaa agattataaa gttgaatcgt cagtttaaat tataaaaatt gcctgatagc	6360
ctgcgcttat caggcttaca agttcagcga tctacattag ccgcatccgg catgaacaaa	6420
gcgcaggaac aagcgtcgca tcatgcctct ttgacccaca gctgcgaaa acgtactggt	6480

-continued

---

```

gcaaaacgca gggttatgat catcagccca acgacgcaca gcgcatgaaa tgcccagtc 6540
atcaggtaat tgccgctgat actacgcagc acgccagaaa accacggggc aagcccggcg 6600
atgataaaac cgattccctg cataaacgcc accagcttgc cagcaatagc cggttgcaca 6660
gagtgatcga gcgccagcag caaacagagc ggaaacgcgc cgcccagacc taaccacac 6720
accatcgccc acaataccgg caattgcacg ggcagccaga taaagccgca gaaccccacc 6780
agttgtaaca ccagcgccag cattaacagt ttgcgccgat cctgatggcg agccatagca 6840
ggcatcagca aagctcctgc ggcttgccc agcgtcatca atgccagtaa ggaaccgctg 6900
tactgcgcgc tggcaccaat ctcaatatag aaagcgggta accaggcaat caggctggcg 6960
taaccgccgt taatcagacc gaagtaaaaca cccagcgtcc acgcgcgggg agtgaatacc 7020
acgcaaacgc gagtggttgt tgtcttgagg gaagaggcga cctcgcgggc gctttgccac 7080
caccaggcaa agagcgcaac aacggcaggc agcgcacca ggcgagtggt tgataccagg 7140
tttcgctatg ttgaactaac cagggcgtta tggcggcacc aagcccaccg ccgccatca 7200
gagccgcgga ccacagcccc atcaccagtg gcgtgcgctg ctgaaaccgc cgtttaatca 7260
ccgaagcatc accgcctgaa tgatgccgat cccccacca ccaagcagtg cgctgctaag 7320
cagcagcgca ctttgcgggt aaagctcacg catcaatgca ccgacggcaa tcagcaacag 7380
actgatggcg acaactgcgac gttcgtgac atgctgatga agccagcttc cggccagcgc 7440
cagcccgcgc atggtaacca ccggcagagc ggtcgcac 7477

```

&lt;210&gt; SEQ ID NO 2

&lt;211&gt; LENGTH: 6610

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Agrobacterium tumefaciens

&lt;400&gt; SEQUENCE: 2

```

accgtgccgc cttcataaca ttgaaacac gccgtatgcc tattactgcc cttgtagcat 60
ggcgcgcaac gccttatccc taagctggaa ggcaagaatt ccaagatagc cagatctgca 120
atagaccgcg ttcacatggt gatcatctgt tcaagactga tgaatcgaa gcaagaagg 180
tcctcaaaac ttggccgtaa aacggcgaca gctggacaca ctgtatagcc gaccgcacga 240
ggattgaaag atgcgtctgg ctactggag gctaatggcc acggatcga cgtttcgccc 300
gggcacggaa gccttcgtat tgtgagaagg ttaaattgat cagcagttcg agacagccat 360
tgttggcggc ggtgtagtgg gcgcccaat tggatcggg ttagccaagt tggggcgccg 420
ggtgatcatc atcgatggcg cggacgatgc gctgcgtgcg tcccgaacga atttcgcct 480
gatctggggt cagggcaagg gcgacgaata tcctgcata catcgatca cgcggcaaa 540
cgctgagctg tggcccgtt ttgcatccga gttgaaagaa ctgacgggtg tcgacgtgga 600
gtaccgccgc aacggtggtt tgctcttctg tgcgtcgaa gctcaggcag aagaagaaaa 660
agcgatcggc gagaagatgg ctgccattca tcctgactat gcgttcgaga tgctggagcg 720
ctcgatcga cgcagcttc ccgacatccg tctcgggccc aaggtgtgtt cagcgtcgtt 780
cagtcatatg gacggtgacg tcaatcctct gctggtcctt cgcgctatgc tggacggatt 840
tgtgcgcctc ggtggaacct tgatgacggg cgaacacgct gacagatcg ccccacgct 900
tggcggcttc cgactggcga cggatggcct cagatagag gcggagcggg tgattattgc 960
agcgggaaat gactccatgg cttttgccc gagcctggac cttccgcttc tcctcgtccc 1020

```

-continued

---

ggaaaagggg	cagcttctca	tcacggaacg	ggtcgccccg	gttttccctt	acgccgccag	1080
cggtatccgt	cagaccatga	cgggcacggt	tcagctcggc	gttacgaatg	aacgtgtggg	1140
ccggtctacc	gacgtgacag	ctgccggcgc	acgtcataatc	gccaatcgcg	cgatcaatgt	1200
gatgcctgac	gtcggcgcgt	tgcgtgtcat	acggcagtgg	gcgggactac	gcgtgctgac	1260
cccggacggg	gtgcctgtct	acgatcggtc	gcggcgatac	ccaggatccc	acgtcgtcgc	1320
atgccattcc	ggcgtcacc	ttgccgtct	gcattgcagg	ccttttgccg	aatggttggc	1380
agccggtgga	agggagccat	cctacgacgc	tttccgccc	agcagattca	agatggggat	1440
cgcgcgctga	cgagtgaat	tctctttccc	tccggccaag	ccgtggaat	caccgtcgac	1500
ggtcgtctcg	tgaccgctcg	cgatggagag	cctctggccg	ccgtcttct	gcgtctggac	1560
gacatccata	cgcgaccc	ctttccttc	gccagcccc	ggcgcccta	ctgcatgatg	1620
ggcgtctgtt	tcgaatgtgt	cgtgcacgtc	gaagacatcg	gcacgatccg	aagctgtcag	1680
caaacggttc	acgcccgc	gagggtcgaa	cgccatgacc	atccgaagag	gctgaaatga	1740
accatcattc	gagcggcgc	ctgttgatcg	tcggagccgg	ccccgctggc	atggccgcag	1800
cgaggcgcgc	tgtagcggc	ggtctgtccg	tcacgtgct	ggattcgcaa	tcacagcctg	1860
gaggtcagat	ctggcgcaat	gccggccgca	acgtacaag	cccaatcata	aacattctgg	1920
ggcggaata	tcgacggcgc	gtccggcaag	tcgaagcatt	tctggcctgt	ggcgccgact	1980
acatcccga	tgctcagga	agcaggctca	gccagggttg	gacggtgga	tacgtctggg	2040
gcggagaaat	tcgtcccgc	aggggacggc	atctcttgc	cgctacggc	gcacaggaac	2100
gcccggttcc	gttcaccggc	tgacgctac	ccggcgtcat	gacggttggc	gccgcacaga	2160
tcctgttga	aacggcagg	cagttgccga	gaggccccgt	tgccgtcgtc	ggcagcgggc	2220
ctctcccgt	gctatatatg	cagcagatgc	ggctcgggg	ggcaaagccg	gtcgtcctc	2280
tcgatacgac	ccctcgggg	ctcatcagcc	gttcgctcgc	cggttttcgc	ggagcactgc	2340
aggagcctgg	ccaaatcctg	aaaggattg	cctggttacc	gcagttcagt	ggcgttcggc	2400
acgtacgcaa	cggtgtcaaa	atcagccca	agggaaagcg	acgactgga	actgtacgg	2460
tcgaaacttc	aagcgaag	tccgacagat	tgagggtcaa	atcccttctt	gttcacgagg	2520
ggctggtgcc	aagccatcag	ttggcgtgt	cgccggcgc	gcaactgatg	tggtatgtcg	2580
gccagtcagc	gttcggcct	gagcgcgacg	agtggatgaa	cgagggtccg	gacggactat	2640
acatagccgg	ggacggtcgc	cgatcggag	gcgccgtcaa	tgccgagatc	gagggagaga	2700
tcgcccgat	cgccgtcctt	cttcgtgacg	gcgccctgac	ggccgctcag	gcggataccg	2760
agacaaaacc	gctgagagcg	cagcatggcc	ggcaaaagtc	gttcggcgca	ttctgggatg	2820
gggcctacc	gccgaatatt	gcgcggacc	cgccggacga	cgcaacgatc	gtctgccgg	2880
gcgaggagtt	gagcgcggc	gtgttgagg	aagctgctgt	ccgtggcgt	tgccgaggcc	2940
ccaaccagct	gaagagtttt	acgcgcggc	gcattgggtc	ctgtcagga	aggcagtcgc	3000
gctaccgggt	gcacgagctg	gtgaagtca	tcgccggctg	accgcccggc	aggtcggcct	3060
gttcaatccg	cgccctccct	tcgttcatt	aacggtgtcg	atgctggcag	ctcagcaggc	3120
cgaggaaagt	gtttcagcgt	cgacacatga	gtgatcgttc	ggccatcgca	actgtgcgg	3180
gaccaaagtt	gcagtcaaga	aatgaacgg	gctggcgaaa	tcaacgcttc	gcggcatcgc	3240
aaggagaatg	agttgaagac	gtgcaaatc	ctgggagctc	ccgttcaaag	tgccgcatcc	3300

-continued

---

caaccggat	gctgatggg	acctgatgcg	tttcggactg	ccggcttgac	gcaagttctg	3360
acggagctgg	gctgggctgt	caccgatctc	ggagatgcga	caccaacggt	cgaaccggaa	3420
ctcagccacc	ccaattccgc	ggtgaagaac	ctcgatgctc	tggtgggatg	gacgcgcagt	3480
ctgtcccaga	aagctctgga	gatggcccgc	agctgcgac	ttccggctct	tctcggcggc	3540
gatcaactcga	tgtctgtctg	caccgtctca	ggtgtggccc	aacgtacagc	cgagcttggc	3600
aaggagcaat	tcgtcctttg	gctggacgcg	catacggacc	tgcataccct	ccacacgacc	3660
gcgagcggca	atctccacgg	cacaccgcta	gcctactata	cgggccaatc	cggtctcgaa	3720
gggtgcccgc	cgctggcccgc	gcctgtaaat	ccccgcaacg	tatccatgat	ggggattcgc	3780
tcagtcgac	cggaagagag	gcgacgggtt	gccgagatcg	gtgttcaagt	cgctgacatg	3840
cgggttctgg	acgaacaagg	ggtcgtacgc	ccgctcgaag	cttttcttga	ccgcgtgagt	3900
aaggtcagcg	cgagattgca	cgtcagcctt	gatgtcgatt	tcctcgatcc	cgcgatcgcg	3960
ccagcagtgg	gcacgaccgt	tcctggcggg	gcgaccttcc	gggaagcgca	cctcatcatg	4020
gagatgctcc	atgacagcgg	ccttgtcacg	tactcagacc	tggcggagct	caatccgttt	4080
ctggatgaga	ggggcgccac	tgcccgcctc	ataaccgac	ttgcctcgag	cctattcggc	4140
cggcgcgtgt	tcgacagggg	gacaacagca	ttttgatcac	cgggtgttgc	ccggtgcgat	4200
cgaggtttgc	ctctcgcacc	gagacaaacc	cttcgaaacc	aggaggaaga	ggatgaaggt	4260
aggaatagcc	ggagcaggat	cgatcgccat	tgggtatgcg	gcgttcctct	tacagaatgg	4320
ccatgagcca	ttgatctggt	cgcgttcacg	cgaaagaact	gctgctttct	ctgccggcgc	4380
accggtgaag	atcacgggtg	cgatcagcac	agagttccac	ccgcataatc	gcacgaacgc	4440
cgaggaattg	accgaagccg	atgtgatcgt	actggcactg	cccgcctatg	gccaccgctt	4500
tgttctcgac	gctcttctgc	cgcacctcag	cgcacgtcat	tcggtgatta	tcagcgtcga	4560
tctttccttt	gcggcgcttt	atctcgccaa	gggactggcc	gagcgtggaa	tcgagatccc	4620
gatcacggca	tggagcacga	ctgtgctgac	ctgcaagccg	cgggcggcgg	atacgttcaa	4680
catcggggcc	atccggggca	aggtcgatat	ggcgcgggta	ccggcccgtt	tcgcagaacg	4740
ggcgcacgaa	acctgtgtcg	ccctgttcgg	cgatcgcttc	gacgtgaagg	acgacgtcct	4800
tacgatcgcga	ttgagcaatc	tcaatcccca	ggaccacatg	gggatcgcgc	tgtgcaatct	4860
gtcgcgtatc	gagcggggcg	aggaatgggg	tcagaacacc	aacgtcacgc	ctgccgtcgg	4920
gcggttcctt	gaggcgctcg	atctgaaaag	gctggcaatc	gcgtcacctt	cggcaaaacg	4980
gtccggtcga	ccttcgatca	cttcggattg	tcgttcgaca	tatccggcca	atccgtctcg	5040
gaaatatcat	cgatcctcgt	cgatcgcgga	agatccggca	ggcccgaaga	gccttgatac	5100
gcgctacgtg	ctcagggacg	ttccttttgg	tctgatcccc	accctgcac	tggcaaaagc	5160
agccggcctg	tctgcaccct	tgcatgaaaag	cggcgtgcgg	atcttaagcg	cctgctacgg	5220
acgcgacttc	tccgaagaaa	acgacctgct	tccagacctc	ggcccgtctg	atctgcccac	5280
gctcaagaaa	cgggtagtta	ccgggtacgc	gatcacggcc	gaaccctgtg	gatcggggct	5340
cgctccggcag	cctatcaaac	caataaaggg	ggccaaaaga	tgccctgact	tgccaacctc	5400
aacatcgttc	cgttcatcag	tgttgaaaac	atgatggatc	tggcggtcag	caccgggatc	5460
gagaacttcc	ttgttcaact	ggcggcttac	attgaagaag	atttccggcg	ctgggaaagc	5520
ttcgacaaga	ttccacgcat	cgctcgcgat	tcgcgggacg	gcgtgattga	actcatgccc	5580

