US 20030027183A1

(19) United States

(12) Patent Application Publication (10) Pub. No.: US 2003/0027183 A1

Kumagai et al. (43) Pub. Date: Feb. 6, 2003

(54) METHOD OF IDENTIFYING A NUCLEIC ACID SEQUENCE IN A PLANT

(76) Inventors: **Monto H. Kumagai**, Davis, CA (US); **Guy R. della-Cioppa**, Vacaville, CA (US); **Robert L. Erwin**, Vacaville, CA (US); **David R. McGee**, Vacaville, CA (US)

Correspondence Address:
HOWREY SIMON ARNOLD & WHITE, LLP
BOX 34
301 RAVENSWOOD AVE.
MENLO PARK, CA 94025 (US)

(21) Appl. No.: **10/120,630**

(22) Filed: **Apr. 10, 2002**

Related U.S. Application Data

(60) Division of application No. 09/359,301, filed on Jul. 21, 1999, which is a continuation-in-part of application No. 09/232,170, filed on Jan. 15, 1999, now abandoned, which is a continuation-in-part of application No. 09/008,186, filed on Jan. 16, 1998.

Publication Classification

(51) **Int. Cl.**[7] **C12Q 1/68**; A01H 1/00

(52) **U.S. Cl.** **435/6**; 800/278; 800/280

(57) ABSTRACT

The present invention provides a method of compiling a plant functional gene profile, a method of changing the phenotype or biochemistry of a plant, a method of determining a change in phenotype or biochemistry of a plant, and a method of determining the presence of a trait in plant. The methods comprise expressing transiently a nucleic acid sequence of a plant into a host plant to affect phenotypic or biochemical changes in the host plant. A viral vector functional genomic screen has been developed to identify nucleotide sequences in transfected plants by systemically knocking out endogenous gene expression in an antisense mechanism. Once the presence of a trait in a plant is identified by phenotypic or biochemical changes in the host plant, the nucleic acid insert in the cDNA clone or in the vector that results in the changes is then sequenced. The present invention exemplifies that genes encoding GTP binding proteins in one plant can silence endogenous gene expression in a different plant.

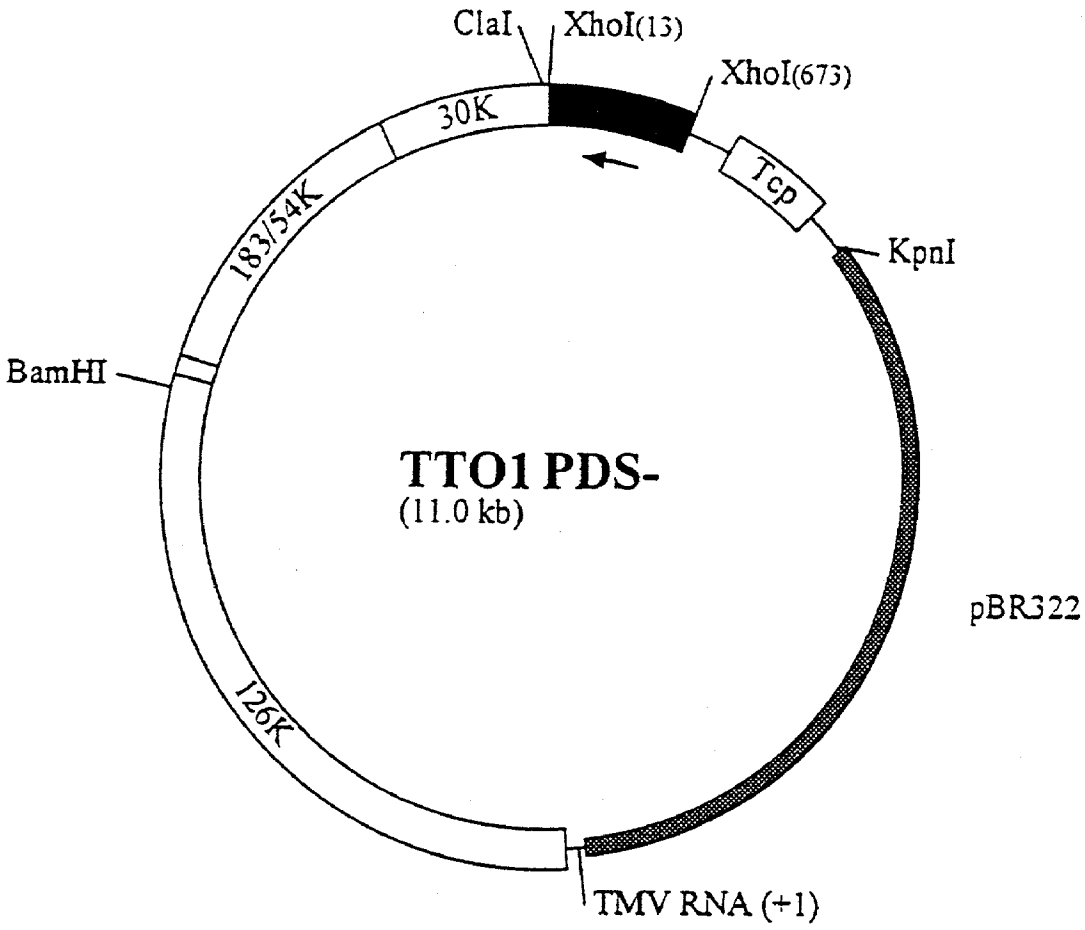
ClaI
XhoI(13)
XhoI(673)
30K
Tcp
KpnI
183/54K
BamHI
TTO1 PDS-
(11.0 kb)
pBR322
126K
TMV RNA (+1)

Figure 1

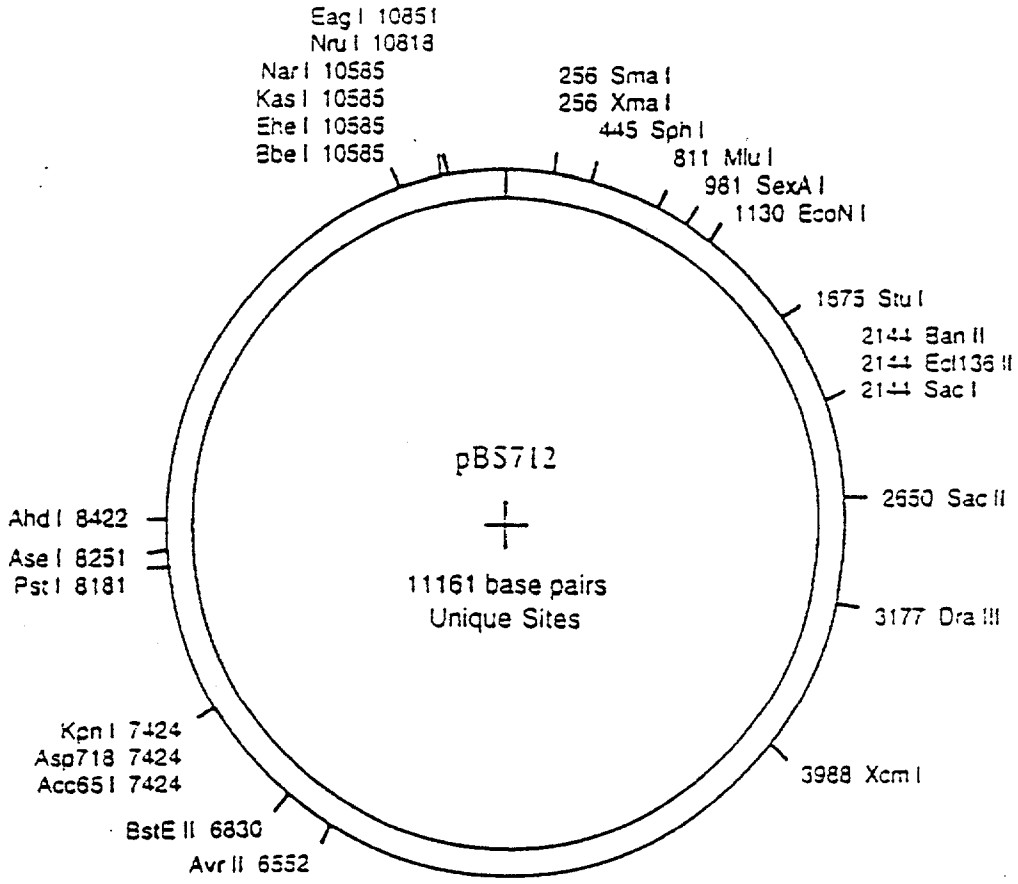
Eag I 10851
Nru I 10818
Nar I 10585
Kas I 10585
Ehe I 10585
Bbe I 10585
256 Sma I
256 Xma I
445 Sph I
811 Mlu I
981 SexA I
1130 EcoN I
1675 Stu I
2144 Ban II
2144 Ecl136 II
2144 Sac I
2650 Sac II
3177 Dra III
3988 Xcm I
Kpn I 7424
Asp718 7424
Acc65 I 7424
BstE II 6830
Avr II 6552
Ahd I 8422
Ase I 8251
Pst I 8181
pBS712
11161 base pairs
Unique Sites

Figure 2

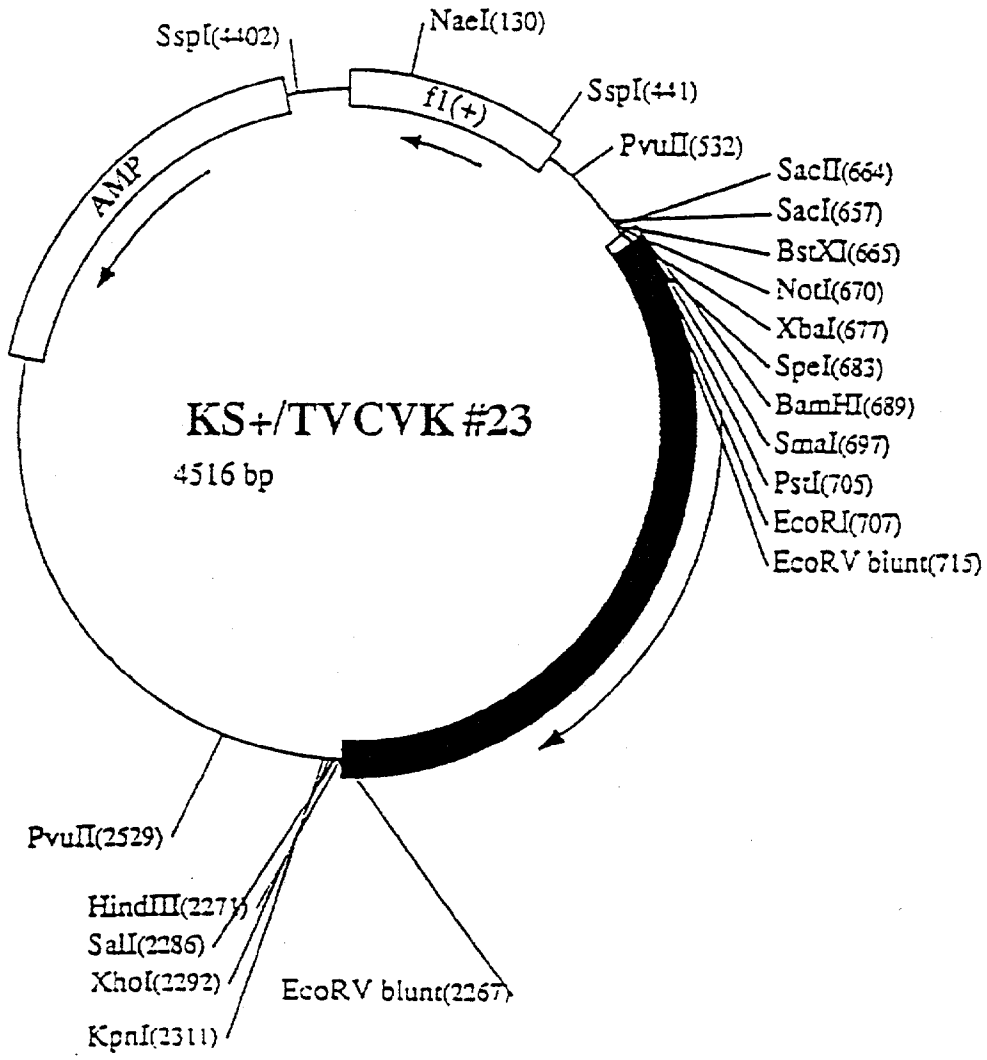
SspI(4402)
NaeI(130)
SspI(441)
f1(+)
PvuII(532)
AMP
SacII(664)
SacI(657)
BstXI(665)
NotI(670)
XbaI(677)
SpeI(683)
BamHI(689)
SmaI(697)
PstI(705)
EcoRI(707)
EcoRV blunt(715)
KS+/TVCVK #23
4516 bp
PvuII(2529)
HindIII(2271)
SalI(2286)
XhoI(2292)
KpnI(2311)
EcoRV blunt(2267)

Figure 3

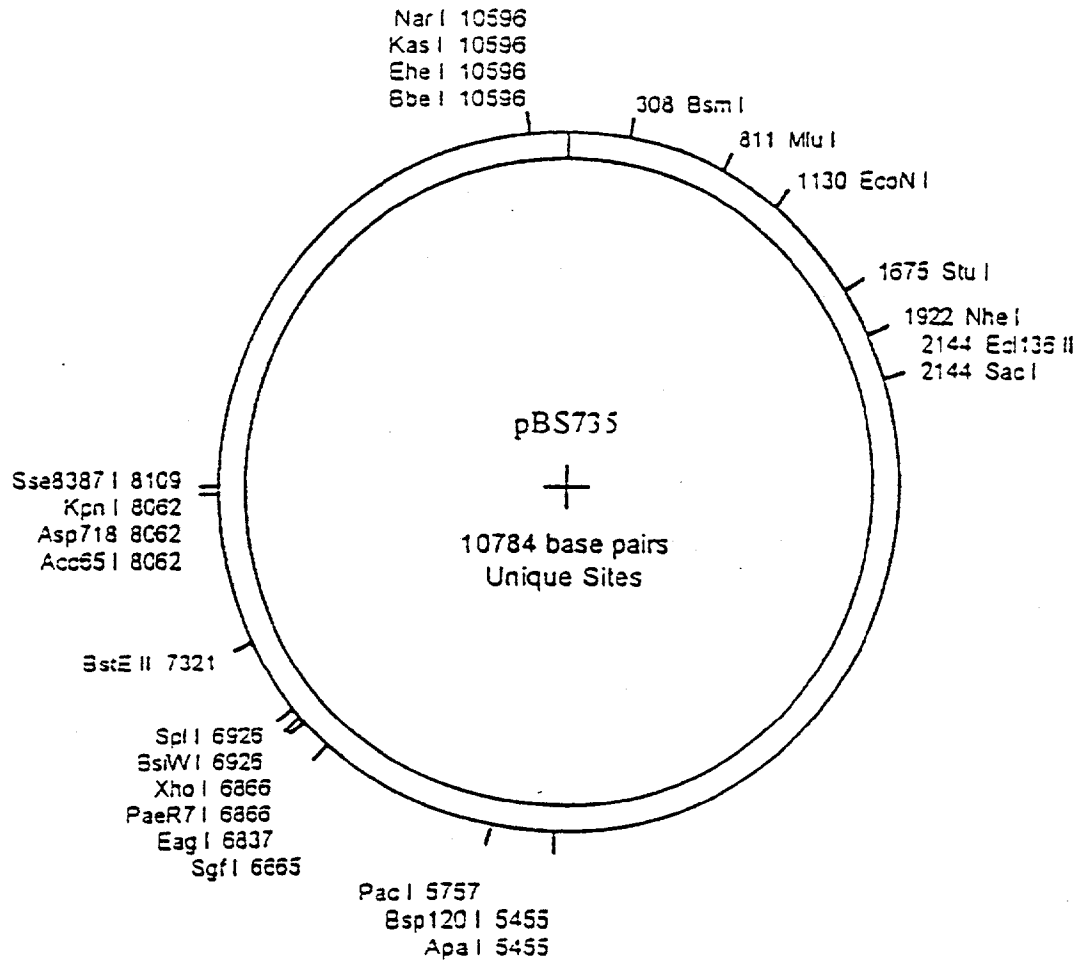
Nar I 10596
Kas I 10596
Ehe I 10596
Sbe I 10596
308 Bsm I
811 Mlu I
1130 EcoN I
1675 Stu I
1922 Nhe I
2144 Ecl136 II
2144 Sac I
pBS735
10784 base pairs
Unique Sites
Sse8387 I 8109
Kpn I 8062
Asp718 8062
Acc65 I 8062
BstE II 7321
Spl I 6925
BsiW I 6925
Xho I 6866
PaeR7 I 6866
Eag I 6837
Sgf I 6665
Pac I 5757
Bsp120 I 5455
Apa I 5455

Figure 4

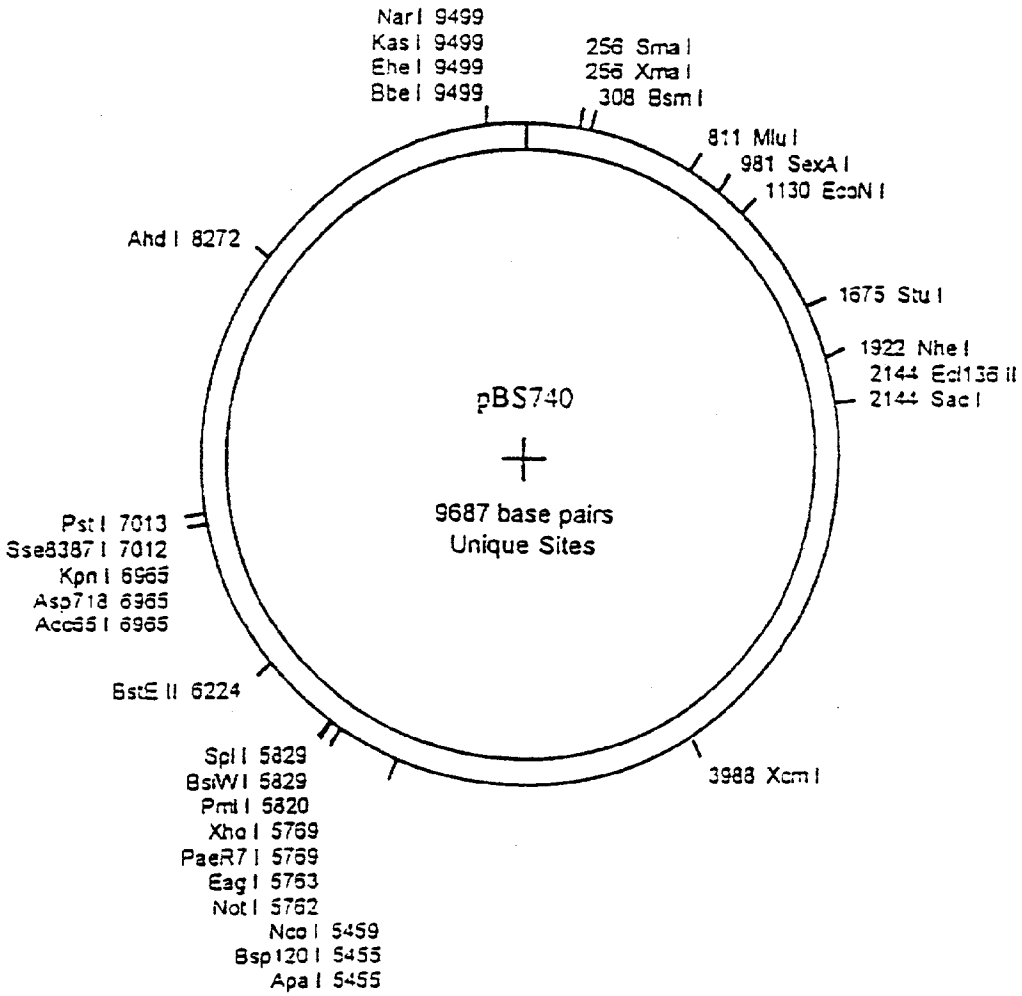
Nar I 9499
Kas I 9499
Ehe I 9499
Bbe I 9499
256 Sma I
256 Xma I
308 Bsm I
811 Mlu I
981 SexA I
1130 EcoN I
1675 Stu I
1922 Nhe I
2144 Ecl136 II
2144 Sac I
3988 Xcm I
BstE II 6224
Spl I 5829
BsiW I 5829
Pml I 5820
Xho I 5769
PaeR7 I 5769
Eag I 5763
Not I 5762
Nco I 5459
Bsp120 I 5455
Apa I 5455
Pst I 7013
Sse8387 I 7012
Kpn I 6965
Asp718 6965
Acc65 I 6965
Ahd I 8272
pBS740
9687 base pairs
Unique Sites

Figure 5

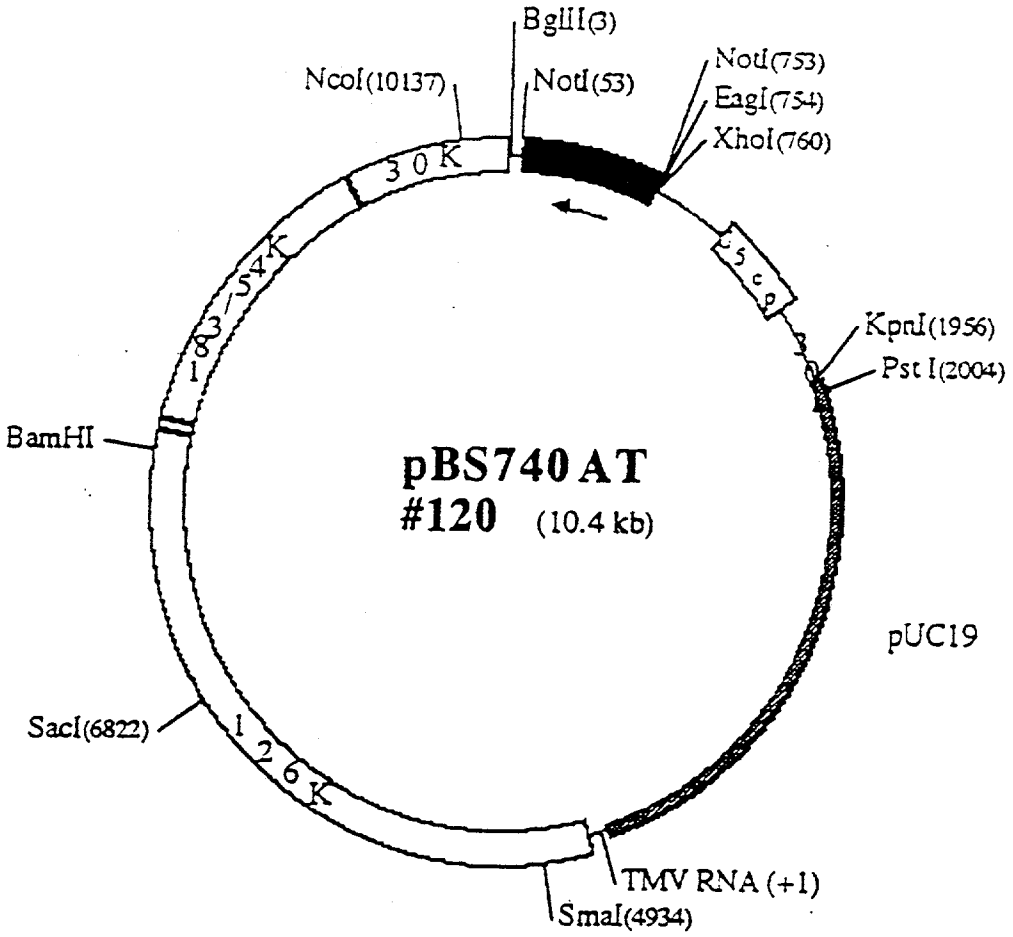
BglII(3)
NcoI(10137)
NotI(53)
NotI(753)
EagI(754)
XhoI(760)
30K
54K
KpnI(1956)
Pst I(2004)
BamHI
pBS740AT
#120 (10.4 kb)
pUC19
SacI(6822)
126K
TMV RNA (+1)
SmaI(4934)

Figure 6

Nucleotide sequence alignment of 740 AT #120 to AA042085

```
740 AT #120   TCCGAAACATTCTTCGTAGTGAAGCAAAATGGGGTTGAGTTTCGCCAAGCTGTTTAGCAG
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AA042085      TCCGAAACATTCTTCGTAGTGAAGCAAAATGGGGTTGAGTTTCGCCAAGCTGTTTAGCAG

740 AT #120   GCTTTTTGCCAAGAAGGAGATGCGAATTCTGATGGTTGGTCTTGATGCTGCTGGTAAGAC
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AA042085      GCTTTTTGCCAAGAAGGAGATGCGAATTCTGATGGTTGGTCTTGATGCTGCTGGTAAGAC

740 AT #120   CACAATCTTGTACAAGCTCAAGCTCGGAGAGATTGTCACCACCATCCCTACTATTGGTTT
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AA042085      CACAATCTTGTACAAGCTCAAGCTCGGAGAGATTGTCACCACCATCCCTACTATTGGTTT

740 AT #120   CAATGTGGAAACTGTGGAATACAAGAACATTAGTTTCACCGTGTGGGATGTCGGGGGTCA
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AA042085      CAATGTGGAAACTGTGGAATACAAGAACATTAGTTTCACCGTGTGGGATGTCGGGGGTCA

740 AT #120   GGACAAGATCCGTCCCTTGTGGAGACACTACTTCCAGAACACTCAAGGTCTAATCTTTGT
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AA042085      GGACAAGATCCGTCCCTTGTGGAGACACTACTTCCAGAACACTCAAGGTCTAATCTTTGT

740 AT #120   TGTTGATAGCAATGACAGAGACAGAGTTGTTGAGGCTCGAGATGAACTCCACAGGATGCT
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AA042085      TGTTGATAGCAATGACAGAGACAGAGTTGTTGAGGCTCGAGATGAACTCCACAGGATGCT

740 AT #120   GAATGAGGACGAGCTGCGTGATGCTGTGTTGCTTGTGTTT
              | ||||| ||||||||||||||||||||||||||||||||
AA042085      GNATGAGNACGAGCTGCGTGATGCTGTGTTGCTTGTGTTT
```