## -continued

---

acgagcgacg	gaacgctcta	tggcttcaaa	tacgtcaacg	ggcatccgaa	aaacacgaag	5640
tcgggacggc	agacggtgac	cgcgttcggc	gtgctctccg	atgtcgatag	cggetatccg	5700
ctgctgctgt	cggaatgac	aatcctgacg	gccttgcgaa	cgcccgac	ctcccgatc	5760
gcggcaaatg	atctggcccc	taaagattcg	cgcacgatgg	cgctgatcgg	caacggtgcc	5820
caaagtgagt	tccaggcttt	ggccttcaag	gcgctgatcg	gcgtcgatcg	gatacgactt	5880
tacgacatcg	accocgaggc	tactgctcgc	tgctcccgca	atctccagcg	gttcgggttt	5940
cagatcgagg	cctgcacctc	tgcggaacag	gcggtcgaag	gtgccgatat	cattacgaca	6000
gcaacggccg	acgagcaca	cgccacgatc	ctgagcgaca	atatgatcgg	gcccgccgtc	6060
cacatcaacg	gcgtcggcgg	cgactgcccc	ggcaagaccg	agatgcaccg	cgatatcttg	6120
ctgctgtcgg	atattttcgt	ggagtccccg	ccacagaccc	ggattgaagg	cgagatccag	6180
caactcgcgc	cagatcatcc	tgtgacggaa	ctttggcgcg	tgatgactgg	ccaggacgtc	6240
ggccgaaaag	gcgacaagca	gatcacgatt	ttcgacagcg	tcggatttgc	gatcgaggat	6300
ttttcggcgc	tgcgctacgt	cccgacccgc	gttgagggat	caagccacag	cagcccactc	6360
gaccttctag	ccgaccacga	cgagccaagg	gatctttttg	gaatgctgct	ccgtcgtcag	6420
gctttccgac	gtttgggtgg	ttgaacagaa	gtcattatcg	tacggaatgc	cagcactccc	6480
gaggggaacc	ctgtggttgg	catgcacata	caaatggacg	aacggataaa	ccttttcacg	6540
cccttttaaa	tatccggtat	tctaataaac	gctcttttct	cttaggttta	cccgccaata	6600
tatcctgtca						6610

&lt;210&gt; SEQ ID NO 3

&lt;211&gt; LENGTH: 987

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Agrobacterium tumefaciens

&lt;400&gt; SEQUENCE: 3

taaaagcttg	ccttccat	cgcaaggcca	gtaataggca	tgcgccatg	tgtatttgtt	60
atggagaaa	caatgattca	atcgcgtcag	ttagaagcgt	tccgggcagt	catgctgaca	120
ggaggtatga	cgctcagcgc	gaatctggtg	aggatcacgc	agcccgcat	cagccggctg	180
atcagggatc	tcgaagagga	aattgggatc	agcctcttcg	aaagaacggg	caaccggtta	240
cgctctacgc	gggaggccgg	tattctgttc	aaggaagtgt	cgcgacattt	caacgggatt	300
cagcacatcg	acaaagtccg	ggctgaactg	aagaagtctc	atatggggtc	cctaagggtc	360
gcctgttata	cagcggcagc	tctgagtttt	atgtccggcg	tcattcagac	gttcacgcgc	420
gatcggcccg	acgtgtcggg	ctacctcgac	acagttcctt	cccagacggg	cctogaattg	480
gtctcgctcc	agcactacga	tctcggaaata	tcgatattgg	ctggcgacta	tcctggtctc	540
accaccgaac	ctgtcccttc	ctttcgtgcg	gtctgctcgc	tgccggccggg	gcatcgtctc	600
gaagacaagg	aaactgttca	tgcgacggac	cttgaaggag	agtcattgat	ttgctctctc	660
ccagtgagcc	ttctacggat	gcaaacggac	gcgcactg	acagctcggg	cgccactgt	720
aatcgcagga	tagaaaagtag	tctggcgctg	aatctctcgc	atctggtaag	caggggaaatg	780
ggggttggtg	tcgtcgacc	cttccactg	gactactaca	gtgcaaatcc	ggttattcag	840
cgctcctttg	atccggttgt	cccctacat	tttgcctatag	ttctccgac	cgacagccca	900
ccgcccgcgt	tggttagcga	gttccggcca	gcgttgcttg	atgctttgaa	agccttgccc	960

-continued

---

 tatgaaacca tttgatcgtc aggatcg 987

<210> SEQ ID NO 4  
 <211> LENGTH: 1246  
 <212> TYPE: DNA  
 <213> ORGANISM: Escherichia coli

<400> SEQUENCE: 4

gatccgctaa tcttatggat aaaaatgcta tggcatagca aagtgtgacg cctgtcaaat 60  
 aatcaatgtg gacttttctg cctgtgattat agacactttt gttacgcggt tttgtcatgg 120  
 ctttggctcc gctttgttac agaatgcttt taataagcgg ggttaccggt ttggttagcg 180  
 agaagagcca gtaaaagacg cagtgcggc aatgtctgat gcaatatgga caattggttt 240  
 cttctctgaa tggcgggagt atgaaaagta tggctgaagc gcaaaatgat cccctgctgc 300  
 cgggatactc gtttaatgcc catctgggtg cgggtttaa gccgattgag gccaacgggt 360  
 atctcgattt ttttatcgac cgaccgctgg gaatgaaagg ttatattctc aatctacca 420  
 ttcgcggtca gggggtggg aaaaatcagc gacgagaatt tgtttgccga cggggtgata 480  
 ttttctgtt cccgccagga gagatcactc actacggctc tcatccggag gctcgcgaat 540  
 ggtatcacca gtgggtttac ttcgtccgc gcgcctactg gcatgaatgg cttaactggc 600  
 cgtcaatatt tgccaatagc gggttctttc gcccgatga agcgcaccag ccgcatttca 660  
 gcgacctgtt tgggcaaact attaacgccg ggcaagggga agggcgctat tcggagctgc 720  
 tggcgataaa tctgcttgag caattgttac tgcggcgcac ggaagcgatt aacgagtcgc 780  
 tccatccacc gatggataat cgggtacgcg aggcttgta gtacatcagc gatcacctgg 840  
 cagacagcaa ttttgatata gccagcgtcg cacagcatgt ttgcttctcg ccgtcgcgtc 900  
 tgtcacatct tttccgccag cagttagga ttagcgtctt aagctggcgc gaggaccaac 960  
 gtatcagcca ggcgaagctg cttttgagca ccaccggat gcctatcgcc accgtcggtc 1020  
 gcaatgttgg ttttgacgat caactctatt tctcgcgggt atttaaaaaa tgcaccgggg 1080  
 ccagcccgag cgagttccgt gccggttgtg aagaaaaagt gaatgatgta gccgtcaagt 1140  
 tgtcataatt ggtaacgaat cagacaattg acggcttgac ggagtagcat agggtttgca 1200  
 gaatccctgc ttcgtccatt tgacaggcac attatgcaag cattgc 1246

<210> SEQ ID NO 5  
 <211> LENGTH: 1335  
 <212> TYPE: DNA  
 <213> ORGANISM: Escherichia coli

<400> SEQUENCE: 5

gcaatgcttg cataatgtgc ctgtcaaatg gacgaagcag ggattctgca aacctatgc 60  
 tactccgtca agccgtcaat tgtctgattc gttaccaatt atgacaactt gacggctaca 120  
 tcattcaact tttcttcaca accggcagc aactcgtctg ggctggcccc ggtgcatttt 180  
 ttaaataccc gcgagaaata gatttgatcg tcaaaaccaa cattgacgacc gacggtggcg 240  
 ataggcatcc ggggtgtgct caaaagcagc ttcgcctggc tgatcgttg gtcctcgcgc 300  
 cagcttaaga cgctaatccc taactgctgg cggaaaagat gtgacagacg cgacggcgac 360  
 aagcaaacat gctgtgcgac gctggcgata tcaaaattgc tgtctgccag gtgatcgtg 420  
 atgtactgac aagcctcgcg taccgatta tccatcggtg gatggagcga ctcgtaatc 480

## -continued

---

gcttccatgc gccgcagtaa caattgctca agcagattta tcgccagcag ctccgaatag	540
cgcccttccc cttgcccggc gttaatgatt tgcccaaaca ggctcgctgaa atgcggctgg	600
tgcgcttcat ccggggcgaaa gaaccccgta ttggcaaata ttgacggcca gttaagccat	660
tcatgccagt aggcgcgccc acgaaagtaa acccactggt gataccattc gcgagcctcc	720
ggatgacgac cgtagtgatg aatctctcct ggcggaaca gcaaaatc acccggtcgg	780
caacaaatt ctctccctg attttccacc acccctgac cggaatggt gagattgaga	840
atataacctt tcattcccag cggtcggtcg ataaaaaat cgagataacc gttggcctca	900
atcgcgctta aaccgcccac cagatgggca ttaaaccgag atcccggcag caggggatca	960
ttttgcgctt cagccatact tttcatactc ccgccattca gagaagaac caattgtcca	1020
tattgcatca gacattgccg tcaactgcgc ttttactggc tcttctcgct aaccaaaccg	1080
gtaacccccg ttattaaaaa cattctgtaa caaagcggga ccaagccat gacaaaaacg	1140
cgtaacaaaa gtgtctataa tcacggcaga aaagtccaca ttgattatct gcacggcgtc	1200
acactttgct atgccatagc atttttatcc ataagattag cggatcctac ctgacgcttt	1260
ttatcgcaac tctctactgt ttctccatac ccggtttttt ggatggagtg aaacgatggc	1320
gattgcaatt ggaat	1335

&lt;210&gt; SEQ ID NO 6

&lt;211&gt; LENGTH: 1113

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Escherichia coli

&lt;400&gt; SEQUENCE: 6

ccggaagaga gtcaattcag ggtggtgaat gtgaaaccag taacgttata cgatgtcgca	60
gagtatgccg gtgtctctta tcagaccgtt tcccgcgtgg tgaaccaggc cagccacgtt	120
tctgcgaaaa cgggggaaaa agtggaaagc gcgatggcgg agctgaatta cattcccaac	180
cgcggtggc acaaactggc gggcaaacag tcggtgctga ttggcgtgac cacctccagt	240
ctggccctgc acgcgcctgc gcaaattgct gcggcgatta aatctcgcgc cgatcaactg	300
ggtgccagcg tgggtggtgc gatggtagaa cgaagcggcg tcgaagcctg taaagcggcg	360
gtgcacaatc ttctcgcgca acgcgtcagt gggctgatca ttaactatcc gctggatgac	420
caggatgcca ttgctgtgga agctgcctgc actaatgttc cggcgttatt tcttgatgac	480
tctgaccaga caccatcaa cagtattatt ttctcccatg aagaaggtag gcgactgggc	540
gtggagcatc tggctgcatt gggcaccag caaatcgcgc tggtagcggg cccattaagt	600
tctgtctcgg cgcgtctcgg tctggctggc tggcataaat atctcactcg caatcaaatt	660
cagccgatag cggaaacggga aggcgactgg agtgccatgt ccggttttca acaaacctg	720
caaatgctga atgagggcat cgttcccact gcgatgctgg ttgccaacga tcagatggcg	780
ctgggcgcaa tgcgcgccat taccgagtcc gggctgcgcg ttggtgcgga tatctcggta	840
gtgggatacg acgataccga agacagctca tgttatatcc cgcgctcaac caccatcaaa	900
caggattttc gcctgctggg gcaaacaccg gtggaccgct tgctgcaact ctctcagggc	960
caggcgggta agggcaatca gctgttgccc gtctcactgg tgaagagaaa aaccaccctg	1020
gcgccaata cgaaaccgc ctctcccgc gcggtggcgg attcattaat gcagctggca	1080
gcagaggttt cccgactgga aagcgggcag tga	1113

-continued

---

<210> SEQ ID NO 7

<211> LENGTH: 24708

<212> TYPE: DNA

<213> ORGANISM: *Agrobacterium tumefaciens*

<400> SEQUENCE: 7

ctgcaggagc agaggtgagc gcgaccgtcc gatccagccc gccggtgctc tcgaagcaca 60  
acacaccgcc aaccacgagc accgcaacct acatcgaccg atgcatcacg tgetgaccgt 120  
tctagcattc gccgcagtgt cgcgttcgtc cgcttctgtc cgacgcttgc ccttcgttga 180  
acgaattcgg ttcaggatgg gcgacagaaa cgcaatgtct tcccatatct cggcaccttc 240  
gaagagcgtc tcgtgtttgc gattgcccga ccgcacctct tctctcgtca cgataggcgg 300  
cgcgatctcg gcaagggtctg cgagcggcgc aaggatcagg acacagaaca tgccgattgc 360  
tttcagggcc gtttcccccga gcagcgggag cacgccaaga gaggcgatag ctgccacacc 420  
gaaaaacggc gtgacgataa tgaacgcgac ccagatcatc agcgggcccga tcagcgagta 480  
ccaggttggt ctccaatgcc tcatgggggc gtcccttcct tctgccatgg cggcgcattg 540  
tcagtgtgct taccggttcc ggttccagtt cgatggaacc tgttctcgtc gacacaggct 600  
ttcatgtcat cacggcgcaa ccagatggcc cggccgcacc ggagcggcgg attgcccag 660  
gtgaagacga tctgctcgtc gctgcccattg tgcagaactt catgcggcag gatcagcggc 720  
cggcactga gctgcttcca tctcgatcgc gaagatccct tcattccggg tgaacgggtg 780  
gtctggtcga cctcgacggg ggtgtctcca cagcgttcg agatatagtc ggcggtatcg 840  
ggatcgtaa ttgcccgaaa cgaatccac gatgcggatt cgaaccattt gctggctcga 900  
tcacgcccgc cataggcctc acgatctgg ccgagcgcact ggaagatcat cgtcagcatg 960  
atgccgtatt tgcggcctgc gtcgagcgc gtttcgagga tgcgcaggta gccgagacgc 1020  
gctacctcgt cgagcaggaa gagggtgcga ccctttacat tgcggttgcg gttatagatc 1080  
gcgttgagca gcgagccgat gacgacacgc gcgagccgg gatgggttc cagcaccttc 1140  
aggtcgagcg cgatgaagat gtcggttccg ccgctggcga gatcgtcggg cgaagagctg 1200  
tcgcccgata caagtccggc atagttggga taggacagcc aatgggtttc cttcaccgca 1260  
ttggcataaa caccggaaaa cgtctccggc gtcattgtga cgaacactga cacattctcc 1320  
ttcacgaagt cggattcaga cccttcgtag atcttcgtca gacgcgcgcg cagcttgggt 1380  
tccggttcgg acaggttggc gcggacgcgg cgcagcgtct ggtcttccgt ttcogtgtgc 1440  
cctgacaggc acacgtcggc gatcaggggc gtcagcagct gcatggcggg agcgcggaag 1500  
aagtctgctc gggcggatgc cgttcgggga ttgtcgggta tgatccatgt ggcgacggcg 1560  
acgatatcct cttccttggg attgccatga cggccgatcc agtcgagggc attgaagccg 1620  
acaccgccgg ctgtcggatc gacacagatc accttgcggc cggcctgacg cctgtgctcg 1680  
cagaccatcg gcgcgacctc gcttgacggg tccagaacga caagcccgc gccccatttg 1740  
agcgcggctc ggatcgtcac cgatgtcgtc ttgaaaccgc cggagccggc gaagacgatg 1800  
ccatgcgacg atccgaagga accgtcgaag cagagcagcg gcgacttgcc gccagcggcc 1860  
cagctttgtc gatcgtcagc gcgaaacggc atggcggcaa cactgtcgcg atcaaccgca 1920  
tagcgtctgc cgatgatgat gccgcctcgt tcgggaaaca gttttgccgc ctcctggatc 1980  
ttcatccagt cggcctcggc atgcacagcc cgcttgcggc cgatccgttt gggcgtggtt 2040

-continued

---

gttgcgaacg	cggcattgcc	cttgatcgcg	acacgcagcg	cgaacattcc	gctgatgaag	2100
gcaatgcttg	cgccgatcat	cgtcaccgga	tcagcatagc	ccagaaccga	ttgttcggcc	2160
ggcacgctgc	ttgcgatgct	attcaggcgg	attacttcgc	gcagtgttgc	gatggtaatg	2220
accacgccc	ttccggccag	cacgctcaga	ccggcggcct	tgatgttggc	cgatccgttc	2280
gcggcgaaga	ggccaattat	gccgatcgcc	accgcagcga	tgtaggggag	cgcgagccc	2340
gcgctccca	gcatcagctt	cgctcggg	gacgttccga	atgccgccag	tcgctgctcc	2400
atcccggatg	tcgcgaacat	cgccgcgagc	attatgattg	ccggcaggat	caacagcaaa	2460
agcctattca	tcgtcatggc	cgaacgcctc	cgcccgatg	gcggtcaagc	gtgaccgctc	2520
ttcttcgctg	cccttgagcc	ggcgccgccc	gtcgaccaac	agcccagca	acagcgccc	2580
cttttcgtaa	cgcagtcctg	ccttgacgat	cagcccggcg	agctcgattt	tttcgagggt	2640
gtccttcttg	cgctcgcttg	acgtcatcga	ctttgccatc	cgttcaagcc	tcgccagacc	2700
cgccctcatc	cgccccagac	gcgagcgcg	cggacggctc	ggtaccggtc	cggtctctc	2760
cggttgtct	ctttccggtc	gccgagccct	tgccgccacg	aaaccgtttg	gcgacttctc	2820
cgaaggctgc	ctgaagctgt	gactcctcga	tgctgatttc	accaagaccg	gccttcaacg	2880
caatcctgcc	gattcgcttg	gcctcgcgtg	tttcggcctg	tttcagctgg	tcctgcaatc	2940
tggcaatttc	ttccctgatc	ttcgatgatg	gtttcttcat	tccggtcgca	tctccctgga	3000
aatcctgcgg	cgctgtgtcc	gctgcaagat	ttcctcaaaa	gcacttcgga	aggaatgtgc	3060
agatctgcac	gtcggcaaa	ccgacacttt	ggaggatgat	cccgcgctc	gacgagagcg	3120
gatccaagg	cgcaattata	cgctcgtgac	gcgacgcctt	gcgtaggggg	ccaaacagg	3180
gcccactgtg	gcctcaccgc	tcccagcga	cgacgttcaa	acgggagctt	ttaccgccgt	3240
ggccatcgcc	cacttctcag	ccagcatcgt	cagccgcggc	gacggccgca	gcgtggtgct	3300
gtctgcggcc	taccagcact	gcgcgaagat	ggaatacgag	cgcgaggccc	gcaccatcga	3360
ctacacccgc	aagcaaggcg	tggtgcatca	ggaattcata	ctcccggcgg	atgccccgaa	3420
atgggtccgc	gcgttgattg	cagactgttc	tgctgccggg	gcctctgagg	ccttctggaa	3480
caaggtcgag	gctttcgaaa	aacgtccga	cgccagctg	gcccgcgacc	tgacgattgc	3540
actgcccgg	gagcttacct	ccgagcagaa	catcgccttg	gtccgtgatt	tcgtggagaa	3600
acatatcctg	gggaagggca	tggtcgcga	ctgggtctat	cacgacaatc	ccggcaatcc	3660
gcacatccac	ctgatgacca	cattgcggcc	gctaaccgag	gacgggttcg	ggccaagaa	3720
ggtagcggtc	attggtgagg	acggccagct	ggttcgcacg	aaatccgca	aaatcctcta	3780
cgaactttgg	gccggttcga	cggacgattt	caacgtcgtg	cgcgatggat	ggttcgaacg	3840
gctgaaccat	cacctgacgc	ttggcggcat	cgatctcaag	atcgatggcc	gctcctacga	3900
gaagcaggga	atcgacctcg	aaccgacgat	ccatctcgga	gttgccgcca	aggcaatctc	3960
gcggaagcg	gaacaacag	gtgttcggcc	ggagctcga	aggatcgaac	tgaacgaaga	4020
acgccgctcg	gaaaacaccc	gccgcatcct	caaaaatcca	gccatcgtgc	tcgacctgat	4080
catgcccggaa	aagagctctc	tcgatgaacg	cgacgtcgcg	aaagtgtctg	accgctatgt	4140
cgacgatccc	gccgtcttcc	agcagctgat	gttgcgcatc	atcctgaatc	cggagggtgct	4200
acggctgcag	cgcgacacga	tcgagttcgc	gaccggagag	aaagtccgg	cgcgctattc	4260
gacacggcgg	atgatccggc	tggaagcgac	gatggcgcga	caggcgatgt	ggctgtcggaa	4320

-continued

---

caaggaaacg	catgccgttt	ccacggcggg	gctgctcgca	gcgacttttg	ggcggcatgg	4380
gcgggtgctg	gaggaacaga	aagcagccat	cgagtgcac	gccggtcg	cacgcacgc	4440
cgcggtcgtc	ggacgcgcg	gtgctggcaa	gaccacgatg	atgaaggcgg	cgcgggaggc	4500
gtgggagctg	gccggatata	gtgctgctgg	tggcgcgctc	gccggcaagg	cgctggagg	4560
tctggataag	gaagccggca	tcgagagccg	cacgctctcc	tcctgggagc	tgcgctgaa	4620
ccgcggtcgt	gacgtgctg	acaataaaac	cgtcttcgct	atggacgagg	ccggcatggt	4680
ggcctcgaaa	cagatggcgg	gctttgtcga	tgcgctcgtc	agggcagggg	caaagatcgt	4740
gctagtccgc	gatcccagc	agcttcagcc	gatcgaagcg	ggggccgctt	tccgcgccat	4800
tgctgatcgg	attggttacg	ctgaactcga	aacgatctat	cgccagcgcg	aggactggat	4860
gcgaaaggcg	tcgctcgatc	tggcgcgcg	caatgtcgag	aaggcgttgg	cgctctacaa	4920
tgccaatgcg	agaattgttg	gcgaactctc	gaaggcggaa	gcggttgagc	gcctgatcgc	4980
cgactggaac	cggtattacg	atcagacgaa	gacgacattg	atcctcgccc	atctgcgcgg	5040
cgacgtgcgc	atgctcaacg	tcattggcgg	cgagaagctg	gtcagcgcgg	gcacgtcgg	5100
cgagggccac	gtcttcagaa	cggtgacgg	cgaacgccgt	ttccacgcag	gcgatcagat	5160
cgctctcctg	aagaatgaga	ccttacttgg	ggtcaagaac	ggcatgatcg	gccatgtcgt	5220
tgaggctgtg	ccaaaccgca	tcgttgccgt	ggtcggggat	cgggatcacc	ggcggcatgt	5280
tggtgctgaa	caacgcttct	acagcaatct	cgatcacggc	tatgcgacga	ccatccacaa	5340
gtctcagggc	gccactgctg	acagggtcaa	ggtgctggcc	tctctctcgc	tcgatcggca	5400
tctgacctat	gtcgcgatga	cccgccatcg	cgaggatctc	cagctctatt	acggatgccg	5460
ttcctctcgc	ttcaacgggt	gcctagccaa	ggttctttcc	cgaagaagc	ccaaggagac	5520
gacgctcgat	tacgagcgcg	gcaaactcta	ccgcgaggcg	ctgcgctttg	ccgaaaaccg	5580
tggcctgcac	atcatgcagg	tggctcgaac	aatgctgcgc	gaccggctcg	actggacgct	5640
gcgcacaaaa	acaaggttt	ccgatctcgt	tcataggctt	cgctgcgctc	gtgagcgctt	5700
cggcctcgac	caatccccaa	aaacccaaac	gatgaaggag	gccgcgccca	tggtcacagg	5760
catcaagaca	ttctccggct	ctgttgccga	tacggtcggt	gacaagctcg	gcgccgatcc	5820
cacgctgaag	cagcaatggg	aggaggtctc	ggcgcgcttc	cgctacgtct	tcgccgatcc	5880
cgaaaccgcc	ttccgcgcgg	tcaacttcga	tacggttctg	gcagataagg	aggcagcgaa	5940
agcggctctg	cagaagcttg	aggcggagcc	ggcatcgatc	ggggcgctga	agggcaagac	6000
cggtattctg	gccagcaaga	cgagcgcgca	agcgcgccgc	gtcgcagagg	tcaatgttcc	6060
cgccctgaag	cgcatctgg	agcaatatct	gcggatgcgc	gagaccgcga	cgctgcggat	6120
agtgaaccgaa	gaacaagccc	tgcgccagcg	cgtctcgatc	gacatcccgg	cgctctcgcc	6180
ggccgcacgc	gtggtgcttg	agcgcgtgcg	cgacgcgatc	gaccggaacg	atctgccggc	6240
ggcccttga	tatgcgctca	gcaatcgca	gaccaagctc	gagatcgacg	ggttcaacca	6300
ggctgttacc	gaacgcttcg	gggagcggac	cctgctgagc	aacgtggcgc	gcgagccgctc	6360
cggaaagctg	tacgaaaaac	tctccgatgg	tatgcggccg	gagcagaaag	agcagttgaa	6420
gcaagcctgg	ccgatcatgc	gcacggccca	gcagctcgcc	gcccacgagc	gcaccgtgca	6480
gtcgcgaaa	cagcccgagg	aattgcggca	gacgttgccg	caggcaccgg	tgctgaaaca	6540
atgagcaaac	gcgctgctat	ccggtttctc	ggtgctcgag	gcctcgtctt	gagcggcgcc	6600