Figure 7

Nucleotide sequence alignment of 740 AT #120 to *Oryza sativa* D17760

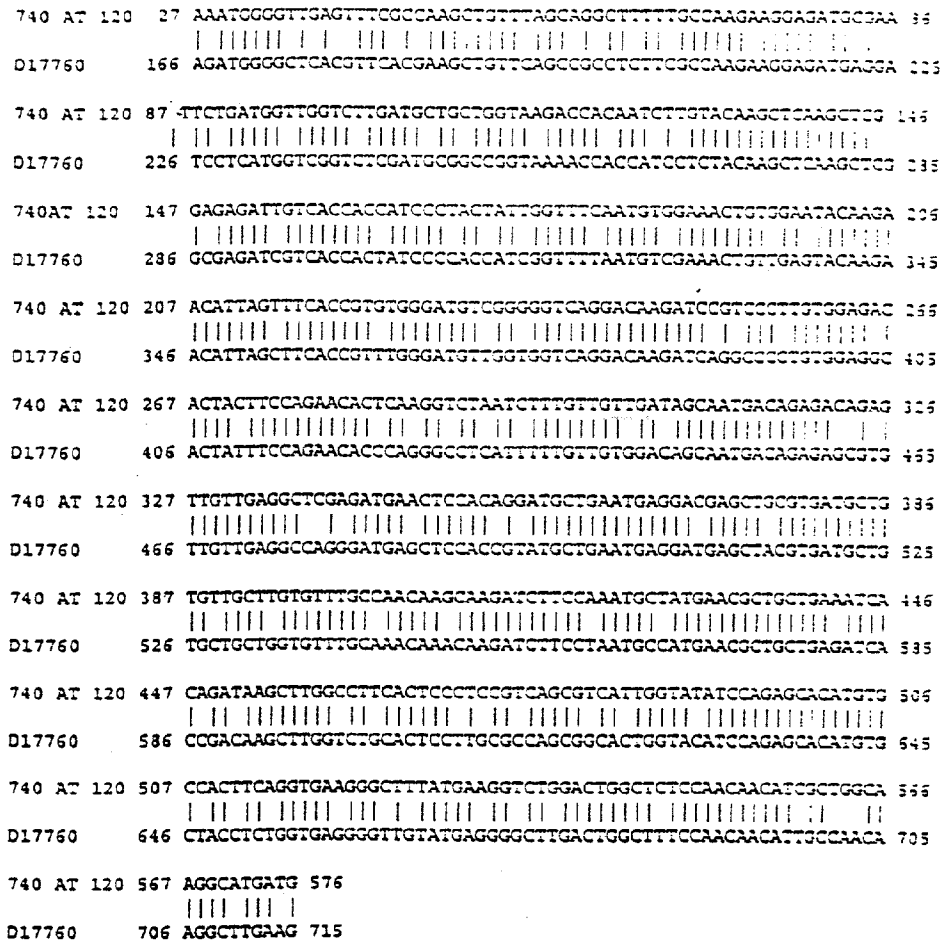

Figure 8

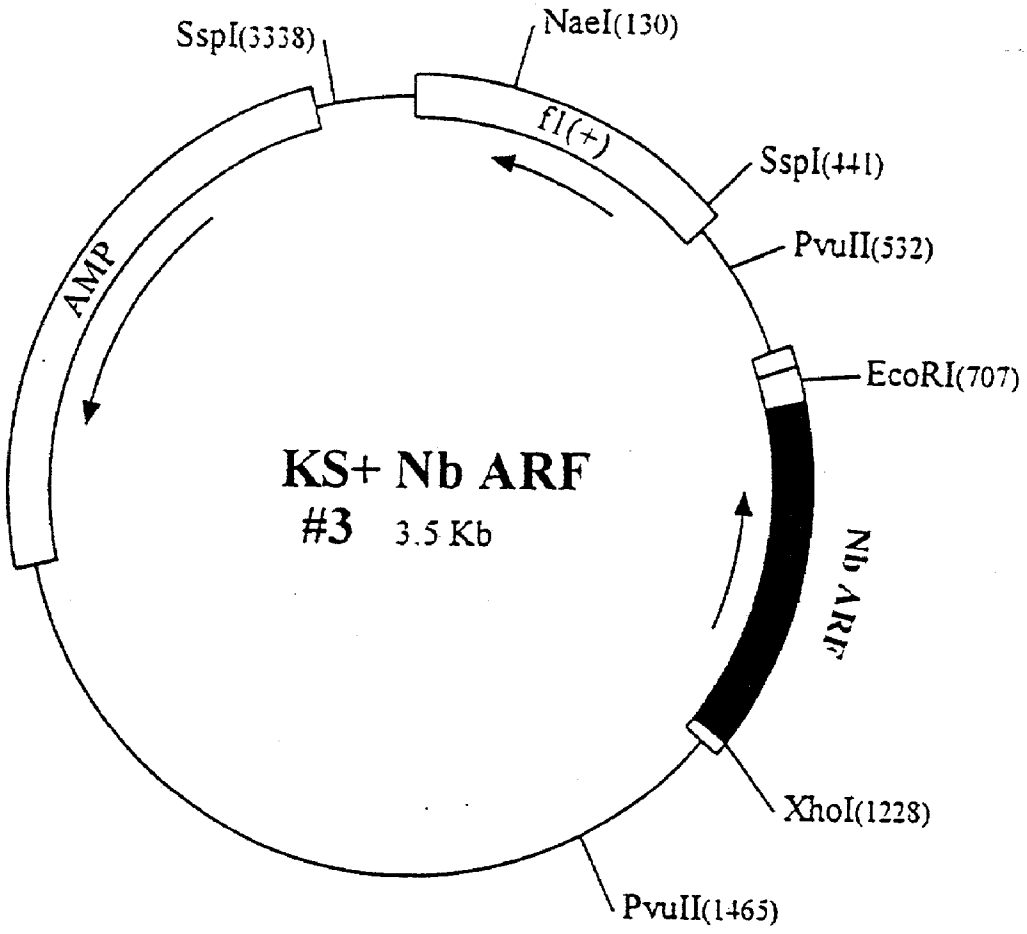
SspI(3338)
NaeI(130)
f1(+)
SspI(441)
PvuII(532)
EcoRI(707)
AMP
KS+ Nb ARF
#3 3.5 Kb
Nb ARF
XhoI(1228)
PvuII(1465)

Figure 9

Nucleotide sequence alignment of 740 AT #120 to KS+ Nb ARF #3

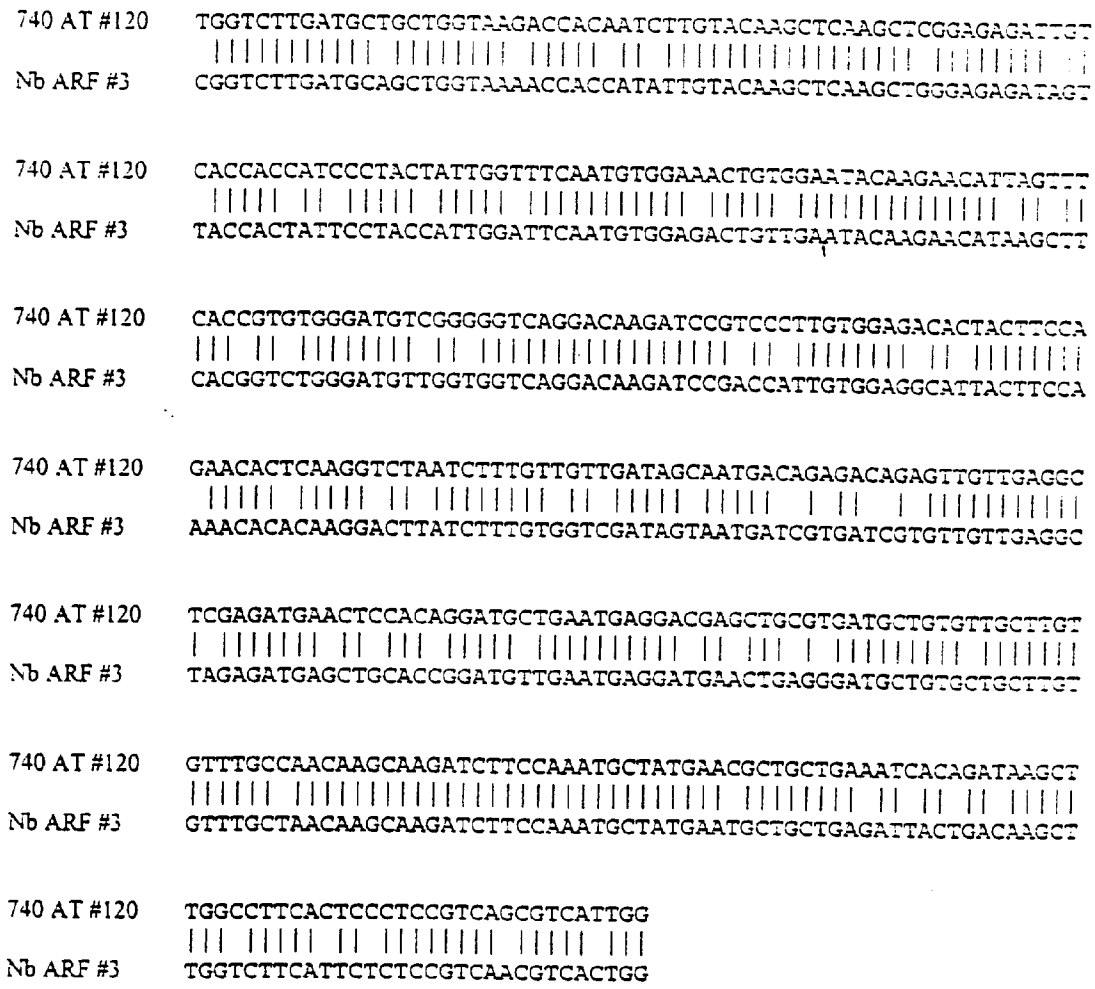

Figure 10

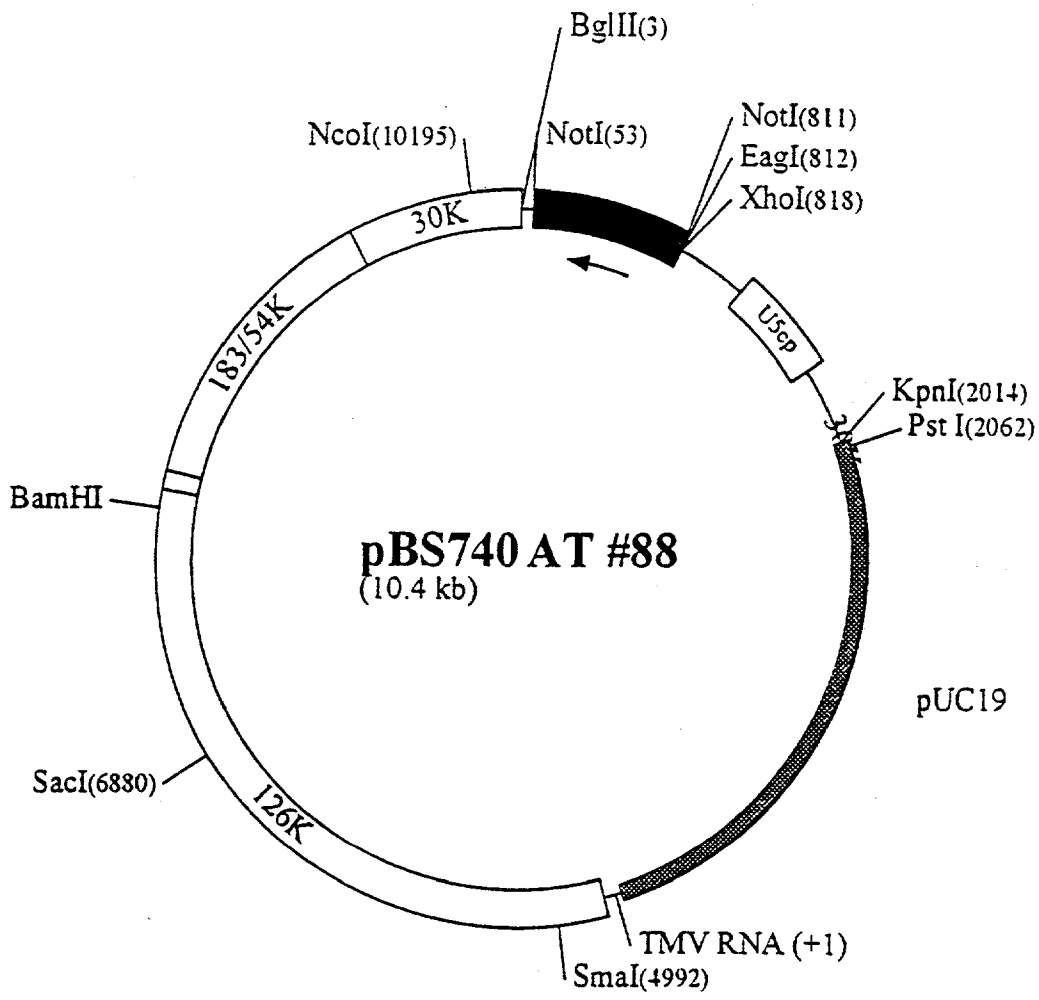
BglII(3)
NcoI(10195)
NotI(53)
NotI(811)
EagI(812)
XhoI(818)
30K
183/54K
U5cp
KpnI(2014)
Pst I(2062)
BamHI
pBS740 AT #88
(10.4 kb)
pUC19
SacI(6880)
126K
TMV RNA (+1)
SmaI(4992)