-continued

---

accgtcatgg	ggctcatcgg	tggttatcga	gtgaacgtca	cgccgagcga	accccttggg	6660
ctctggcgtg	tggaaaagct	ccgacgccc	gttcagtcg	gcgatctggt	gtttgtctgc	6720
ccgcccaca	acgccttctt	cgaggaagcc	tggctacg	gctatctccg	ttccggctc	6780
tgccagggtt	ggttcgcgcc	gctcataaag	tccgttctg	cgcttcccgg	tcagcgcgtc	6840
gaaatgccg	atcgattgt	catcgacgt	catcccgtat	cggttccac	tgccagcgcg	6900
acggacagcg	agggcagggc	gatcgcgcc	ttcgcggcg	gcgtcgtgcc	gccaggattt	6960
ctatttctcc	attcttcctt	tgcgagtcc	tatgattcac	gatacttcgg	cccgatccc	7020
gacagcggc	ttctcggcct	ggcgcggcc	gtgttcacgt	ttgaccctg	accgctggcg	7080
tgtggcgctc	ctgatctgc	ttcgcagtc	cgccggagc	gtcggctgga	gcggcgaggc	7140
cctccttctt	ccgacagcca	tgctctttcc	gttgcctatg	gcgcaatctc	catcacgact	7200
ggttcggggc	gcggtgtctg	ccgggtactt	cctcaccgcc	tcgcgcggcc	tgccgcaggg	7260
tgtggcgaac	ttctacgccc	ccgatttctg	gcacggcctc	ctgctatggc	tgccggcctc	7320
ggccggcttt	gtcgcctg	atgcggcgtt	ctggcccgcg	cgcttgcaaa	agaggcttcc	7380
gggaagggg	gctctggggt	gggaaaagcc	ggtgcgctac	cttgcggccg	cggtgctgat	7440
ggggctgcc	cccttcgcga	tcaccgggtg	ggcgcaccca	ctaaccgag	cgggcattct	7500
gtttcccggc	tttggatggt	gggggctggg	tgccaccaca	gccggcctcg	cgatgatgac	7560
atcgcgatac	tggccgctg	ctgccattgc	cctgggaggc	ttctgggtct	ggtccgccgc	7620
tacgtggagc	caacctgtcc	taccggatg	atgaaaagcc	gtcgcacctg	aacagggtca	7680
aacgcttga	cgatgatggt	cgcttgacca	tcaccgtgac	ctgatcgcaa	cggtgcgcgc	7740
ggcgcagga	gcgaaaacc	gtgtcatcgt	tcttcccga	agcgcctc	gcttatggac	7800
cccactgtc	gcacggcttt	ggcaggcggg	cttgcgagg	gcggacgtca	ccgtgatcgc	7860
cgccgcggcc	gtcatcgatc	cgggcggcta	cgacaacgtc	atggtgacgg	tctcggagg	7920
tgaaacacgg	atcctctatc	gcgagcggat	gccgatccc	gtatcgatgt	ggcagccatg	7980
gttgcaatgg	accgggcag	gcggtggcgc	tcaagcgc	ttcttcgcca	accggccgt	8040
ggatctcgt	ggcagcagga	tcgcgcgct	gatctgtac	gagcagctca	tcgtctggcc	8100
catactgcac	tcgatgctgt	ttccccccgc	cgccatcgtc	gccacgggca	acggatggtg	8160
gaccgagga	acgtcgatcg	tcgccatcca	acaagcagc	gtgatcgct	gggcgaaact	8220
cttcgggcga	cccgtcgtca	cggtttcaa	caataagga	gactttcgcg	atggtcgatt	8280
ccgcctcat	caaggaatgc	agcgatccc	gcctgaaacc	cgcatcgtc	gaacagttca	8340
tcgaaacggc	aggggcgcg	gaccgcctc	cggtcactgt	ccgctccggc	aaccgcatcg	8400
tgcttggtcc	gagccgcaca	acgcccggac	agggcctcgc	attgatacgc	cagaacctcg	8460
gtcgcaatat	gttcgggctc	ggcttgacgc	agtatcccgc	cggtctcgg	atcgtcgagg	8520
cgggccagtt	gaagccggac	ctgatcgagc	cttgcgaaaa	catcaaatg	ggcagggccc	8580
tgcttcgcaa	agtctatcgt	atcgtcacca	aatggtacgc	taacctacc	gaaagcgagg	8640
tccttccgca	ggtttttgaa	gacgcgatca	tcgcatggca	aacggggtat	ttcgagggga	8700
cagcgggtgt	ccgggcagcc	gatccggg	acgtgaaact	gccagagccg	gatoctgggg	8760
tggaagcgg	aacgtccgaa	gccggtagtg	acaagatgaa	ggcgggcaac	gtcgatactg	8820
cactcgaggc	ggaaccagct	gaccgcaaca	agggcgggtat	tcggatcgac	ctttccggca	8880

-continued

---

tcggcgcgcg	cggaagtga	ggaacaggtt	cctcgcgtct	gcgtcgatga	acgcgggcca	8940
ggagcgaaaa	attgctcggg	ccgtgtgcca	agcgaatcgc	agttaatgag	accgtggcgg	9000
aaaacaacta	agtgctcgtc	cgcttcaaca	atgcattcta	accctggctg	caagtgcgct	9060
tgatttcttc	ttcgaagca	taggccttga	ggtattttgt	gccttgcatc	gagggtggca	9120
tgtgaagatt	tgcaccctcg	ttctgtcgat	cgcgctggc	atctctttgg	aatggctctt	9180
ataaaggcca	tgacgcagta	cgcgctgtgt	tcggccttag	aatccgcaga	tgaacggttg	9240
gcgctgaccg	aaacggtgcc	aggaaccctc	tgacctgtg	gccgggggcc	tgtgactcat	9300
gtccaggcgc	ctcgggtcta	aaggctctcag	ggccgattca	tctcggattc	taacccccgg	9360
ctagggttca	aagggttcaa	cgtcagcggg	ggcgaccgc	agacaaaacc	agggttaggt	9420
caggaatgga	atcggaaagt	gcaacattga	cgaaaaaagt	tgagctccgg	ccactggctg	9480
gcttaactcg	tggaactacac	ccggctgacc	tcgagaaact	gaccatagac	gcaatccgtg	9540
cgcatcgacg	actggctcga	aaggccgacg	aacttttcca	ggccctgccg	gaaagctata	9600
agtcgggaaa	ggaagtcgga	ggcccacagc	atctgtgtta	catcgaggcc	agtatcgaga	9660
tgcacgcgca	gatgagcgct	gtcagcacgc	tgatcagcat	tcttggctat	atcccgaacg	9720
ccaccgtgaa	ctgaaccggt	ttccgtttgt	ctcatggcta	atctttatag	cgggctatga	9780
tgtcataccc	tgaactattg	cggcagctca	gaccgctat	cttcatcgga	gaaaggcgg	9840
cctgacacga	gctcctccgt	tcggtctctt	tgataacgac	tcgggtgga	ccctacgcag	9900
cgaaagacctg	ggcagcgcg	gtcgcactgt	tggcgactgg	cgcgacgat	ggctgatgaa	9960
cctggccagc	aaacgcttgc	gttcggccgc	tatcatcgaa	cctgaaatgc	tccagcagct	10020
cgaaatagcgc	ggcggcctcg	cgagcaaggc	tgtggctggc	cgccgtctgc	tcctcgacca	10080
tcgcgcgctt	ttgttgcggt	ccctgatcga	cgctgttcac	ggccttgttg	atctcaccga	10140
gagcgttcga	ctgctcgcgt	gccgcttcga	caatcgcgac	gacattcgtg	ttgatgtcgt	10200
ggacctgggt	tgcgatttcc	tcgagcgcgg	aaccggcttt	cgtgacaagc	gataccccgc	10260
tctggacctg	gctgcccagc	gaggatgatca	gcgtcttgat	ttccttggcg	gccttggccg	10320
atcgctgcgc	gagctcgcgg	acttctctcg	cgacgaccgc	gaatcccttg	cccgttctgc	10380
cggcacgggc	cgcttcgaca	ccggcattca	gcgccaggtg	gttggctctg	aacgcgatct	10440
catcgatgac	accgatgatg	ttggagatct	cgcgagagga	cgtttcgac	tgatccatgg	10500
cgccgatcgc	gtcgcgcacc	acctccccgg	aatgctctgc	atgttcgcgg	gtcogtgcaa	10560
cgagttggcc	agcctcctcc	gcccagcagc	tggaatcctt	gaccgttctg	gtgatttcct	10620
cgagcgcagc	ggcggctctc	tcgaccgagg	cggttgttg	ctcggtgctg	tttgccagct	10680
cgctggcggc	ggtccgaatc	tcgctggatc	ccgaggagat	cgccctggcg	ttgtcggcga	10740
cagtcgccat	cgcgcgcttc	agcttctcgg	aggcggcgtt	gaaatcggtg	cgaaaacgct	10800
cgagcgcagg	gatgaaccgt	tttctgattt	gctgcgccag	gtcaccgtcc	gccagactgt	10860
tcaggcagcc	cgcgagctgt	tcgacgtttt	cgaccgggt	ggtcacgtcg	gtcgcgaatt	10920
tcaccacctt	gaaaactctg	ccgttcatat	cgaagatcgg	attataggat	gcctggataa	10980
agaccttgcg	accacccttg	ccgagccgca	tgaactcgtc	cgccatcaaa	tccccgctg	11040
caagcatctt	ccagaaattc	cgatagtcgg	gcgacgcgg	ataggatggt	tcgcagaaca	11100
tcgaatggtg	tttgcctcgg	atctcgtcga	gagaatatcc	gagggcggac	aggaaattct	11160

-continued

---

catttgcggt	caggacgtcg	ccggtcggcg	tgaactcgat	aatcgctgc	gcccgcgaca	11220
aagcctcgat	cttgccgca	tcttctgcaa	ctttcagctt	ctggcggtg	atgtcggtg	11280
cgatcttgac	aaccttgacc	ggcttgccac	ggcgcgatgac	agggttatag	gacgcctcga	11340
tccagacttc	cttgccgccc	ttgccgatgc	gcttatattg	ctgctggtcg	aagttgcccg	11400
cgcgagttt	cgcccagaac	gccttatagt	ctgccgattt	cacaaatgcc	ggttcaacga	11460
acatgctgtg	atgtttgccg	ataatttctg	atagctcata	gccaagcgc	cggcagaaat	11520
tgtcgttggc	agtgagtatc	ctgccggaca	gatcgaattc	gatcatggcc	tgagacttag	11580
agagggcagc	aagtaccgcg	catgcattgg	caccacgatt	gagaatattc	accaggcgc	11640
ctcccgaaga	tacttaccgg	accgaaggtc	gcggcagagc	gcgaatagtt	gcatacgatg	11700
tttaatcttc	atttaaacat	ggaactggca	tccatcaatt	atgcccgctt	aaagcgctcg	11760
aagagaagga	ggagattgta	tcaacaatca	tgtgcccgtc	gataagggct	tccatgttgc	11820
gattgagcaa	agacagcagc	agctcttttg	tgacggcccg	ggtctcgatt	ttttcgagat	11880
agagcggacg	ccacgtgagg	tcgagatcct	ggcttaggtc	ttcgtgcaac	tgactaact	11940
cctgcccggc	cgcttccagg	ctatcgaaaa	actgaacggt	ggtgtgcccg	tgacgccagt	12000
gcgcctgac	atcggcagtc	gcctgagaat	aggtgtagaa	cgtatagaa	gttactgaca	12060
tcattcctct	cccctggttt	gtctgtcacg	ttttcagatc	agctttcttc	tgattgcgag	12120
ggcggtgaga	tgggccttgc	tgcaacgctc	gaagcgcttc	atggcctcgc	ggagcttgac	12180
acggacgctg	ttgtacttga	cgccctccac	gtctgcgact	tcctccattg	tcagcccgac	12240
ggcgatccat	ctgagatagg	tcgcctcttt	cggatcgagc	caggcggcat	cttccactgt	12300
cggagtggtc	tgaaggaag	agatgcgggc	atggagctgc	ccgacggcgc	ccgcggtgc	12360
ggccgcgctg	atctcacggt	cgaggtcgat	cgccggcctt	tccgacgcca	gcgtgaacat	12420
cgacattgat	ccgttggcgg	tcttgatcgg	aatggtgatg	ccggagcggg	tgccgaaatc	12480
ggccgcatgc	gcgtagaagg	cacgctcttc	cttcgatagc	cgcgatcgtt	cctgttcgcc	12540
ggaccaggca	aagacgtgct	tcctggattt	cgcgcgcttg	acgaccgat	cgagcttgtc	12600
gaagttgttc	tcgaagtaag	ccgatcgcca	gtcacgatga	taattggtga	ccgogatggt	12660
gtgtttgtgc	tggtatagga	gataggcata	gccggtgaag	ccgaaatggt	cggcaaggtc	12720
ggcaaggcca	tccttcagga	tgcaactgct	gcctgaatt	cgcgcaagat	cggtaactt	12780
gtccagccag	tgctgcattc	cataccccc	tatgctggtg	tgcgatgcct	actctaactg	12840
ttgtcccag	cagccgaccg	ccgggctgaa	tctgctttca	tcctctttta	cccagacaac	12900
tcaggctgag	ccgaagcatt	tttccacgag	gcatacatgc	tccaacgctt	cacgaatata	12960
cgtactatcc	ccctgtacat	gctatcgagc	cggggaagcg	gcacttgcgt	gcagttcact	13020
cgacgatctg	catctcgctc	tgaccagatg	cggcaatgct	cgcttgcgg	tcggcaaatg	13080
agatcttcaa	ttcagcaaaa	gcacgcgctg	gttgccatga	aagcaccac	gctccggcat	13140
cagcaatggt	ggccttgagc	gctccaccga	aggcgggatg	ggtgttcctg	aaacggatca	13200
gatttgaaa	gcgtaggaag	agcggcgtct	caaggccgt	accgatttcg	tcgttgccgt	13260
aatagtccg	gttaatgtca	cgccccacgc	ccgttttggc	gagcaggtcc	atatcgttta	13320
cgccgcccaa	taaaccgaca	taatagactt	gcggaattcc	tgagcaaaa	aactgaattg	13380
cccagcgat	gagataatca	ttatcgttgc	gaccaagcgc	gtcatagtag	gtgcagtta	13440

-continued

---

cctgataaag	gtccagattc	gatgcagcgg	ccccgcgcg	ctggcggctc	tggccatccg	13500
atcgccatg	gatctctccc	acaagattgt	cgatcgcctg	gggttcgagg	aggccccccc	13560
ggccgcgct	atgggcaccg	acatctataa	cgccaatccc	gtcatgggtg	tcgagaacgg	13620
tgatcgcgtt	acgggggctt	attgccagcc	actttgccag	cggcgtggcg	tctcccgtaa	13680
acaatgcatg	caagatgagc	ggtggtagt	cgaaatcgta	caccgcgtcg	acctttctcg	13740
cgatctcgat	ttggtcacga	tgataactat	gaatctcgac	cagaacctcc	atgccgcggt	13800
tgccggcctt	tccagccact	tcttcgagga	atgcgtaggt	atcgtcgatc	atgaagcaac	13860
tgtcggccg	tttcttgatc	gcgatcccc	cggcgtcgag	gcggatcgcc	gtcacattcg	13920
cctccgaaaa	gcgatcgagg	atcgtttcga	gataggttgt	gcctttggcg	ctgtggacat	13980
cgatgtcgat	ctgctccggt	gtgaaggttg	tccagagcat	ccgctgagta	ccatcagcca	14040
gggtcacctt	gaaaacggg	aaaccggggc	gcggtcgata	gatcctcaga	aggtcttctt	14100
cggtgccggg	atcgggaaat	accttatcga	acgtcatgaa	catgtcggca	aattcggacg	14160
ccgaaccctt	ggcaatgaag	tccgtgaatg	cgccggattg	cgatgagaca	tggttgacga	14220
taagatcggc	catgatttcc	acgatcccc	acaaagcccg	cacatcctcc	cagtcgccga	14280
ggcgcggatc	gacgatcgtg	tgatctgttg	gatcgaatcc	ggcatcagcg	ccatcgatcg	14340
gattgaagaa	aggcaagaca	tggacgccgc	cgaacaggcc	agcgaacttg	ccatcgacca	14400
gttccttcaa	ttgcctgaaa	ccacctctgg	tcaggcggtc	aacatagggtg	ataaactgca	14460
ctttgttttg	catttcaatt	tccagtatgt	gtcgtcttac	gaggacgagt	tgattttctt	14520
tcactcagtc	tgacagaccg	cccgcgatga	gacggctctc	gccagtaat	ccttgctcgt	14580
ccatgatcgg	gtgacgaacc	gagacgatat	gcgcgttgat	ccgcttgaga	tcacgcaaga	14640
tgtcgagggtg	aagcagcgtg	gtttgaaggc	tgctcggtcg	cccgtcgcgc	agacgttcca	14700
ggtggcgcctc	ggaagactgt	ttttccatgt	gccggacttc	gactttccgc	tccatgagct	14760
gacgcgcgag	attgatgtcc	cgggtgataa	ggattgtttg	tgcgatccgg	agattatcga	14820
cggtgagcgtg	gaagagtgtg	gttagttcgg	agaagccgtc	atcagagaat	ttcaaaccgt	14880
ggcgatctt	cttcgtgacc	tcgggcagca	gacctttctc	gacaatatcc	ccgatatggt	14940
cgaggttgat	ggcgtagtcg	atgatgacga	tcgagcgtcg	gccttcgtct	tccgacagtc	15000
cagtacggcc	gagtttcgat	aagtagacct	tgacctctg	ctgtaaccgg	tcgacctggc	15060
gctcaagctg	cgcaatatcc	ttcaatctcg	cggcgttggt	ctggtggaac	gcgtccgaca	15120
ctctcaagag	cattctttcg	atcaggtctc	caacgccaa	cacttcgcgt	gcagcactcg	15180
tcagtgctac	gactgggtgt	gatagctcat	ccggatcgag	gaatttcggt	gtgttgctcag	15240
cctgttcgga	cgcggggatg	aggaccccc	tccaccgtgc	aagaagcggg	gaaaacggca	15300
tggccaggag	agccaggccg	atatttaacg	cgagatgcgc	gtccacggga	agtttgagcg	15360
ctgggaccgg	caacatcgaa	agccagctgg	ccaaaatc	ggccactggc	aggatcagca	15420
ggcagccgat	tgcccggaca	atcagattgc	ctagtgtgac	ccgcctggcc	gccgccgatg	15480
ctgataacga	cgcatcacc	ggaggtattg	cgccaccgaa	attggctcca	agaaccagcg	15540
ctacgacaag	cccgccagag	aggattccgg	tggaggtcaa	cgacaaaatc	agaaccaccg	15600
cagccaagct	tgaagatgac	gcgaatgcga	tggccgcgga	aaagatcatc	gcgaccggcc	15660
aggcgttttc	aagcagccg	ataaacgctg	cgagcggccg	cgaatggcga	agcggctcgg	15720

-continued

---

tcgcggtgct	gagtaggtgt	agtgacagca	gcatcagcgc	gataccgagc	aacgctgttc	15780
cgccgccctg	cctggaggtc	gaaccgcctc	gatagagaat	gatcccgacg	aagagttagg	15840
caggggacac	ccattcgatg	cccgttgcca	cgatccaggc	ggtgaccgcc	gttccgacat	15900
tggcgccgag	gaggacgacc	tgcgccatgc	gaggcttgat	aagctctcgt	tccacaaacg	15960
aggagaccat	cagcgccctg	gctgtagaac	tctgaagcgc	gactgtggcg	acgaaccctg	16020
atagaaacga	ccgggcgctg	cttttcgtgc	cggaggcgag	gcccgttcgc	aatctggccc	16080
caagcgcacg	ggaggtcccg	tccttcacct	gggccaggcc	gaacagcagc	aacgcgacgg	16140
cgccgcagag	atztatgaga	acgattgtcg	attccaagtc	gcgcctctcg	ctaattctgt	16200
tggaagtaag	ccttgatcag	cgaagcgctt	gctcgcaaca	gagatcgccc	agacgggtgc	16260
cgctctgatt	tgtctcgaga	gggaagcggg	tctcgtcgca	gacctgtccc	agcttgccgt	16320
tgatgagggg	gccgggtggg	gagcaggagg	gagagctgcc	caccggcatc	ttgcccgccg	16380
cgatcgcag	gcaagatgaa	ttccaactgg	tttcgtattg	tggtcgtcgc	gagtagcgct	16440
gatccgcgcc	ggttttgttg	acggcagttc	gacattgaat	gagtcacac	tcctcctaga	16500
agcttacgat	tttgggtcgc	tgcgaaagtg	agcgatttcg	atcacactaa	caaaggttc	16560
cgcttagcta	cccgaagcct	cttgatcccc	aacaatgcgt	aatcgtttgt	aatgtattgt	16620
tgcgttgcaa	caaaactgtc	acaaaacttt	catttgacga	agaacgatta	cgctcatagt	16680
atgagcgata	tagaccatth	cttgttgaaa	gggaagcgtt	ggtgttcaac	tcaactcagg	16740
accgccagc	gaaaatcgtt	gagcttttgc	gcgatgagca	gtttttggct	atcggcagat	16800
tgacggagca	cttccagatt	tcgggtggca	ctgagggcgc	cgacttgagc	gaactccacg	16860
aggccggact	cttgctcgca	actcacggcg	gtgaggtcag	cgtcacgcag	ggttacacag	16920
acaagccgaa	cgctgccccg	gccgtctgga	atcgggcgga	gaaggcgccc	atagccggag	16980
tcggtgccgg	tatgatcgtc	gagggtgaca	cggttcttct	ggacgccggg	acaactgcgc	17040
tggaagttgc	caagaagctc	gctgaccgca	gaaacctcac	cttcatctcg	aatggtctcg	17100
acattgtcga	ggaattgacg	agaggggagg	gcaaaagcat	ctattcgctc	gggggcgagt	17160
ataccgaaac	gaaccgctcc	ttccgcggcc	ctttggcgga	gcagtttata	cgccagttca	17220
atgtcgacaa	actcattctc	aacgcggcgt	cgatcgatgt	tgatcgcgga	ttgatctgta	17280
cgctcgtccc	cgtgaaccgc	agtgtcgctc	gcgccatgat	cgaagtttcg	agccgcgtga	17340
tcgtcgtcgc	ggatcattcg	aaattcacia	aatcgagcct	ctcggtgacg	gccaggatcg	17400
aggatgtcgg	cgtgatcgct	accgactctg	gaaccgaaac	catcattgag	acaataccgg	17460
aaaagctcgc	gaagaagttc	gttgttgcca	actaggttcg	acggccgcgc	gcgaattaaa	17520
gcctcattca	acagagctgg	gagggaaaacc	acatgcttaa	gaccaatcgc	aggaatttta	17580
tgatgggcac	cgcgccgata	gcggtcgctt	caactgccgg	tgcaagttc	accttcgcgc	17640
aagaacgcgc	ggcgcttcgc	ctgggtgta	acggattgcc	gaactcgcct	gagccggtca	17700
atgcaatcag	caatgtcggc	ccgcgtatcg	tcaaccagat	attcgacacg	ctgatagcgc	17760
gcgatttttt	gcgaagggga	gctccccgca	atgccatcga	cctcgttcgg	gcacttgccg	17820
agagctggga	gcgcatcgat	gagaaaatcg	ttcgttcaa	gctcggcgag	aaggtcatgt	17880
tccacgatgg	tggtgaactg	accgcggacg	atgtcgcata	caccttttct	tccgaacgcc	17940
tttggggctc	cgaagcgatc	aagaaaatcc	cgcttgggaa	atcctattcg	cttgatttcg	18000