Figure 11

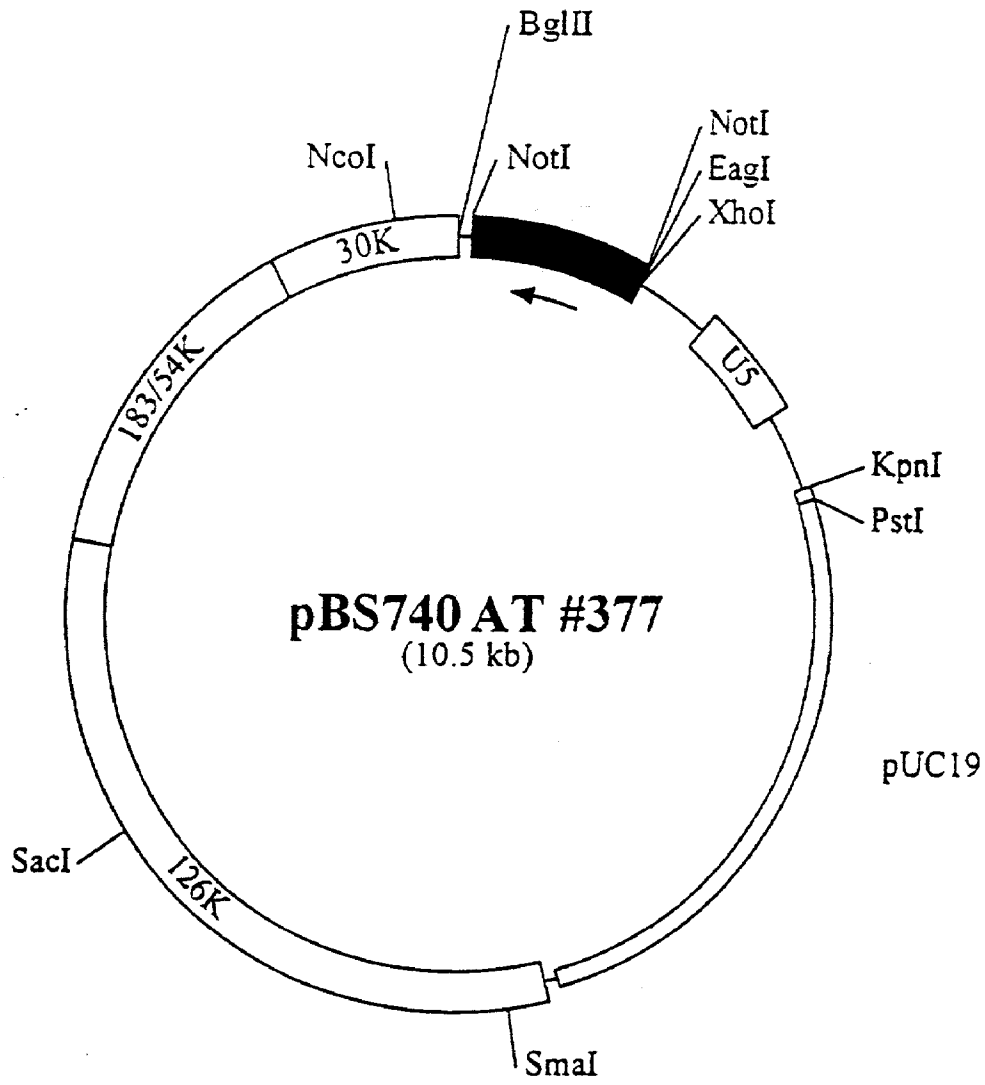
BglII
NotI
EagI
XhoI
NcoI
NotI
30K
183/54K
U5
KpnI
PstI
pBS740 AT #377
(10.5 kb)
pUC19
SacI
126K
SmaI

Figure 12

CTTAAAAGCAATATGACAGTAGAGAAGATCTCTCACAAAAGACCCAAAATCGAGTCGTGC
AAAATTGTACGAACAACAAAATTTAAAATTCAGTCCTTATCAAAGATCCAATCCAGCTGC
AACTAGCAACATTGGCTTAACGCTTCTTAGACACACCAACAGTCTTTCCTCTGCGACCAG
TTGTCTTGGTGTGTTGTCCACGAACACGGAGACCCCAGTAATGTCTCAGACCACGATGGT
TTCTGATTTTCTTGAGACGCTCAAGATCATCCCTGAGCTTCATGTCAAGGGCATTGGAGA
CAACTTGAGAGTACTTCCCATCCTTGTAATCTTTCTGTCTGTTCAAAAACCAGTCTGGAA
TCTTGAACTGTCTTGGGTTTGCAACAATAGTCATGAGGTTGTCAATCTCAGCTGCAGATA
ACTCACCAGCCCTCTTGTTCATGTCGACATCGGCTTTCTTGCAGACAATGTTGGCCAATC
TCCTTCCAATACCTTTGATAGAGGTAAGGGCAAACATAATCTTTTGCTTACCATCAACGG
TAG

Figure 13

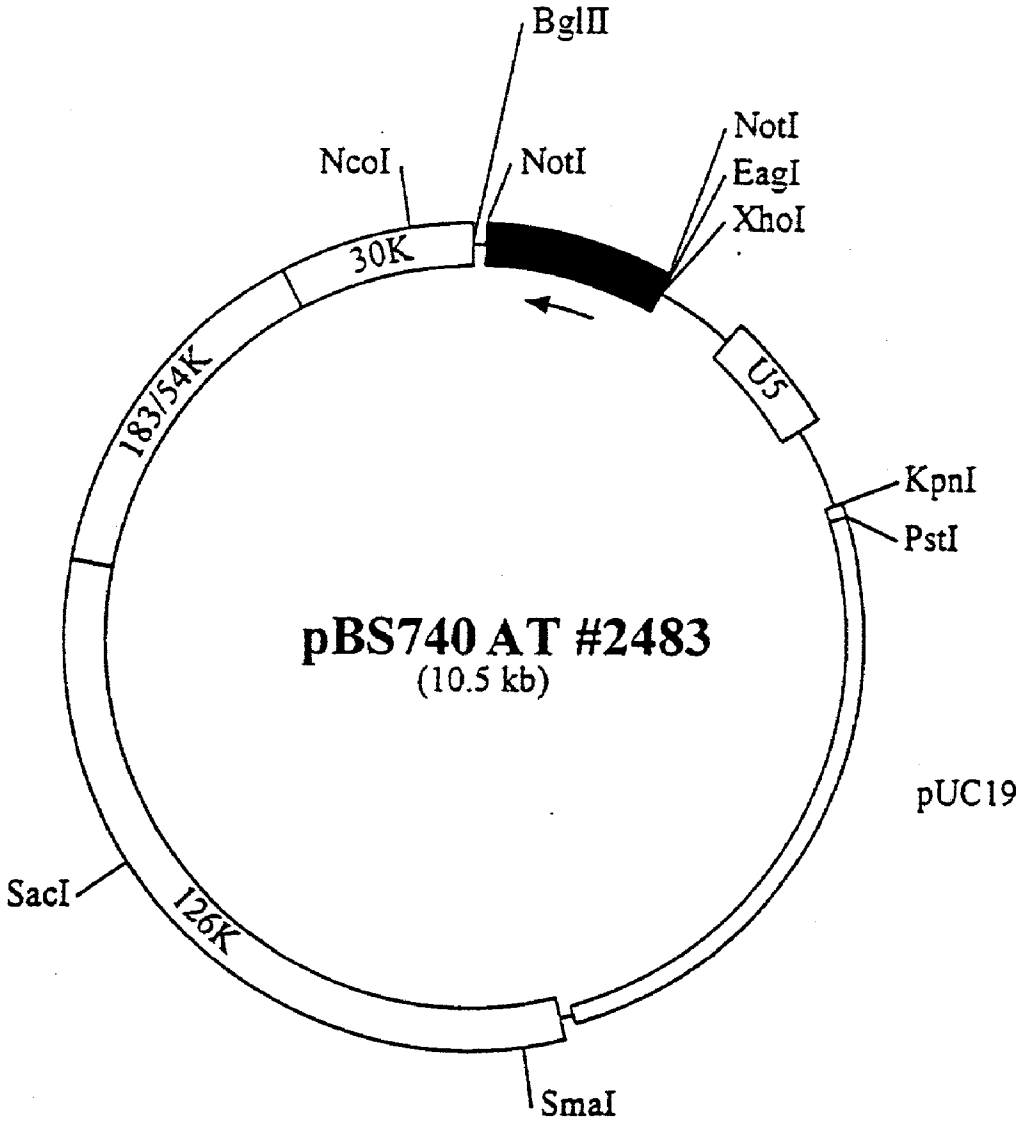
BglII
NotI
EagI
XhoI
NcoI
NotI
30K
183/54K
U5
KpnI
PstI
pBS740 AT #2483
(10.5 kb)
pUC19
SacI
126K
SmaI

Figure 14

CTGACATAAGTTATGTTCTTTGCGAAAATAAAAGTTATTCCACAAACGCATTCGATAAAA
CATTCAAAACCTTCTTCAGAGTCTAATCCGTGAACTGATGATCGATATAGCTTCACACTA
TATATCCTCTTCACTTCTTAGACTTCTTCTTCGGTACAGCTGCAGTTGGAGCAGGTGTAG
CAGCAGGTGCTGGAGCAGCTACAGGCGCAACATCTCCACCGGGACCCTTAGCTAAACGCT
CCTCTCTCCTAGCATGCTTCCTTTCTCGGCTAGCCTTGTTCTTCGCCCTCTTAGCCTCAA
ACTGATCAAGACAGAGTCTTCTCCCTAGCCTTCTCAAGCCTTTGACTTGTGGATACTCTC
CATCAAGACACGCTTGTTCTTGAACACATTTACCCTTAACACGCATGTACATGGTCATGG
TACATGTGCTTGTCAATCTTCTTTCGTCTCTCTGGATTTCTTCAAACAGACGCCTAAGAA
ACACGCCTTCCTACGCATTCCACAGTACCTTTGGTTGGGAACCTAACTTACGGGTACCCC
TTCCTTTTAACCGATTCCAGAGTGGCGACCCTTTATCTTGGCAATCTTCATTTTGCGAGC
TTGGAACAAGAGTGAATCTTGGGTGGCTTCTGATGATGAAACCTCTTTAAACTTTCTGAG
GTTTTGGCGGAAATGGCTGAAACGGATTTGTGGGACCAACCAAATTGCCTTTCGGCTTAA
TACTGATGCGACCGTTTGAGTAAAAACCGCCTTCAGG

Figure 15

METHOD OF IDENTIFYING A NUCLEIC ACID SEQUENCE IN A PLANT

[0001] This application is a divisional application of U.S. patent application Ser. No. 09/359,301, filed Jul. 21, 1999; which is a Continuation-In-Part application of U.S. patent application Ser. No. 09/232,170, filed on Jan. 15, 1999, which is a Continuation-In-Part application of U.S. patent application Ser. No. 09/008,186, filed on Jan. 16, 1998. All the above applications are incorporated herein by reference.

FIELD OF THE INVENTION

[0002] The present invention relates generally to the field of molecular biology and plant genetics. Specifically, the present invention relates to a method for determining the presence of a trait in a plant and a method of changing the phenotype or biochemistry of a plant, by expressing transiently a nucleic acid sequence in an antisense orientation in a host plant. This invention is exemplified by a nucleic acid sequence comprising a GTP binding protein open reading frame having an antisense orientation.

BACKGROUND OF THE INVENTION

[0003] Great interest exists in launching genome projects in plants comparable to the human genome project. Valuable and basic agricultural plants, including by way of example but without limitation, corn, soybeans and rice are targets for such projects because the information obtained thereby may prove very beneficial for increasing world food production and improving the quality and value of agricultural products. The United States Congress is considering launching a corn genome project. By helping to unravel the genetics hidden in the corn genome, the project could aid in understanding and combating common diseases of grain crops. It could also provide a big boost for efforts to engineer plants to improve grain yields and resist drought, pests, salt, and other extreme environmental conditions. Such advances are critical for a world population expected to double by 2050. Currently, there are four species which provide 60% of all human food: wheat, rice, corn, and potatoes, and the strategies for increasing the productivity of these plants is dependent on rapid discovery of the presence of a trait in these plants, and the function of unknown gene sequences in these plants.

[0004] One strategy that has been proposed to assist in such efforts is to create a database of expressed sequence tags (ESTs) that can be used to identify expressed genes. Accumulation and analysis of expressed sequence tags (ESTs) have become an important component of genome research. EST data may be used to identify gene products and thereby accelerate gene cloning. Various sequence databases have been established in an effort to store and relate the tremendous amount of sequence information being generated by the ongoing sequencing efforts. Some have suggested sequencing 500,000 ESTs for corn and 100,000 ESTs each for rice, wheat, oats, barley, and sorghum. Efforts at sequencing the genomes of plant species will undoubtedly rely upon these computer databases to share the sequence data as it is generated. *Arabidopsis thaliana* may be an attractive target discovery of a trait and for gene function discovery because a very large set of ESTs have already been produced in this organism, and these sequences tag more than 50% of the expected Arabidopsis genes.

[0005] Potential use of the sequence information so generated is enormous if gene function can be determined. It may become possible to engineer commercial seeds for agricultural use to convey any number of desirable traits to food and fiber crops and thereby increase agricultural production and the world food supply. Research and development of commercial seeds has so far focused primarily on traditional plant breeding, however there has been increased interest in biotechnology as it relates to plant characteristics. Knowledge of the genomes involved and the function of genes contained therein for both monocotyledonous and dicotyledonous plants is essential to realize positive effects from such technology.

[0006] The impact of genomic research in seeds is potentially far reaching. For example, gene profiling in cotton can lead to an understanding of the types of genes being expressed primarily in fiber cells. The genes or promoters derived from these genes may be important in genetic engineering of cotton fiber for increased strength or for "built-in" fiber color. In plant breeding, gene profiling coupled to physiological trait analysis can lead to the identification of predictive markers that will be increasingly important in marker assisted breeding programs. Mining the DNA sequence of a particular crop for genes important for yield, quality, health, appearance, color, taste, etc., are applications of obvious importance for crop improvement.

[0007] Work has been conducted in the area of developing suitable vectors for expressing foreign DNA and RNA in plant hosts. Ahlquist, U.S. Pat. Nos. 4,885,248 and 5,173,410 describes preliminary work done in devising transfer vectors which might be useful in transferring foreign genetic material into a plant host for the purpose of expression therein. All patent references cited herein are hereby incorporated by reference. Additional aspects of hybrid RNA viruses and RNA transformation vectors are described by Ahlquist et al. in U.S. Pat. Nos. 5,466,788, 5,602,242, 5,627,060 and 5,500,360, all of which are incorporated herein by reference. Donson et al., U.S. Pat. Nos. 5,316,931, 5,589,367 and 5,866,785, incorporated herein by reference, demonstrate for the first time plant viral vectors suitable for the systemic expression of foreign genetic material in plants. Donson et al. describe plant viral vectors having heterologous subgenomic promoters for the systemic expression of foreign genes. Carrington et al., U.S. Pat. No. 5,491,076, describe particular potyvirus vectors also useful for expressing foreign genes in plants. The expression vectors described by Carrington et al. are characterized by utilizing the unique ability of viral polyprotein proteases to cleave heterologous proteins from viral polyproteins. These include Potyviruses such as Tobacco Etch Virus. Additional suitable vectors are described in U.S. Pat. No. 5,811,653 and U.S. patent application Ser. No. 08/324,003, both of which are incorporated herein by reference.

[0008] Construction of plant RNA viruses for the introduction and expression of non-viral foreign genes in plants has also been demonstrated by Brisson et al., *Methods in Enzymology* 118:659 (1986), Guzman et al., Communications in Molecular Biology: Viral Vectors, Cold Spring Harbor Laboratory, pp. 172-189 (1988), Dawson et al., *Virology* 172:285-292 (1989), Takamatsu et al., *EMBO J.* 6:307-311 (1987), French et al., *Science* 231:1294-1297 (1986), and Takamatsu et al., *FEBS Letters* 269:73-76 (1990). However, these viral vectors have not been shown capable of systemic spread in the plant and expression of the non-viral foreign genes in the majority of plant cells in the

whole plant. Moreover, many of these viral vectors have not proven stable for the maintenance of non-viral foreign genes. However, the viral vectors described by Donson et al., in U.S. Pat. Nos. 5,316,931, 5,589,367, and 5,866,785, Turpen in U.S. Pat. No. 5,811,653, Carrington et al. in U.S. Pat. No. 5,491,076, and in co-pending U.S. patent application Ser. No. 08/324,003, have proven capable of infecting plant cells with foreign genetic material and systemically spreading in the plant and expressing the non-viral foreign genes contained therein in plant cells locally or systemically. All patents, patent applications, and references cited in the instant application are hereby incorporated by reference.

[0009] With the recent advent of technology for cloning, genes can be selectively turned off. One method is to create antisense RNA or DNA molecules that bind specifically with a targeted gene's RNA message, thereby interrupting the precise molecular mechanism that expresses a gene as a protein. The antisense technology which deactivates specific genes provides a different approach from a classical genetics approach. Classical genetics usually studies the random mutations of all genes in an organism and selects the mutations responsible for specific characteristics. Antisense approach starts with a cloned gene of interest and manipulates it to elicit information about its function.

[0010] Post-transcriptional gene silencing (PTGS) in transgenic plants is the manifestation of a mechanism that suppresses RNA accumulation in a sequence-specific manner. There are three models to account for the mechanism of PTGS: direct transcription of an antisense RNA from the transgene, an antisense RNA produced in response to over expression of the transgene, or an antisense RNA produced in response to the production of an aberrant sense RNA product of the transgene (Baulcombe, *Plant Mol. Biol.* 32:79-88 (1996)). The PTGS mechanism is typified by the highly specific degradation of both the transgene mRNA and the target RNA, which contains either the same or complementary nucleotide sequences (Waterhouse et al *Proc. Natl. Acad. Sci.* USA 10: 13959-64 (1998)). Antisense RNA has been used to down regulate gene expression in transgenic and transfected plants. The effectiveness of antisense on the inhibition of eukaryotic gene expression was first demonstrated by Izant et al. (*Cell* 36(4):1007-1015 (1984)). Since then, the down-regulation of different genes from transgenic plants has been reported. Kumagai et al (*Proc. Natl. Acad. Sci.* USA 92:1679 (1995)) report that gene expression in transfected *Nicotiana benthamiana* was cytoplasmic inhibited by viral delivery of a RNA of a known sequence derived from cDNA encoding tomato phytoene desaturase in a positive sense or an antisense orientation. The host plant, *Nicotiana benthamiana,* and the donor plant, tomato (*Lycopersicon esculentum*), belong to the same family. There is also evidence that inhibition of endogenous genes occurs in transgenic plants containing sense RNA (Van der Krol et al., *Plant Cell* 2(4):291-299 (1990), Napoli et al., *Plant Cell* 2:279-289 (1990) and Fray et al., *Plant Mol. Biol.* 22:589-602 (1993)).

[0011] The antisense technology can be used to develop a functional genomic screening of a plant of interest. The antisense technology is applied in this invention to provide a method of discovering the presence of a trait in a plant and to determine the function and sequence of a nucleic acid of a plant by expressing the nucleic acid sequence that has not been identified in a different host plant. GTP-binding proteins exemplify this invention. In eukaryotic cells, GTP-binding proteins function in a variety of cellular processes, including signal transduction, cytoskeletal organization, and protein transport. The heterotrimeric and monomeric GTP-binding proteins that may be involved in secretion and intracellular transport are divided into two structural classes: the rab and the ARF families. The ARFs are highly conserved and found in all eukaryotic cells including human, yeast, plants, and slime mold. The cDNAs encoding GTP binding proteins have been isolated from a variety of plants including rice, barley, corn, tobacco, and *A. thaliana.* For example, Verwoert et al. (*Plant Molecular Biol.* 27:629-633 (1995)) report the isolation of a *Zea mays* cDNA clone encoding a GTP-binding protein of the ARF family by direct genetic selection in an *E. coli* fabD mutant with a maize cDNA expression library. Regad et al. (*FEBS* 2:133-136 (1993)) isolated a cDNA clone encoding the ARF from a cDNA library of *Arabidopsis thaliana* cultured cells by randomly selecting and sequencing cDNA clones. Dallmain et al. (*Plant Molecular Biol.* 19:847-857 (1992)) isolated two cDNAs encoding small GTP-binding proteins from leaf cDNA libraries using a PCR approach. Dallmann et al. prepared leaf cDNAs and use them as templates in PCR amplifications with degenerated oligonucleotides corresponding to the highly conserved motifs, found in members of the ras superfamily, as primers. The present invention provides advantages over the above isolation methods in that it only sequences clones that have a function and does not randomly sequence clones. Haizel et al., *Plant J.,* 11:92-103 (1997)) isolated cDNA and genomic clones encoding Ran-like small GTP binding proteins from Arabidopsis cDNA and genomic libraries using a full-length tobacco Nt Ran1 cDNA as a probe.

[0012] The present invention provides a method for discovering the presence of a trait in a plant by expressing a nucleic acid sequence in an antisense orientation in a host plant. Once the presence of a trait is identified by phenotypic changes, the nucleic acid insert in the cDNA clone or in the vector is then sequenced. The present method provides a rapid method for determining the presence of a trait and identifying a nucleic acid sequence and its function in a plant by screening a transfected host plant for its change of function.

SUMMARY OF THE INVENTION

[0013] The present invention is directed to a method of making a flnctional gene profile in a host plant by transiently expressing a nucleic acid sequence library in a host plant, determining the phenotypic or biochemical changes in the host plant, identifying a trait associated with the change, identifying the donor gene associated with the trait, and identifying the homologous host gene, if any. The invention is also directed to a method of changing the phenotype or biochemistry of a plant, a method of determining a change in phenotype or biochemistry in a plant, and a method of determining the presence of a trait in a plant. The method comprising the steps of expressing transiently an unidentified RNA or DNA sequence of a donor plant in an antisense orientation in a host plant, identifying changes in the host plant, and correlating the expression with the phenotypic changes. Alternatively, the method comprises the steps of expressing transiently a nucleic acid sequence of a donor plant in an antisense orientation in a host plant, identifying changes in the host plant, and correlating the expression with

the phenotypic changes, wherein the donor plant and the host plant belong to different genus, family, order, class, subdivision, or division. The present invention is also directed to a method of determining the function of a nucleic acid sequence, including a gene, in a donor plant, by transfecting the nucleic acid sequence into a host plant in a manner so as to affect phenotypic changes in the host plant. In one embodiment, recombinant viral nucleic acids are prepared to include the nucleic acid insert of a donor. The recombinant viral nucleic acids infect a host plant and produce antisense RNAs in the cytoplasm which result in reduced expression of endogenous cellular genes in the host plant. Once the presence of a trait is identified by phenotypic changes, the function of the nucleic acid is determined. The nucleic acid insert in a cDNA clone or in a vector is then sequenced. The nucleic acid sequence is determined by a standard sequence analysis. This invention is exemplified by a nucleic acid sequence comprising a GTP binding protein open reading frame having an antisense orientation.

[0014] The present invention is also directed to a method of increasing yield of a grain crop. The method comprises expressing transiently a nucleic acid sequence of a donor plant in an antisense orientation in a grain crop, wherein said expressing results in stunted growth and increased seed production of the grain crop. A preferred method comprises the steps of cloning the nucleic acid sequence into a plant viral vector and infecting the grain crop with a recombinant viral nucleic acid comprising said nucleic acid sequence.

[0015] One aspect of the invention is a method of identifying and determining a nucleic acid sequence in a plant of interest, whose function is to silence endogenous genes in a host plant, by introducing the nucleic acid into the host plant by way of a viral nucleic acid such as a plant viral nucleic acid suitable to produce expression of the nucleic acid in the transfected host. This method utilizes the principle of post-transcription gene silencing of the endogenous host gene homolog, for example, antisense RNAs. Particularly, this silencing function is useful for silencing a multigene family frequently found in plants.