-continued

---

atgagcccg	ggtcgaggac	aaatacacag	tcacacttcg	cacaaagacc	ccgagctatc	18060
tcatcgagac	cttcgctgca	tcctggatga	gccgcattgt	tcccaaagaa	tactacaaga	18120
aactgggagc	agtagatttc	ggtaacaagc	ccgttggaac	aggcccgta	aagttcgtgg	18180
aatttgttgc	cggcgaccgt	gtggtgctgg	aagccaacga	cgcctactgg	ggtccaaagc	18240
cgactgcatc	gaagatcacc	taccagatcg	tcgcagagcc	agccacacgc	gtcgcgggtc	18300
tgatcagcgg	tgaatacgat	atcatcacta	ccctgacgcc	ggacgacatt	cagctgataa	18360
atagctatcc	agacctcgaa	acgcgcggga	cgctcatcga	aaacttcac	atgttcacct	18420
tcaatatgaa	ccaagaagtt	ttcaaggata	agaagcttcg	gcgcgctctt	gctctggctg	18480
tcaaccggcc	aatcatggtt	gaagcgtct	ggaagaaaca	agcatctatc	ccggccggct	18540
tcaatttccc	gaactatggg	gaaaccttcg	atccaaagcg	taaggccatg	gaatataatg	18600
tcgaggaggc	caagcgtctc	gtcaaggaga	gcggtacga	tggaaacg	atcacctatc	18660
acacgatggg	caactattat	gccaacgcca	tgctgctg	gatgatgatg	atcgaaatgt	18720
ggaagcagat	cggtgtgaac	gtggtcatga	aaacctacgc	accgggtagc	ttcccccg	18780
acaaccagac	ctggatgaga	aactggtcca	acggccaatg	gatgaccgat	gcctacgcca	18840
cgatcgtccc	cgaattcggg	ccgaatggtc	aggttcaaaa	gcgttggggt	tggaaaggcg	18900
cgcccgagtt	caacgaatta	tgtcagaagg	tcacggttct	gccgaatggc	aaggagcgt	18960
tcgagccta	caatcgcatg	cgggacattt	tcgaagagga	ggcaccagcc	gtcattctct	19020
atcagccgta	tgacgtttat	gccgcgcgca	aggatgtcca	ctggaagccc	gtgagcttcg	19080
agatgatgga	attccgcaac	aatctcagct	tcggctgaac	tcgttaatcg	agaaggcggc	19140
gcgcgcgaga	tgctgcgctc	gccctggagg	agcctaagtt	gtcaaactca	ctaacagttc	19200
gcgacctgat	cgttcaggtg	ccggcgcg	atatcacaat	catcaacggt	gtcagttttt	19260
cactggatgc	agggcagacg	cttgactc	tcggggagtc	gggctgcgga	aagagcatga	19320
catgttacgc	aatagccaag	gcattgccgc	gagggatttc	gcagacgggc	ggaactattg	19380
agctggacag	tggaccaaa	gctgatggca	agccggctat	tgccatgatc	tatcaggacc	19440
cgaccagcag	cctcaatccg	gttcattcga	tcggttatta	tctcgagagc	agcctttatc	19500
gccatcaggg	tctggagggc	aacgatgcac	ggctcgaggc	catgcgtctt	ctcgaacgtg	19560
tcggtatcga	ccgcgccaaa	agtcggctgc	gatcctatcc	tcatgagtac	tcgggcggta	19620
tgaaccagcg	cgatcatgatt	gcccacgcgc	tcgctgcgaa	gccaagctg	atgatcgccg	19680
acgagcccac	gacagcgcct	gatgtcacga	cgcaggcgca	gatcctcat	cttctggagg	19740
aactcagatc	tgaaacagcg	atggcgcttt	tgatcgtgtc	gcacgatctc	ggcgtgatcg	19800
ccgcctcgc	ggatcgtg	gcggtcattg	attgcggaaa	gatcgtcgaa	acggcgccg	19860
ttgcggaatt	gttgagcgt	gcagcgcac	cttacgctc	cgccctgatc	ggctgcatgc	19920
caacgatcga	tgcggagcag	ctggagccgc	cagtaccaat	cccgggttct	gttctcttc	19980
tcgataacct	gcctgccgga	tgattattatc	acccacgctg	tcccgcgcg	agcgaacaat	20040
gttcctgtgag	ctttcccg	cctgtcaacg	tcggcgcg	ccatgacgcg	gcctgttacc	20100
atccggtgag	cccagatg	cgctctct	tcaggccgc	ggggtacca	aggtctttcg	20160
cctgaaatcc	gggcttttcg	taaagccatc	ggtacatgtg	gcggttaatt	caattgatct	20220
cgaacttgcg	ccgcgtgagc	gggtaggtgt	cgctcggtgaa	tctggaagcg	ggaatcgac	20280

-continued

---

actcggtaga	cttctgctcg	gcttgacctc	caccaccgag	ggagaggctc	tgttcgaagg	20340
actcgcgcac	acagatcgta	gcgcgcgaga	ctggtcgacg	ttccgaaagg	agacctccct	20400
ggttcagcaa	aaccgcgtg	ccgcccttaa	tccacaaatg	acgatcggcg	agcaaatcgc	20460
tgaagccgtc	tgggttcacc	gggttgccaa	cgcgggcgcg	cgtgatcggg	cggcaacct	20520
gctggaccgg	gtcggcctct	cgagcgcgat	gatgaaccgc	tatccacatc	agatgtcggg	20580
agggcagcga	caaagagtgg	ttattgcccg	agctttgatc	ctcgacccca	aattggtcgt	20640
attcgatgag	gcggtatctg	cgctcgcagt	ttctgtgcaa	gccaggctcg	ttgcgttgct	20700
tcggcagctc	tggcaggagc	ttgatctggc	gtatgtgttc	atcaccacac	acctgcgcac	20760
cgttcccat	ctcgttgatc	gtattgtcgt	gatgatctct	ggcgcgctcg	tcgaaaccgg	20820
cgatatcaaa	tctgtctatg	cgtggccacg	ccatccctat	accgcgcgat	tgcttgcttc	20880
cgccccact	atggaccoga	cggtccgcaa	tattgaacca	ccgatcaagg	gggaactcga	20940
cgctggcaag	gtaatgactg	gctgtgcctt	ccggtcacgc	tgtccttttg	caatcgaaag	21000
atgtgcgaaa	gaaacgccgc	tgcttcggcc	tgcgggcggg	caactggcag	cgtgccatag	21060
agcggaaag	ttcccgagat	ctatagtcgc	gcagactgat	ccggcgcaga	tgatggaggt	21120
ggcgcgatga	ttagtttcat	ccttgccacg	ctcgttcgag	cgatcattgt	catgatgctg	21180
accgtgacgt	ttgtcttcat	cattctcagc	ttgggtggag	atccagctcg	ggccatggtg	21240
ggcgagaacg	gcagcccgca	ggcgcctggc	gcattccgaa	cagcatgggg	cctggacaga	21300
tccattcctg	aacagttcct	gatctatctc	gccagcgcgc	tcagaggcga	cttcggggta	21360
tcggcagccg	acggctcgtg	ggtattcacc	gtcatctccg	aacgtatacc	caagacattg	21420
acgtgacgca	tctcagcatt	cgtgctaagc	atactcgtcg	gtatccccgc	gggaatcacc	21480
gcagcagtc	cccgcgacag	tgccgcgcac	aaatccgtca	tggcaggcgc	tgttctcggc	21540
tatagcgtac	caaatttctt	gctcggactg	acgctcatct	tcgtgtttgc	cgtctggttt	21600
cgtgtgctgc	cgagctcggg	aagttccacc	tggcagcatg	cgatcctgcc	cgttgccgaca	21660
ttcggctctg	tcaatgcggc	tgcgattgcg	cttcccttcc	cggtgatcga	agttctggaa	21720
cagccctata	tcgcagcggc	tcgcgcggat	ggtattccca	agtggaagt	ggttatccgc	21780
cacgcgctcc	ggaacgcggc	gatccccgat	gtgaccatgc	tgggtttcgt	tcagactggc	21840
ttgctggggg	ggtcggctat	cgtagagacc	gtattcgctc	ggccccggct	tggtagcggg	21900
ttgtttcag	caatcacgct	cagcgcctc	aatgtcgttc	aggcaatgat	ccttctcttc	21960
acggcgttca	tggtcaccgat	caacttgatc	gtcagatgtc	tctatgcggg	gctgaaccgc	22020
aaaatcaggc	ttcaaaaaaa	tggctaatac	gacactctca	ttaccaaaag	ggcaattcct	22080
gaaatctgcc	cgcgcattc	cgatcagtat	ctggctatgc	ctgatctgga	ttgcggttgt	22140
cgtctctcgt	gacatcacgg	cgcaatgggt	ggcaccgttc	gactatctgc	agcagtcacc	22200
tctggcgcga	tttgccgcgc	cggtgcccgc	gggacatttg	ctgggaaactg	attatctcgg	22260
tcgtgacgtg	ctcagcaaca	ttctggtcgc	taaccagacg	tctttgatga	tcgctcttgt	22320
cggttctctc	atctgcgcgt	tcacccggac	gaccctcggg	tttcttgccg	cacatttcgg	22380
tggctgggtc	gacaatctca	ttatgggctt	cgagacgcgc	aaggcttcgc	tgccctttat	22440
aatcttcgcg	cttgggtgct	tggcgtttct	cggtcaagc	ccgctgatac	ttcttatgcg	22500
cgttggtgct	gcacattcgc	agcgttatgc	gcggctgacg	cgtggcctcg	tcttgagtgc	22560

-continued

---

```

gaaccaggag ggatatgctg aggcggcgcg catcgtcggc gtaccatcgc ttcggatcta 22620
tgtgcccac gtccttccaa acattgctgg tccgctcatc gttcaaatca ccctgaattt 22680
cccggatata atcctgctcg aaagcggatt gagcttcctt gggcttgga tccaaccgcc 22740
ggaaacgagc cttgggctga tggtcgcaga cggcagaaac tataatcgca ttgctggtg 22800
gctcatcgcg ttgcccggcg ttgttattgt tctcaccacc ctgtccatca gcctgctcgg 22860
cgactacttg cgcgatcgtc tcgacgcgcg cctgcgataa acgaaaggaa tgaaaatgaa 22920
ccctctcaga atgacaccgg agacgatccg cttggggcgc catcgcggcc acagcgcgg 22980
cgcgccgaa acacgctcgc ggcgttctcg aaggcgtttg aattcggcgg acgtaacgtc 23040
acatcgaaa ccgatctcgg aattaccgc gatggagaac ttgtgctgat tcacgacaag 23100
accgtcgacc gcacaacaga cggtcacggc atcgttcacg acatgaccta ttcggagttg 23160
tcgaaactcg acgacgggag ttggttcagt tcggagtttg tcggagagcg tgtgccgcta 23220
cttaaggacg cgcttcagct cgcctgtag cttggcatta tctatcagct tgaactgaa 23280
atctacgacg cgaacgacgt cttcttccca aagttgagga cgctcatcga tgaattgggc 23340
tgtgcccgat tctccagtt ctcttccttc gactacgttc agtgaaggc cgtcaaggaa 23400
gcaattccc aggtaccgac cgttgccctg atgcattcgc gcctcatcga tccagccgcc 23460
cttgcccgtc aggccaatct cgatgccatg aacatcgaga tctaccattt cgcagcgggt 23520
gaagcccgc aactccacaa cgagggattt gccgcttttt gctacctacc gacggggtag 23580
catgaaaaac tcgcgcagta cggcgtggac gtgaaaacgc agctcgttca gtgggttcgt 23640
gaaggtcagt tggatcagtt gctcggagac gacgtcgcac aggtcggcag gctcaaggac 23700
cgtgcaaatg gctgatcggg acacatcgac ctctcacgca ggagaaacag cttcgattgt 23760
ggaagagatc gcgcgcaagg caggcgacct tgccctcgt catttccgct ccctatcgag 23820
cctgtcggtc gagaccaagg gacatctcga tcttgtgacg aaggcggaca aggaggtcga 23880
aacatttctt atcgcgcagt tgcgggagcg gtttcccgca gacggcatat ttggagaaga 23940
agggggcgaa atcaaacgcg gttcaggtcg tatctgggtg atcgatccga ttgatggaac 24000
attcaatttc gtccggcgcg gtcagaactg ggcaatttcc atcggctctct atgaaaacaa 24060
gcgtcctaca ttcggggtta tcttcgctcc agtccgaaac ctgatgtttg ttggagggaa 24120
aacggtggag acgaagctca acggcatggc ggtcaagcca cttccgcgc tcgatatgtc 24180
gcgtgcatcg accgggttca gctaccatcc ttoggcttcg acagcggaca ggctcgaagt 24240
catccgttat atctcggatg atctcaatat cagcttccgt ttctgctcgc ctgogacact 24300
ctcgatggtc gaggtggcca tggcgagac agatggctac gtttcattgg gagactcgac 24360
gtgggacgtt atggcagcgt tgcctatctt gagcaatctg ggtgttcgg atacgatcga 24420
ctgggacagg accgacttgt cggccaaact cgggttgct tgtggcagcc acgacttcct 24480
ggagaagggtg aagcgttgc tcgataaagt ggcctggcc gcataatgcg gcgcgcctct 24540
gacaataacg gtcacagatc gccactggtc gacgcggacc aatgcggcgc gatttgcgta 24600
cgctcgaaa ccctgttaa cagcgagatg gaagcgttc tgatcaccag ccgcgaaact 24660
ggccgctggg taatocccaa gggctggtcg gagggcagga aaaagctt 24708

```

```

<210> SEQ ID NO 8
<211> LENGTH: 8870
<212> TYPE: DNA

```

-continued

<213> ORGANISM: *Agrobacterium tumefaciens*

<400> SEQUENCE: 8

ttactccatt aggccgttct gcttcatcaa gtcacatga aacctccaga attcggttgt	60
gaaacgatcg acgatggttg agggagcgcc aattgctgac cggacttcga ggaatccggc	120
gtcaatgaag atcgagaacg gtcgcagtac gatcctgtcc gtgaactcga tcgcccgcc	180
tggatcgata attgcatgcc cggcgccctc gcggacgaga cttagcgagc tatgacgacg	240
gctcacttca attgacggcc ggcggtgaat accaccaatc gccacctcta cccgcatggc	300
gaagagagtg ccagctctct gttttataat acgctcaccg gcaaggctct gtggcgtagc	360
acggtaacgg ccagcaagtc gatgtcccat cgggacagcg acaacagcgg gaagcgaccg	420
ggtttcgatt agaaaacctt ggcgctcctg tgggccatcg gcataaccga tgcggccct	480
gccggacgca acggtctcca tgaccatgct tgagggcagc cccattaggc agacctggag	540
atgtggtctg tcacggatga actgagcaag aaaccgccc aagaggccgt ttgccagagc	600
aggcattgca gcaatgcgga gcgtccccgc tgctgcctg ccgatgtcgg cagccaggtt	660
gcctatatga ttaagcccga cgaacgccc atcgacctct tccacagcg tctttgctc	720
ctgtgtcggg ataatatggt tcccacgct ctcgaagagc tgcagtttg tcgctgttc	780
aaagtccttg attagggcag tgatggccgg ctgagtcacc agcattagtt cagcccccgc	840
cgatcatttc cccgtcagca tgaactgccc gaacgcccgc acctgcctga gattcataaa	900
cgcaccataa catctgctta ttcttcccgc gtcattatga attgaccga atgcatatcg	960
aatgtaaagc tcacctata aatcacaaact cttccgggcc aaccgggacg agacgtgagg	1020
cgaaacggtc gcgcgacggg aacgctatgg attattctca actcatgggc ttcggccctg	1080
atggctgggg ctatgatatg ctacgggcca ccgcatgac gatggctgtg gctttcagcg	1140
gcttcacaaat cggccttctc ttcggctgtc ttggagccgc cgcaagcttg tcaagctctg	1200
gagcgttca ggcagcgcca tctggataga cgaccgctc tcgtggcatc cccgacctac	1260
ttgtcatcta tctattctac ttcgggtcga gttctgtcat ttccaatgct gcttactct	1320
tcggaagtag tgggttcggt ggcgcttcca cgtttctgat aggtgccctc gccatcggcg	1380
tggtttcggg tgcctaccaa acgaggtcc tgcgcggtgc ggttctggct ctcaacaagg	1440
gcgagatcga ggctggtcga gcgtatggca tgggcgctt gctgtgttc cggcgatcg	1500
tcctgctca ggcggcgccg tacgccctgc ccggtgtcgg caatgtctgg caactggtcc	1560
tgaagaatc cgcctgatc tcggtgattg gcctcgtcga actgatgccc caggcccagg	1620
ttggctcggg ctcgaccgc caaccgttta gttctacct gactgcccgc gcgctctacc	1680
tgcttataac cttcgtctcg ggcaggtct tccgtctggc agagacgccc tctatgccc	1740
gactgcagag ggcggtgta tgcggtcga tcccgcattt ctctggcaga ccttcgtcgc	1800
ttgtcttcc ggcattccgc ttgcaactga actcgtgtg ttctcgggtg cgtcggaac	1860
ggttctcgtc ttcggcctgg cgtgatgccc cgtgtctcgc ctctggtggc ttgatctgcc	1920
ggcacgcttc tacattttcg cgtttcgtg cacgcccgtt ctggttcaga tctacatcat	1980
ttactacggg ctgagccaat ttcccgatgt tcggcatagc ttcatctggc cattcctcgc	2040
cgatgcttac tgggtgcgca tggcggcact ggcattgaat accgctgctt acacagcaga	2100
gatcatgctg ggaggtttgc ttcggtgcc cgcgggacag atcgaggctg cgaagcgtg	2160

## -continued

---

cggatgggc	agggtaaagc	tgtttcgccg	gatcgttatt	ccgcaggcga	tacgacagat	2220
gctgcccggc	tacagcaatg	aagtaatcct	gatggtcaag	tcgacatcgc	tcgcttcacc	2280
catcaccatc	atggaaatta	ccgggatcgc	agccaaactg	atctcggaaa	gctaccggac	2340
agtcgaagtc	ttctcctgtg	ccggggctat	ctaccttata	ctcaacttca	tcgtcgcgcg	2400
tcttttcacg	ttgcttgaat	gggctctctg	gcccgaacgg	cgaacaata	gacttacaac	2460
agatcctgtc	gaccgaaagg	gcgaactcca	tgcttaaccc	agttcgccca	gctgttcagc	2520
tgaaggacat	caggaagaat	ttcggtaacc	tggaggttct	ccacggcgta	tcgctgagcg	2580
ccaatgaggg	agaagtgatc	tccattctcg	gctcgtccgg	ctccggcaag	tcgacgcttc	2640
tgcatgtgtg	caacatgctc	gaggtcctca	atgccggaag	cgttgcgac	atgggcgaag	2700
agatcgtctc	cgagcatcgg	gccggccggc	tcgctcggcc	gaaggatttg	aagcaggctc	2760
acagactccg	cgagcggggc	gcaatggtct	tccaaggctt	taatctctgg	tcacatcaga	2820
cgatcctgca	aatgtcatg	gaggcggcg	tccacgtcca	ggggcgggat	cgtaaggcgt	2880
gccgcgacga	agctgagggc	ctgcttgaac	gcgtaggcat	tgcttccaaa	cgatgctcgt	2940
atccgtccga	gctatctggc	ggccagcaac	aacgcgctgc	gattgcccgt	gccctggcga	3000
tgccggccga	cgatcatgct	ttgacgagc	ccacatcggc	cctcgatccg	gaattggctg	3060
gggaagtctc	caaggtgatg	cgtgacctgg	cggctgaggg	gcgcacctatg	ctgatagtca	3120
cgacgaaaat	ggatcttggc	cgagatgtct	catcgcgaac	ggtcttcttg	caccagggcg	3180
tgattgcccga	agaaggcccc	tcctccgaaa	tgctcggcca	ccccgcacg	gatcgtctcc	3240
ggcaattcct	gcccgcgac	ggcggtagat	cccactgacg	acgaactcgt	cacggctcgc	3300
gtgtccagaa	aacaactgaa	cagaggtgaa	cttatgaaac	tcaaaactat	tctgtgcgca	3360
gcgcttctcc	ttgttgcggc	gcaggcgcca	gctcagggaa	agtcgattac	gatcgcgacg	3420
gagggcggct	atgcgccctg	gaacttctcc	ggccggggcg	gcaaaactcga	cggtttcgag	3480
atcgatctgg	ccaacgctct	ctgcgaaaag	atgaaggcga	aatgccagat	cgctcgtcag	3540
aactgggatg	gcatcatgcc	ttccctgaca	ggcaaaaat	acgatgcgat	catggccgcg	3600
atgagcgtca	cgcaaaaacg	gcaagaggtg	atcggcttct	ctatccccta	cgctgctggc	3660
atcaatgggt	ttgcgctgat	gggtgacagc	aagttggccg	aatgccagg	cttgggtgaa	3720
acctattctc	tcgatagcca	ggcggacgct	gccaagaagg	ctattgccga	tatttcgagc	3780
ttcctcaacg	ggacaactgt	tggcgtgcag	ggctctacca	cggcgtcgac	cttcttgac	3840
aagtatttca	aggggtcggc	cgacatcaag	gaatacaaat	cagtogaaga	gcacaacctc	3900
gacctgacga	gtggctcggc	ggatgctggt	ctggcgaacg	cgaccgtgct	tgccgcagca	3960
atcgagaagc	cggaaatgaa	ggcgcggaag	ctcgtcgggc	cgctgttctc	cggtgccgag	4020
ttcggcgtag	tcgcccgttg	tcttcgcaag	gaagataccg	ccctcaaggc	tgatttcgac	4080
gcggccatca	agggcgcaag	cgaagacggc	acgatcaaaa	cactgtcgtc	gaagtgggtc	4140
aaggtggatg	tcacccccca	atgatcgatc	gcttttaggg	atacagcgtt	cgcgccgtat	4200
cccggccccg	taaataccaa	acaagtccaa	gatgcaatgt	acgaagtga	tatgaccatt	4260
atcggagggc	ggctcgttgg	cgcgtccatt	gcctggggtc	tcgcacgctc	aggaacgaag	4320
cctctcgtcc	tcgacggcgc	ggatctggat	ctgcgcgat	ctcgcgcaa	tttcgcgtg	4380
gtctgggttc	agggtaaagg	actgcatgcg	cctcattacg	ctctctggtc	cgacgcgtcc	4440

-continued

---

gccaggagat	ggccgacgat	ggcaaacacc	ctgctggacg	acagcggcat	cgacgtgggt	4500
cttcagcaag	acggagcatt	tacattcgct	ctttctgaag	aggaactcga	ggcaaatcgc	4560
caagatatgg	aaagcattga	gctcgaaaaca	aatggtcgag	ctccacaatt	tgaagttttg	4620
gatcgcacg	agacgcttga	ccgcgtgctg	ggtatcggcc	ccgaagtgtg	aggctcaatc	4680
tattgcgctg	cggacgggca	cgtgaatgcg	ctccgcctgt	ttcatgcgct	gcatgccgcg	4740
atggaaaggc	aaggagcgac	ttatcgtccg	aaccatcccc	tgacgtccat	cgagccaacg	4800
actggaggct	tcattctcaa	ggcggaggcg	ttttctatcc	tgtaacgccc	tatcgtcctt	4860
gctgccggcc	tcgacaataa	aaggctcgcg	cccatggctg	gcctgtcatg	cccactcaag	4920
cgcagcaagg	gtcagatcct	cgtaaccgag	aagaccaga	ccgctctacc	ttgcctctcg	4980
gcaggaatgc	ggcaggccga	tgaggcggg	atcatgatcg	gcgacagcga	agaaacggac	5040
aacaccagaa	tttcatcgtc	accagacatt	agcgcctgcc	tcgccagccg	cgcttgagg	5100
atattccctg	ccctttcggg	tctgaatgtc	gtgcgtagct	ggaccggctt	cagggtcaag	5160
acggccgacg	gagtacctat	ctacgatcat	tccgagcgtc	atccgggtgc	ctttctgtc	5220
gcttgccatt	caggagtcaac	cctggccgcc	aatcacgctt	tgatcgtggc	tcagcagatc	5280
gcggcagggc	agctcgaaga	cgagctttcc	gtttttccg	cccggaggtt	ccatgctcaa	5340
caggctgtat	aggatagaac	acccattcca	attcagcttc	gatggcacgg	tcattccaggc	5400
cgaaacggcg	gacatgctcg	caaccgcgtt	gctcgcagca	ggggttgatc	gcgttcgcca	5460
gtccatcgtc	agcggtagcc	ctcgcgcacc	ctactgccta	atggcgctct	gcttcgagtg	5520
tctggtcacc	gtggatggcg	ttcaaaaccg	ccaggcttgt	ttgaccgagg	ttgaaaacgg	5580
ctagacgggt	ctgagccaga	agggcgccac	cacgatcccc	ggacttgcca	gaagcgatga	5640
gggaacggcg	ctatgacctc	tcgcgaagtt	tcggtagcaa	cggatctgtc	ggatttctac	5700
gacctgctcg	tgatcggcgc	aggacctgca	ggtatggctg	ccgcagtcga	agcatctgcg	5760
agtggcggcc	gggtagcggg	actggatgaa	aaccctcgtc	cgggcggaca	gatttatcgc	5820
gagatcacac	gaaacagctc	cgacagacga	acctatctcg	gccccgacta	ttggaaggga	5880
cagccactcg	cagagcggtt	ctgtctcagc	aacgtcgact	acgcatcacg	cgctacggtg	5940
tggagcctgg	agacccggga	caagacggcg	ggcaggcgc	gcaatgtcgt	cggcgtgacc	6000
gtggccggct	cggcgcgaat	ggtagaaacc	aacgcctcgc	tgcttgcaac	cggcgcgacg	6060
gagcgtccca	tgctgtccc	cggctggacg	ctgccaggcg	tgatgaccgc	cggcgtgccc	6120
cagattgcgc	tcaaggcagc	cggagcgatg	ccagcggggc	ctggtgtcct	gatcggctgc	6180
gggcgccttc	tctatctcct	ggcaagtcag	ctcgttgatg	cagggtgtcc	cgatctcacc	6240
gtcctcgata	ccgctcaatc	tcctttcccg	ggcgcggctg	tccggcacat	gcctgaattc	6300
ctgctttcgc	cttacgtcct	gaaagcctc	ggcctgctat	tgaaggtcag	gcgtcatgca	6360
caagttggtt	acggcgtcag	atcgtcgcg	attatcggaa	gccagcacgc	ggaagcgtc	6420
cgctacggcg	tcggacaagg	cgagcgttca	ataccagcca	agtcgggtgct	tctgcatcaa	6480
ggcgtgatcc	cttcagacaag	cttgtccaat	gcggccggat	cgcagttgca	atggaatgac	6540
gaacagcgtg	cgtttcagcc	agtcacagat	catgacggtc	ggaccacaaa	ggctggcatt	6600
tatgtcggcg	gcgacggcgc	ggggatcgcg	ggcgcacaag	ccgcagaggt	cagtgggcga	6660
cttgctgcac	tcgcccact	tgctgacttg	aagctggtaa	gcactcagac	atcgtcaacg	6720