[0016] Another aspect of the invention is to discover genes having the same function in different plants. The method starts with a library of cDNAs, genomic DNAs, or a pool of RNAs of a first plant. Then, a recombinant viral nucleic acid comprising a nucleic acid insert derived from the library is prepared and is used to infect a different host plant. The infected host plant is inspected for phenotypic changes. The recombinant viral nucleic acid that results in phenotypic changes in the host plant is identified and the sequence of the nucleic acid insert is determined by a standard method. Such nucleic acid sequence in the first plant has substantial sequence homology as that in the host plant: the nucleic acid sequences are conserved between the two plants. This invention provides a rapid means for elucidating the function and sequence of nucleic acids of interest; such rapidly expanding information can be subsequently utilized in the field of genomics.

[0017] In one embodiment, a nucleic acid is introduced into a plant host wherein the plant host may be a monocotyledonous or dicotyledonous plant, plant tissue or plant cell. Preferably, the nucleic acid is introduced by way of a recombinant plant viral nucleic acid. Preferred recombinant plant viral nucleic acids useful in the present invention comprise a native plant viral subgenomic promoter, a plant viral coat protein coding sequence, and at least one non-native nucleic acid sequence. Some viral vectors used in accordance with the present invention may be encapsidated by the coat proteins encoded by the recombinant plant virus. Recombinant plant viral nucleic acids or recombinant plant viruses are used to infect a plant host. The recombinant plant viral nucleic acid is capable of replication in the host, localized or systemic spread in the host, and transcription or expression of the non-native nucleic acid in the host to produce a phenotypic or biochemical change. Any suitable vector constructs useful to produce localized or systemic expression of nucleic acids in host plants are within the scope of the present invention.

BRIEF DESCRIPTION OF THE FIGURES

[0018] **FIG. 1** depicts the vector TT01/PDS−.

[0019] **FIG. 2** depicts the plasmid pBS #712.

[0020] **FIG. 3** depicts the plasmid KS+TVCVK #23.

[0021] **FIG. 4** depicts the plasmid pBS #735.

[0022] **FIG. 5** depicts the plasmid pBS #740.

[0023] **FIG. 6** depicts the plasmid pBS #740 AT #120 (ATCC No.: PTA-325, deposited Jul. 12, 1999, American Type Culture Collection, 10801 University Blvd., Manassas, Va. 20110).

[0024] **FIG. 7** shows the nucleotide sequence comparison of *A. thaliana* 740 AT #120 and *A. thaliana* est AA042085

[0025] **FIG. 8** shows the nucleotide sequence alignment of 740 AT #120 to rice D1 7760 (*Oryza sativa*) ADP-ribosylation factor.

[0026] **FIG. 9** shows the KS+Nb ARF #3 plasmid (ATCC No.: PTA-324, deposited Jul. 12, 1999, American Type Culture Collection, 10801 University Blvd., Manassas, Va. 20110) map.

[0027] **FIG. 10** shows the nucleotide sequence comparison of *A. thaliana* 740 AT #120 and *N. benthamiana* KS+Nb ARF#3.

[0028] **FIG. 11** shows the plasmid pBS #740 AT #88 (ATCC No.: PTA-33 1, deposited Jul. 12, 1999, American Type Culture Collection, 10801 University Blvd., Manassas, Va. 20110).

[0029] **FIG. 12** shows the plasmid pBS #377 (ATCC No.: PTA-334, deposited Jul. 12, 1999, American Type Culture Collection, 10801 University Blvd., Manassas, Va. 20110).

[0030] **FIG. 13** shows the nucleotide sequence of 740 AT #377.

[0031] **FIG. 14** shows the plasmid pBS #2483 (ATCC No.: PTA-328, deposited Jul. 12, 1999, American Type Culture Collection, 10801 University Blvd., Manassas, Va. 20110).

[0032] **FIG. 15** shows the nucleotide sequence of 740 AT #2483.

DETAILED DESCRIPTION OF THE INVENTION

[0033] The present invention is directed to a method of changing the phenotype or biochemistry of a plant, a method

of determining a change in phenotype or biochemistry in a plant, a method of determining the presence of a trait in a plant, and a method of determining the function of a nucleic acid sequence. The methods comprise the steps of expressing transiently a nucleic acid sequence in an antisense orientation in a host organism such as a plant, a plant tissue or a plant cell, identifying changes in the host organism and correlating the expression and the changes. The presence of a trait is identified either in the infected host plant or in an uninfected host plant. The nucleic acid sequence, which is expressed in the host plant, does not need to be identified, isolated, or characterized prior to the expression. The donor plant and the host plant can belong to different genus, family, order, class, subdivision, or division. In one preferred embodiment, the method comprising the steps of (a) preparing a library of cDNAs, genomic DNAs, or a pool of RNAs of a donor plant, (b) constructing recombinant viral nucleic acids comprising a nucleic acid insert derived from said library, (c) infecting each said host plant with one of said recombinant viral nucleic acids, (d) growing said infected host plant, and (e) determining changes in said host plant.

[**0034**] The invention is directed to a method of compiling a plant antisense functional gene profile. The method comprises (a) preparing a vector library of DNA or RNA sequences from a donor plant, each sequence being in an antisense orientation; (b) infecting a plant host with a vector; (c) transiently expressing the donor plant DNA or RNA sequence in the growing plant host; (d) determining one or more phenotypic or biochemical changes in the plant host, if any; (e) identifying an associated trait where a phenotypic or biochemical change occurs; (f) identifying a donor plant gene associated with the trait; (g) identifying a plant host gene, if any, associated with the trait; and (h) repeating steps (b)-(g) until an antisense functional gene profile of the plant host and/or of the donor plant is compiled.

[**0035**] The invention is also directed to a method of compiling a plant functional gene profile. The method comprises (a) preparing a vector library of DNA or RNA sequences from a donor plant, each sequence being in either an antisense or a positive orientation; (b) infecting a plant host with a vector; (c) transiently expressing the donor plant DNA or RNA sequence in the growing plant host; (d) determining one or more phenotypic or biochemical changes in the plant host, if any; (e) identifying an associated trait where a phenotypic or biochemical change occurs; (f) identifying a donor plant gene associated with the trait; (g) identifying a plant host gene, if any, associated with the trait; and (h) repeating steps (b)-(g) until a functional gene profile of the plant host and/or of the donor plant is compiled. A detailed discussion of positive sense expression of nucleic acids is presented in a co-pending and co-owned U.S. patent application Ser. No. 09/359,305 (Kumagai et al., Attorney Docket No. 08010137US05, filed Jul. 1, 1999), the entire disclosure of which is incorporated herein by reference.

[**0036**] The present method has the advantages that the nucleic acid sequence does not need to be known, identified, isolated, or characterized prior to infecting a host plant with a recombinant viral nucleic acid comprising the nucleic acid sequence. Once changes in the host plant is observed, the nucleic acid sequence can be determined by further identifying the recombinant viral nucleic acid that results in changes in the host, and analyzing the sequence of the nucleic acid insert in the recombinant viral nucleic acid that results in changes in the host.

[**0037**] The present invention provides a method of infecting a host plant by a recombinant plant viral nucleic acid which contains one or more non-native nucleic acid sequences, or by a recombinant plant virus containing a recombinant plant viral nucleic acid. The non-native nucleic acids are subsequently transcribed or expressed in the infected host plant. The products of the non-native nucleic acid sequences result in changing phenotypic traits in the host plant, affecting biochemical pathways within the plant, or affecting endogenous gene expression within the plant.

[**0038**] In one embodiment, a nucleic acid is introduced into a plant host by way of a recombinant viral nucleic acid. Such recombinant viral nucleic acids are stable for the maintenance and transcription or expression of non-native nucleic acid sequences and are capable of systemically transcribing or expressing such non-native sequences in the plant host. Preferred recombinant plant viral nucleic acids useful in the present invention comprise a native plant viral subgenomic promoter, a plant viral coat protein coding sequence, and at least one non-native nucleic acid sequence.

[**0039**] In a second embodiment, plant viral nucleic acid sequences are characterized by the deletion of a native coat protein coding sequence. The plant viral nucleic acid sequence comprises a non-native plant viral coat protein coding sequence and a non-native promoter, preferably the subgenomic promoter of the non-native coat protein coding sequence. Such plant viral nucleic acid sequence is capable of expressing in a plant host, packaging the recombinant plant viral nucleic acid, and ensuring a systemic infection of the host by the recombinant plant viral nucleic acid. The recombinant plant viral nucleic acid may contain one or more additional native or non-native subgenomic promoters. Each non-native subgenomic promoter is capable of transcribing or expressing adjacent genes or nucleic acid sequences in the plant host and incapable of recombination with each other and with native subgenomic promoters. One or more non-native nucleic acids maybe inserted adjacent to the native plant viral subgenomic promoter or the native and non-native plant viral subgenomic promoters if more than one nucleic acid sequence is included. Moreover, two or more heterologous non-native subgenomic promoters may be used. The non-native nucleic acid sequences may be transcribed or expressed in the host plant under the control of the subgenomic promoter to produce the products of the nucleic acids of interest.

[**0040**] In a third embodiment, plant recombinant viral nucleic acids comprise a native coat protein coding sequence instead of a non-native coat protein coding sequence, placed adjacent one of the non-native coat protein subgenomic promoters.

[**0041**] In a fourth embodiment, plant recombinant viral nucleic acids comprise a native coat protein gene adjacent its native subgenomic promoter, one or more non-native subgenomic promoters, and at least one non-native nucleic acid sequence. The native plant viral subgenomic promoter initiates transcription of the plant viral coat protein sequence. The non-native subgenomic promoters are capable of transcribing or expressing adjacent genes in a plant host and are incapable of recombination with each other and with native subgenomic promoters. Non-native nucleic acid sequences

may be inserted adjacent the non-native subgenomic plant viral promoters such that the sequences are transcribed or expressed in the host plant under control of the subgenomic promoters to produce a product of the non-native nucleic acid. Alternatively, the native coat protein coding sequence may be replaced by a non-native coat protein coding sequence.

[0042] The viral vectors used in accordance with the present invention may be encapsidated by the coat proteins encoded by the recombinant plant virus. The recombinant plant viral nucleic acid or recombinant plant virus is used to infect a host plant. The recombinant plant viral nucleic acid is capable of replication in the host, localized or systemic spread in the host, and transcription or expression of the non-native nucleic acid in the host to affect a phenotypic or biochemical change in the host.

[0043] In one embodiment, recombinant plant viruses are used which encode for the expression of a fusion between a plant viral coat protein and the amino acid product of the nucleic acid of interest. Such a recombinant plant virus provides for high level expression of a nucleic acid of interest. The location or locations where the viral coat protein is joined to the amino acid product of the nucleic acid of interest may be referred to as the fusion joint. A given product of such a construct may have one or more fusion joints. The fusion joint may be located at the carboxyl terminus of the viral coat protein or the fusion joint may be located at the amino terminus of the coat protein portion of the construct. In instances where the nucleic acid of interest is located internal with respect to the 5' and 3' residues of the nucleic acid sequence encoding for the viral coat protein, there are two fusion joints. That is, the nucleic acid of interest may be located 5', 3', upstream, downstream or within the coat protein. In some embodiments of such recombinant plant viruses, a "leaky" start or stop codon may occur at a fusion joint which sometimes does not result in translational termination. A more detailed description of some recombinant plant viruses according to this embodiment of the invention may be found in co-pending U.S. patent application Ser. No. 08/324,003 the disclosure of which is incorporated herein by reference.

[0044] The present invention is not intended to be limited to any particular viral constructs, but rather to include all operable constructs. Specifically, those skilled in the art may choose to transfer DNA or RNA of any size up to and including an entire genome in a plant into a host organism in order to determine the presence of a trait in the plant. Those skilled in the art will understand that the recited embodiments are representative only. All operable constructs useful to produce localized or systemic expression of nucleic acids in a plant are within the scope of the present invention.

[0045] The chimeric genes and vectors and recombinant plant viral nucleic acids used in this invention are constructed using techniques well known in the art. Suitable techniques have been described in Sambrook et al. (2nd ed.), Cold Spring Harbor Laboratory, Cold Spring Harbor (1982, 1989); *Methods in Enzymol.* (Vols. 68, 100, 101, 118, and 152-155) (1979, 1983, 1986 and 1987); and *DNA Cloning,* D. M. Clover, Ed., IRL Press, Oxford (1985). Medium compositions have been described by Miller, J., *Experiments in Molecular Genetics,* Cold Spring Harbor Laboratory, New York (1972), as well as the references previously identified, all of which are incorporated herein by reference. DNA manipulations and enzyme treatments are carried out in accordance with manufacturers' recommended procedures in making such constructs.