---

-continued

---

tcgatcaaat cccggcacgc gcaggcccgc cggttcctgc gcggaagagc gtttctggac 6780  
gcactctata cgccgcggca aagctttctc gcgccatcgg caccgaaac gatcgtctgc 6840  
cgctgtgaag aaatcacagt gcgcaagctt cgcgaggcaa tagcgttg accgcccggc 6900  
ccgaaccagt tgaaaacctt cgttcgttgc gggatggggc agtgccaggg gcgcctttgt 6960  
gctgcaaccg tcaccgaaat catggtcgcg gaagaacgaa aagtcagtcc cgcggatgtg 7020  
ggaacataac ggcttcgttc accggtcaaa cccgttcgcc ttgctgaact tgcacacctc 7080  
ccgcacacgg cccgcgcgct gaaagcagta accggcaggg atcctgtcga tcacgacct 7140  
accgaaacag gacatatctt atgacagaac gccttatccg gatgccgaca ttgcatcgcg 7200  
tcgtcaagca caacggcatt gcctatatcg gcggtatcgt tgcgtgacgat gaaagcctcg 7260  
atatggaggg ccagaccgcg caggtgctga ccaagctcga cgcatatctt aaggaggctg 7320  
gttcggcgcg tgccaacctc ctgtccgcc a gatctttat caccgacatg aacgtgaagc 7380  
cgcaaatgga tgcggtgtgg aaggaatggt tggccaccgg agcagttgcc tgcccgcgcg 7440  
actatcggcg ttgcccagct tggtagacag acgctgatcg agcttgttgg ccacggccca 7500  
ttactgaaaa ccggcagcct gttggaagac tgcggaacgc ggtcgttatt ggagatctcc 7560  
atggaatgca ccctggtcg agcgcgctg caagtggag ccgggcatct cggatgtgag 7620  
atggcccccg cagcgttcag gatcgtggc cttgccacgg cgcttgaaga gctgggcaca 7680  
gggcacggga tgcgcccaca ctgggtccga agtttgttcc ggcgaatgcc catccggatc 7740  
cttccgtcca tcgcccacg cgtgcgcct gaattggccc tcagaaaatt ggaggaagaa 7800  
tgccaatcga ccctaaactg aacgtagtcc ctttcatcag tgcgaccac atgatgaaac 7860  
tggtagtcaa ggtaggcatc gacaccttc tgactgaact cgccgcccag atcgagaagg 7920  
atctccgccc gtggccaatc ttcgataaaa aaccacgcgt tggcagccac tcacaagacg 7980  
gcgttatcga actgatgccg accagcgacg gttcgtctca cgcttcaaa tatgtgaacg 8040  
gtcatccgaa gaacacacat caggggcgct aaaccgtgac ggccttcggc gtcctttcgg 8100  
atgttggtaa cggctatccc ctgctgcttt cggaaatgac gatcctgaca gcacttcgca 8160  
cggcggccac atccgctctc gccgcaaat atctggctcg tcccactcg aagacgatgg 8220  
ccattatcgg caatggcgcc cagagcgagt ttcaggctcg tgccttcagg gcaattctcg 8280  
gaattcagaa gctcaggctt ttcgacatcg acacgtcggc caccagaaaa tgcgctcgaa 8340  
acctgaccgg gccgggattt gacatcgtcg aatgcggctc cgtcgcgag gctgtcgaag 8400  
gagccgatgt catcacaacc gtcacggccc acaagcagtt cgcaacaatc ctttccgata 8460  
atcatgtcgg gccgggctc catatcaatg cggtcggcgg ggactgccg ggaagaccg 8520  
aaatcagtat ggaagtgctg ctgcgctccg acattttcgt agaatatccg ccgcagacct 8580  
ggatcgaagg cgatattcag cagcttcac ggaccatcc cgtcaccgag ctttggcagg 8640  
tcatgaccgg tgaaaaaaca ggacgcgctg gtgaccggca gatcacaatg ttcgacagcg 8700  
tcggatttgc gatcagggat ttttcggcgc ttogttacgt ccgcgcgaag ataaccgatt 8760  
tcgaaatgtt taccgaactc gacctcctgg ccgatcccga cgagccgca gacctctatg 8820  
gcatgctgct tcgctgcgag aaaaagctgg agccgacagc cgtcggataa 8870

<210> SEQ ID NO 9  
<211> LENGTH: 2767  
<212> TYPE: DNA

-continued

<213> ORGANISM: *Agrobacterium tumefaciens*

<400> SEQUENCE: 9

```

gaatttgctc aagaagactt ggagtagtca tccgctgctt tcaacatcca ttgcggtctt    60
ctcactcggc cgcaactacg gcagcttcgg gctcggaaac ggatcatggag aacggccccc    120
ctcatggcag ctattttcat ccttagcctc caagactgga ccatccgttc acggtcccat    180
ccaggccatt ccatagagtt atactttcga aaagcgcaac attcattttt acgtcaacaa    240
atgctggattc aaaattgtcg aatattataa ttgcggtgat cctgatccgc acgacacaga    300
agagagtgat gccagcgaca gtggcatggt aaggttttaa aaagacatgt tgtataccgg    360
aaactgattg cacgtcccg cccagagtctt cgcctatcta gactgtttct cccgcagctt    420
ctccttcata ccggccagca ggtaacttag cccactctcg aacacggctt cgggtgtgac    480
cgttccaacc gtggggaagc ttcaagagc ttttgcaagc agggggaact gtgcgagttc    540
tgccatgatg tcggtattgg gcgagtgtg atccatccga tcccgttcgc gctctccctg    600
ttcttcgagc acatagccga ccacatagcg cgtgatgcaa atggcgatcc caaatgcctc    660
ctctcgggaa aatcccgtct cgacatagag ctgcaagacg ctttcacat cggcaaacctc    720
ggcgacggaa ggccgtgtac ctgcggttag ccgcgcgcca tcgcgcacgg tgagcaaggt    780
cttgcaatg ctgcgggcat tggccagcgt gaacgtaatc caatcctggc cgggtttggg    840
caggcgatcg cagtataaac gcagcagcat ttcggaattg atggcgtcga gcaattccgc    900
cttgttgagg aagtgccagt agagcgcggc ctgctggagc cccagccgtt ccgcgagctt    960
gcgggtggac agcttgtcta ttcccacctc gttcaggagc cgcagggcct catcgacaat   1020
gcgatcacgt ccgatagcca tggatagaat ccctcttgac ttatcattga taagggagct   1080
tatcgccgat aaatttatca ctgataaagc tttatagga tgccatacca tgaacaaggc   1140
ccttatcgtt attctctcaa ccgttgccct cgacgccatt ggccgaggcc tgatcttccc   1200
gatcctgccc gacatcttgg tcgaggtgac tggcgcgggc gacatcgggt tcctctatgg   1260
ggatcatgct ggggtattcg ccgcatgca atttgtgttc tcgccgatcc ttggtgcgct   1320
cagcgaccgg ttccggtcggc gcccggtctt gttgctttct ttggccggtc ccctgcttga   1380
ttaccctggt atggcaatct ccccgctcgg ctgggtgctc gtcgctcggc gggccatggc   1440
ggggatcacc agcgcaataa tggcgggtgc aagcgctac atcaactgaca tcaccccagc   1500
cgagcagcgc gcgagcggg ttggcacggg tgggtccgtg atgagcctgg gctttatcat   1560
cgggtcccgc attggtggcg tcattggcgc ctggtggctt cgggcacat ttcttgtggc   1620
agccctgttc aatggcctca acctgttctg cgcgctgttt gttctgccc aaagccgaaa   1680
ggccggtccg ggcaagtttg cgttcaagga acttaaccgc ttggcgccat tgggtgtggc   1740
ttggaatttc aagccgctcc tgccacttgt aaccgtctct gtcgtctctg gctctgtggc   1800
cgccatcccg ggaacgatct ggggtctcta tggcgccgag cggttcggat gggattcggg   1860
gcataatggc ctgtcgtat cggttttcgg cgtcagtggc gccctggcgc aggoccttct   1920
cgtcggggcc ctctcgcgcc gctttgtgta tttgggcacg ttgatgatcg gcgttggtct   1980
tgacatgctg gcttatatgc tgatggcctt cgccaaccag agctggatgg gctacgaggg   2040
agcggccctg tttgcattgg gcggcgttgc catgcccggc ctgcaatctc tggtaacccg   2100
ccgctgtgag gatgatcagc agggccagtt gcaggcgtg ctcgccagcc tcatgagcct   2160

```

-continued

---

ggcgggtata atagggccgg tgctgaccac cgcagtgttc tttccacca aaagcatctg	2220
gatcgggacg atctggctgg tgggtgccgc actttatctt ctcgccttgc cgtgttcgc	2280
aacggtgaaa accccgaagg ctgtggcggc ttaaagactc cggacgtcat gcgcctcgtg	2340
tttatcttcg ccgccgattt gatgtcatcg cttcgaatca aacaggtacg agcgttgaac	2400
tggaaatcag aacgccagtt caacggcatc tccaagcgac gcttttcggc cgatcacgat	2460
tcgcagcgca tattagctat ggctggaagg tcatgaagga tcatggctgc cgcgccgaaa	2520
gccacggtat ccgcagcaga gacttcgggc aaaagaattg ccggctgaag aactttcgca	2580
ccctcgaacg gatcgggaaac tgaaaaatcg agccaggggc ataattgttc ccggatcattg	2640
cggggcaggg agcccccgac atatattgcc tcaggggcca gaacagcgtc gagggaaaat	2700
cccagtctgg cagtttctgg cccggcccga tttatccacg ttatctacaa gctcttgatt	2760
gctgttc	2767

---

What is claimed is:

**1-13.** (canceled)

**14.** A method for transforming a plant cell comprising,

- a) obtaining an *Agrobacterium* whose genome contains an inducible regulatory sequence operatively linked to a nucleotide sequence encoding a levansucrase;
- b) introducing a DNA construct into a T-DNA element of said *Agrobacterium*; and
- c) inoculating at least one plant cell with the *Agrobacterium* of (b) for a time sufficient for mobilization of the T-DNA element from the *Agrobacterium* to the plant genome.

**15.** A method for transforming a plant cell comprising,

- a) obtaining an *Agrobacterium* comprising a first recombinant nucleic acid construct containing an inducible regulatory sequence other than sacR, operatively linked to a nucleotide sequence encoding a levansucrase;
- b) introducing a second DNA construct into a T-DNA element of said *Agrobacterium*; and
- c) inoculating at least one plant cell with the *Agrobacterium* of (b) for a time sufficient for mobilization of the T-DNA element from the *Agrobacterium* to the plant genome.

**16.** (canceled)

**17.** The method of claim 14, wherein said regulatory sequence comprises the *E. coli* lactose operon (SEQ ID NO: 1).

**18.** The method of claim 14, wherein said regulatory sequence comprises the pi 2(noc) promoter (SEQ ID NO: 2) and noc 1 operon (SEQ ID NO: 2).

**19.** The method of claim 14, wherein said regulatory sequence comprises the P<sub>BAD</sub> promoter (SEQ ID NO: 5) and the araC cis element (SEQ ID NO: 4).

**20.** The method of claim 14, wherein said sequence encoding a levansucrase is a sacB open reading frame.

**21-37.** (canceled)

**38.** The method of claim 14, wherein the regulatory sequence comprises the traCDG promoter (SEQ ID NO: 7) and the occ promoter (SEQ ID NO: 8).

**39.** The method of claim 20, wherein the nucleotide sequence encoding a levansucrase contains a second copy of a sacB open reading frame.

**40.** The method of claim 15, wherein the regulatory sequence comprises the traCDG promoter (SEQ ID NO: 7) and the occ promoter (SEQ ID NO: 8).

**41.** The method of claim 14, wherein the nucleotide sequence encoding a levansucrase contains a second copy of a sacB open reading frame.

**42-43.** (canceled)

**44.** The method of claim 15, wherein the nucleotide sequence encoding a levansucrase contains a second copy of a sacB open reading frame.

\* \* \* \* \*