[0046] The first step in producing recombinant plant viral nucleic acids is to modify the nucleotide sequences of the plant viral nucleotide sequence by known conventional techniques such that one or more non-native subgenomic promoters are inserted into the plant viral nucleic acid without destroying the biological function of the plant viral nucleic acid. The subgenomic promoters are capable of transcribing or expressing adjacent nucleic acid sequences in a plant host infected by the recombination plant viral nucleic acid or recombinant plant virus. The native coat protein coding sequence may be deleted in some embodiments, placed under the control of a non-native subgenomic promoter in other embodiments, or retained in a further embodiment. If it is deleted or otherwise inactivated, a non-native coat protein gene is inserted under control of one of the non-native subgenomic promoters, or optionally under control of the native coat protein gene subgenomic promoter. The non-native coat protein is capable of encapsidating the recombinant plant viral nucleic acid to produce a recombinant plant virus. Thus, the recombinant plant viral nucleic acid contains a coat protein coding sequence, which may be native or a nonnative coat protein coding sequence, under control of one of the native or non-native subgenomic promoters. The coat protein is involved in the systemic infection of the plant host.

[0047] Viruses suitable for use according to the methods of the present invention include viruses from the tobamovirus group such as Tobacco Mosaic virus (TMV), Ribgrass Mosaic Virus (RGM), Cowpea Mosaic virus (CMV), Alfalfa Mosaic virus (AMV), Cucumber Green Mottle Mosaic virus watermelon strain (CGMMV-W) and Oat Mosaic virus (OMV) and viruses from the brome mosaic virus group such as Brome Mosaic virus (BMV), broad bean mottle virus and cowpea chlorotic mottle virus. Additional suitable viruses include Rice Necrosis virus (RNV), and geminiviruses such as Tomato Golden Mosaic virus (TGMV), Cassava Latent virus (CLV) and Maize Streak virus (MSV). Each of these groups of suitable viruses is characterized below. However, the invention should not be construed as limited to using these particular viruses, but rather the present invention is contemplated to include all plant viruses at a minimum.

TOBAMOVIRUS GROUP

[0048] The tobacco mosaic virus (TMV) is of particular interest to the instant invention because of its ability to express genes at high levels in plants. TMV is a member of the tobamovirus group. The TMV virion is a tubular filament, and comprises coat protein sub-units arranged in a single right-handed helix with the single-stranded RNA intercalated between the turns of the helix. TMV infects tobacco as well as other plants. TMV virions are 300 nm×18 nm tubes with a 4 nm-diameter hollow canal, and consist of 2140 units of a single structural protein helically wound around a single RNA molecule. The genome is a 6395 base plus-sense RNA. The 5'-end is capped and the 3'-end contains a series of pseudoknots and a tRNA-like structure that will specifically accept histidine. The genomic RNA functions as mRNA for the production of proteins involved in

viral replication: a 126-kDa protein that initiates 68 nucleotides from the 5'-terminus, and a 183-kDa protein synthesized by readthrough of an amber termination codon approximately 10% of the time. Only the 183-kDa and 126-kDa viral proteins are required for the TMV replication in trans. (Ogawa et al., Virology 185:580-584 (1991)). Additional proteins are translated from subgenomic size mRNA produced during replication (Dawson, *Adv. Virus Res.,* 38:307-342 (1990)). The 30-kDa protein is required for cell-to-cell movement; the 17.5-kDa capsid protein is the single viral structural protein. The function of the predicted 54-kDa protein is unknown.

[**0049**] TMV assembly apparently occurs in plant cell cytoplasm, although it has been suggested that some TMV assembly may occur in chloroplasts since transcripts of ctDNA have been detected in purified TMV virions. Initiation of TMV assembly occurs by interaction between ring-shaped aggregates ("discs") of coat protein (each disc consisting of two layers of 17 subunits) and a unique internal nucleation site in the RNA; a hairpin region about 900 nucleotides from the 3'-end in the common strain of TMV. Any RNA, including subgenomic RNAs containing this site, may be packaged into virions. The discs apparently assume a helical form on interaction with the RNA, and assembly (elongation) then proceeds in both directions (but much more rapidly in the 3'- to 5'- direction from the nucleation site).

[**0050**] Another member of the Tobamoviruses, the Cucumber Green Mottle Mosaic virus watermelon strain (CGMMV-W) is related to the cucumber virus. Nozu et al., *Virology* 45:577 (1971). The coat protein of CGMMV-W interacts with RNA of both TMV and CGMMV to assemble viral particles in vitro. Kurisu et al., *Virology* 70:214 (1976).

[**0051**] Several strains of the tobamovirus group are divided into two subgroups, on the basis of the location of the assembly of origin. Subgroup I, which includes the vulgare, OM, and tomato strain, has an origin of assembly about 800-1000 nucleotides from the 3'-end of the RNA genome, and outside the coat protein cistron. Lebeurier et al., *Proc. Natl. Acad. Sci.* USA 74:149 (1977); and Fukuda et al., *Virology* 101:493 (1980). Subgroup II, which includes CGMMV-W and cornpea strain (Cc) has an origin of assembly about 300-500 nucleotides from the 3'-end of the RNA genome and within the coat-protein cistron. The coat protein cistron of CGMMV-W is located at nucleotides 176-661 from the 3'-end. The 3' noncoding region is 175 nucleotides long. The origin of assembly is positioned within the coat protein cistron. Meshi et al, *Virology* 127:54 (1983).

BROME MOSAIC VIRUS GROUP

[**0052**] Brome Mosaic virus (BMV) is a member of a group of tripartite, single-stranded, RNA-containing plant viruses commonly referred to as the bromoviruses. Each member of the bromoviruses infects a narrow range of plants. Mechanical transmission of bromoviruses occurs readily, and some members are transmitted by beetles. In addition to BV, other bromoviruses include broad bean mottle virus and cowpea chlorotic mottle virus.

[**0053**] Typically, a bromovirus virion is icosahedral, with a diameter of about 26 μm, containing a single species of coat protein. The bromovirus genome has three molecules of linear, positive-sense, single-stranded RNA, and the coat protein mRNA is also encapsidating. The RNAs each have a capped 5'-end, and a tRNA-like structure (which accepts tyrosine) at the 3'-end. Virus assembly occurs in the cytoplasm. The complete nucleotide sequence of BMV has been identified and characterized as described by Ahlquist et al., *J. Mol. Biol.* 153:23 (1981).

RICE NECROSIS VIRUS

[**0054**] Rice Necrosis virus is a member of the Potato Virus Y Group or Potyviruses. The Rice Necrosis virion is a flexuous filament comprising one type of coat protein (molecular weight about 32,000 to about 36,000) and one molecule of linear positive-sense single-stranded RNA. The Rice Necrosis virus is transmitted by *Polymyxa oraminis* (a eukaryotic intracellular parasite found in plants, algae and fungi).

GEMINIVIRUSES

[**0055**] Geminiviruses are a group of small, single-stranded DNA-containing plant viruses with virions of unique morphology. Each virion consists of a pair of isometric particles (incomplete icosahedral), composed of a single type of protein (with a molecular weight of about $2.7\text{-}3.4\times10^4$). Each geminivirus virion contains one molecule of circular, positive-sense, single-stranded DNA. In some geminiviruses (i.e., Cassava latent virus and bean golden mosaic virus) the genome appears to be bipartite, containing two single-stranded DNA molecules.

POTYVIRUSES

[**0056**] Potyviruses are a group of plant viruses which produce polyprotein. A particularly preferred potyvirus is tobacco etch virus (TEV). TEV is a well characterized potyvirus and contains a positive-strand RNA genome of 9.5 kilobases encoding for a single, large polyprotein that is processed by three virus-specific proteinases. The nuclear inclusion protein "a" proteinase is involved in the maturation of several replication-associated proteins and capsid protein. The helper component-proteinase (HC-Pro) and 35-kDa proteinase both catalyze cleavage only at their respective C-termini. The proteolytic domain in each of these proteins is located near the C-terminus. The 35-kDa proteinase and HC-Pro derive from the N-terminal region of the TEV polyprotein.

[**0057**] The nucleic acid of any suitable plant virus can be utilized to prepare a recombinant plant viral nucleic acid for use in the present invention, and the foregoing are only exemplary of such suitable plant viruses. The nucleotide sequence of the plant virus can be modified, using conventional techniques, by insertion of one or more subgenomic promoters into the plant viral nucleic acid. The subgenomic promoters are capable of functioning in a specific host plant. For example, if the host is a tobacco plant, TMV, TEV, or other viruses containing suitable subgenomic promoter may be utilized. The inserted subgenomic promoters should be compatible with the viral nucleic acid and capable of directing transcription or expression of adjacent nucleic acid sequences in tobacco.

[**0058**] The native or non-native coat protein gene is included in the recombinant plant viral nucleic acid. When non-native nucleic acid is utilized, it may be positioned adjacent its natural subgenomic promoter or adjacent one of

the other available subgenomic promoters. The non-native coat protein, as is the case for the native coat protein, is capable of encapsidating the recombinant plant viral nucleic acid and providing for systemic spread of the recombinant plant viral nucleic acid in a host plant. The coat protein is selected to provide a systemic infection in the plant host of interest. For example, the TMV-O coat protein provides systemic infection in *N. benthamiana,* whereas TMV-U1 coat protein provides systemic infection in *N. tabacum.*

[**0059**] The recombinant plant viral nucleic acid is prepared by cloning a viral nucleic acid. If the viral nucleic acid is DNA, it can be cloned directly into a suitable vector using conventional techniques. One technique is to attach an origin of replication to the viral DNA which is compatible with the cell to be transfected. If the viral nucleic acid is RNA, a full-length DNA copy of the viral genome is first prepared by well-known procedures. For example, the viral RNA is transcribed into DNA using reverse transcriptase to produce subgenomic DNA pieces, and a double-stranded DNA made using DNA polymerases. The cDNA is then cloned into appropriate vectors and cloned into a cell to be transfected. Alternatively, the cDNA is ligated into the vector and is directly transcribed into infectious RNA in vitro, the infectious RNA is then inoculated onto the plant host. The cDNA pieces are mapped and combined in a proper sequence to produce a full-length DNA copy of the viral RNA genome, if necessary. DNA sequences for the subgenomic promoters, with or without a coat protein gene, are then inserted into the nucleic acid at non-essential sites, according to the particular embodiment of the invention utilized. Non-essential sites are those that do not affect the biological properties of the plant viral nucleic acids. Since the RNA genome is the infective agent, the cDNA is positioned adjacent a suitable promoter so that the RNA is produced in the production cell. The RNA can be capped by the addition of a nucleotide in a 5'-5' linkage using conventional techniques (Dawson et al., *Proc. Natl. Acad. Sci.* USA, 83:1832 (1986). One or more nucleotides may be added between the transcription start site of the promoter and the start of the cDNA of a viral nucleic acid to construct an infectious viral vector. In a preferred embodiment of the present invention, the inserted nucleotide sequence contains a G at the 5'-end. In one embodiment, the inserted nucleotide sequence is GNN, GTN, or their multiples, $(GNN)_x$ or $(GTN)_x$. The capped RNA can be packaged in vitro with added coat protein from TMV to make assembled virions. These assembled virions can then be used to inoculate plants or plant tissues.

[**0060**] Alternatively, an uncapped RNA may be employed in the embodiments of the present invention. Contrary to the practiced art in scientific literature and in an issued patent (Ahlquist et al., U.S. Pat. No. 5,466,788), uncapped transcripts for virus expression vectors are infective on both plants and in plant cells. Capping is not a prerequisite for establishing an infection of a virus expression vector in plants, although capping increases the efficiency of infection.

[**0061**] One feature of the recombinant plant viral nucleic acids useful in the present invention is that they further comprise one or more non-native nucleic acid sequences capable of being transcribed in a host plant. These nucleic acid sequences may be native nucleic acid sequences that occur in a host plant. Preferably, these nucleic acid sequences are non-native nucleic acid sequences that do not normally occur in a host plant. These nucleic acid sequences are derived from a donor plant, which is preferably genetically different from the host plant. The donor plant and the host plant may be genetically remote or unrelated: they may belong to different genus, family, order, class, subdivision, or division. Donor plants and host plants include plants of commercial interest, such as food crops, seed crops, oil crops, ornamental crops and forestry crops. For example, wheat, rice, corn, potatoes, barley, tobaccos, soybean canola, maize, oilseed rape, Arabidopsis, Nicotiana can be selected as a donor plant or a host plant. Host plants include those capable of being infected by an infectious RNA or a virus containing a recombinant viral nucleic acid. Preferred host plants include Nicotiana, preferably, *Nicotiana benthamiana,* or *Nicotiana cleaviandii.* Plant are grown from seed in a mixture of "Peat-Lite Mix™ (Speedling, Inc. Sun City, Fla.) and Nutricote™ controlled release fertilizer 14-14-14 (Chiss-Asahi Fertilizer Co., Tokyo, Japan). Plants are grown in a controlled environment provided 16 hours of light and 8 hours of darkness. Sylvania "Gro-Lux/Aquarium" wide spectrum 40 watt flourescent grow lights. (Osram Sylvania Products, Inc. Danvers, Mass.) are used. Temperatures are kept at around 80° F. during light hours and 70° F. during dark hours. Humidity is between 60 and 85%.

[**0062**] To prepare a DNA insert comprising a nucleic acid sequence of a donor plant, the first step is to construct a library of cDNAs, genomic DNAs, or a pool of RNAs of the plant of interest. Full-length cDNAs can be obtained from public or private repositories, for example, cDNA library of *Arabidopsis thaliana* can be obtained from the Arabidopsis Biological Resource Center. Alternatively, cDNA library can be prepared from a field sample by methods known to a person of ordinary skill, for example, isolating mRNAs and transcribing mRNAs into cDNAs by reverse transcriptase (see, e.g., Sambrook et al., *Molecular Cloning: A Laboratory Manual* (2nd ed.), Vols. 1-3, Cold Spring Harbor Laboratory, (1989), or *Current Protocols in Molecular Biology,* F. Ausubel et al., ed. Greene Publishing and Wiley-Interscience, New York (1987)). Genomic DNAs represented in BAC (bacterial artificial chromosome), YAC (yeast artificial chromosome), or TAC (transformation-competent artificial chromosome, Liu et al., *Proc. Natl. Acad. Sci.* USA, 96:6535-6540 (1999)) libraries can be obtained from public or private repositories, for example, the Arabidopsis Biological Resource Center. The BAC/YAC/TAC DNAs or cDNAs can be mechanically size-fractionated or digested by an enzyme to smaller fragments. The fragments are ligated to adapters with cohesive ends, and shotgun-cloned into recombinant viral nucleic acid vectors. Alternatively, the fragments can be blunt-end ligated into recombinant viral nucleic acid vectors. Recombinant plant viral nucleic acids containing a nucleic acid sequence derived from the cDNA library or genomic DNA library is then constructed using conventional techniques. The recombinant viral nucleic acid vectors produced comprise the nucleic acid insert derived from the donor plant. The nucleic acid sequence of the recombinant viral nucleic acid is transcribed as RNA in a host plant; the RNA is capable of regulating the expression of a phenotypic trait by an antisense mechanism. The nucleic acid sequence may also code for the expression of more than one phenotypic trait. Sequences from wheat, rice, corn, potato, barley, tobacco, soybean, canola, maize, oilseed rape, Arabidopsis, and other crop species may be used to

assemble the DNA libraries. This method may thus be used to search for useful dominant gene phenotypes from DNA libraries through the gene expression.

[0063] A further alternative when creating the recombinant plant viral nucleic acid is to prepare more than one nucleic acid (i.e., to prepare the nucleic acids necessary for a multipartite viral vector construct). In this case, each nucleic acid would require its own origin of assembly. Each nucleic acid could be prepared to contain a subgenomic promoter and a non-native nucleic acid.

[0064] In some embodiments of the instant invention, methods to increase the representation of gene sequences in virus expression libraries may also be achieved by bypassing the genetic bottleneck of propagation in bacterial cells. For example, cell-free methods may be used to assemble sequence libraries or individual arrayed sequences into virus expression vectors and reconstruct an infectious virus, such that the final ligation product can be transcribed and the resulting RNA can be used for plant, plant tissue or plant cell inoculation/infection. A more detailed discussion is presented in a co-pending/co-owned U.S. patent application Ser. No. 09/359,303 (Padgett et al., Attorney Docket No. 08010137US03, filed Jul. 21, 1999), incorporated herein by reference.

[0065] The host can be infected with a recombinant viral nucleic acid or a recombinant plant virus by conventional techniques. Suitable techniques include, but are not limited to, leaf abrasion, abrasion in solution, high velocity water spray, and other injury of a host as well as imbibing host seeds with water containing the recombinant viral RNA or recombinant plant virus. More specifically, suitable techniques include:

[0066] (a) Hand Inoculations. Hand inoculations are performed using a neutral pH, low molarity phosphate buffer, with the addition of celite or carborundum (usually about 1%). One to four drops of the preparation is put onto the upper surface of a leaf and gently rubbed.

[0067] (b) Mechanized Inoculations of Plant Beds. Plant bed inoculations are performed by spraying (gas-propelled) the vector solution into a tractor-driven mower while cutting the leaves. Alternatively, the plant bed is mowed and the vector solution sprayed immediately onto the cut leaves.

[0068] (c) Vacuum Infiltration. Inoculations may be accomplished by subjecting a host organism to a substantially vacuum pressure environment in order to facilitate infection.

[0069] (d) High Speed Robotics Inoculation. Especially applicable when the organism is a plant, individual organisms may be grown in mass array such as in microtiter plates. Machinery such as robotics may then be used to transfer the nucleic acid of interest.

[0070] (e) High Pressure Spray of Single Leaves. Single plant inoculations can also be performed by spraying the leaves with a narrow, directed spray (50 psi, 6-12 inches from the leaf) containing approximately 1 % carborundum in the buffered vector solution.

[0071] (f) Ballistics (High Pressure Gun) Inoculation. Single plant inoculations can also be performed by particle bombardment. A ballistics particle delivery system (BioRad Laboratories, Hercules, (A) can be used to transfect plants such as *N. benthamiana* as described previously (Nagar et al., *Plant Cell,* 7:705-719 (1995)).

[0072] An alternative method for introducing a recombinant plant viral nucleic acid into a plant host is a technique known as agroinfection or Agrobacterium-mediated transformation (sometimes called Agro-infection) as described by Grimsley et al., *Nature* 325:177 (1987). This technique makes use of a common feature of Agrobacterium which colonizes plants by transferring a portion of their DNA (the T-DNA) into a host cell, where it becomes integrated into nuclear DNA. The T-DNA is defined by border sequences which are 25 base pairs long, and any DNA between these border sequences is transferred to the plant cells as well. The insertion of a recombinant plant viral nucleic acid between the T-DNA border sequences results in transfer of the recombinant plant viral nucleic acid to the plant cells, where the recombinant plant viral nucleic acid is replicated, and then spreads systemically through the plant. Agro-infection has been accomplished with potato spindle tuber viroid (PSTV) (Gardner et al., *Plant Mol. Biol.* 6:221 (1986); CaV (Grimsley et al., *Proc. Natl. Acad. Sci.* USA 83:3282 (1986)); MSV (Grimsley et al., *Nature* 325:177 (1987)), and Lazarowitz, S., *Nucl. Acids Res.* 16:229 (1988)) digitaria streak virus (Donson et al., *Virology* 162:248 (1988)), wheat dwarf virus (Hayes et al., *J. Gen. Virol.* 69:891 (1988)) and tomato golden mosaic virus (TGMV) (Elmer et al., *Plant Mol.* 30 Biol. 10:225 (1988) and Gardiner et al., *EMBO J.* 7:899 (1988)). Therefore, agro-infection of a susceptible plant could be accomplished with a virion containing a recombinant plant viral nucleic acid based on the nucleotide sequence of any of the above viruses. Particle bombardment or electrosporation or any other methods known in the art may also be used.

[0073] Infection may also be attained by placing a selected nucleic acid sequence into an organism such as *E. coli,* or yeast, either integrated into the genome of such organism or not, and then applying the organism to the surface of the host organism. Such a mechanism may thereby produce secondary transfer of the selected nucleic acid sequence into a host organism. This is a particularly practical embodiment when the host organism is a plant. Likewise, infection may be attained by first packaging a selected nucleic acid sequence in a pseudovirus. Such a method is described in WO 4/10329, the teachings of which are incorporated herein by reference. Though the teachings of this reference may be specific for bacteria, those of skill in the art will readily appreciate that the same procedures could easily be adapted to other organisms.

[0074] After a host is infected with a recombinant viral nucleic acid comprising a nucleic acid insert derived from a cDNA library or a genomic library, one or more biochemical or phenotypic changes in a host plant is determined. The biochemical or phenotypic changes in the infected host plant is correlated to the biochemistry or phenotype of a host plant that is uninfected. Optionally, the biochemical or phenotypic changes in the infected host plant is further correlated to a host plant that is infected with a viral vector that contains a control nucleic acid of a known sequence in an antisense

orientation; the control nucleic acid has similar size but is different in sequence from the nucleic acid insert derived from the library. For example, if the nucleic acid insert derived from the library is identified as encoding a GTP binding protein in an antisense orientation, a nucleic acid derived from a gene encoding green fluorescent protein can be used as a control nucleic acid. Green fluorescent protein is known not be have the same effect as the GTP binding protein when expressed in plants.

[0075] Those of skill in the art will readily understand that there are many methods to determine phenotypic or biochemical change in a plant and to determine the function of a nucleic acid, once the nucleic acid is localized or systemic expressed in a host plant. In a preferred embodiment, the phenotypic or biochemical trait may be determined by observing phenotypic changes in a host by methods including visual, morphological, macroscopic or microscopic analysis. For example, growth change such as stunting, hyperbranching, and necrosis; structure change such as vein banding, ring spot, etching; color change such as bleaching, chlorosis, or other color; and other changes such as marginal, mottled, patterening, punctate, and reticulate are easily detected. In another embodiment, the phenotypic or biochemical trait may be determined by complementation analysis, that is, by observing the endogenous gene or genes whose function is replaced or augmented by introducing the nucleic acid of interest. A discussion of such phenomenon is provided by Napoli et al., *The Plant Cell* 2:279-289 (1990). In a third embodiment, the phenotypic or biochemical trait may be determined by analyzing the biochemical alterations in the accumulation of substrates or products from enzymatic reactions according to any means known by those skilled in the art. In a fourth embodiment, the phenotypic or biochemical trait may be determined by observing any changes in biochemical pathways which may be modified in a host organism as a result of expression of the nucleic acid. In a fifth embodiment, the phenotypic or biochemical trait may be determined utilizing techniques known by those skilled in the art to observe inhibition of endogenous gene expression in the cytoplasm of cells as a result of expression of the nucleic acid. In a sixth embodiment, the phenotypic or biochemical trait may be determined utilizing techniques known by those skilled in the art to observe changes in the RNA or protein profile as a result of expression of the nucleic acid. In a seventh embodiment, the phenotypic or biochemical trait may be determined by selection of organisms such as plants capable of growing or maintaining viability in the presence of noxious or toxic substances, such as, for example herbicides and pharmaceutical ingredients.

[0076] Phenotypic traits in plant cells, which may be observed microscopically, macroscopically or by other methods, include improved tolerance to herbicides, improved tolerance to extremes of heat or cold, drought, salinity or osmotic stress; improved resistance to pests (insects, nematodes or arachnids) or diseases (fungal, bacterial or viral), production of enzymes or secondary metabolites; male or female sterility; dwarfness; early maturity; improved yield, vigor, heterosis, nutritional qualities, flavor or processing properties, and the like. Other examples include the production of important proteins or other products for commercial use, such as lipase, melanin, pigments, alkaloids, antibodies, hormones, pharmaceuticals, antibiotics and the like. Another useful phenotypic trait is the production of degradative or inhibitory enzymes, for example, enzymes preventing or inhibiting the root development in malting barley, or enzymes determining response or non-response to a systemically administered drug in a human. The phenotypic trait may also be a secondary metabolite whose production is desired in a bioreactor.

[0077] Biochemical changes can also be determined by analytical methods, for example, in a high-throughput, fully automated fashion using robotics. Suitable biochemical analysis may include MALDI-TOF, LC/MS, GC/MS, two-dimensional IEF/SDS-PAGE, ELISA or other methods of analyses. The clones in the plant viral vector library may then be functionally classified based on metabolic pathway affected or visual/selectable phenotype produced in the plant. This process enables the rapid determination of gene function for unknown nucleic acid sequences of a plant origin. Furthermore, this process can be used to rapidly confirm function of full-length DNA's of unknown gene function. Functional identification of unknown nucleic acid sequences in a plant library may then rapidly lead to identification of similar unknown sequences in expression libraries for other crop species based on sequence homology.

[0078] One useful means to determine the function of nucleic acids transfected into a host is to observe the effects of gene silencing. Traditionally, functional gene knockout has been achieved following inactivation due to insertion of transposable elements or random integration of T-DNA into the chromosome, followed by characterization of conditional, homozygous-recessive mutants obtained upon backcrossing. Some teachings in these regards are provided by WO 97/42210 which is herein incorporated by reference. As an alternative to traditional knockout analysis, an EST/DNA library from an organism, for example *Arabidopsis thaliana,* may be assembled into a plant viral transcription plasmid. The nucleic acid sequences in the transcription plasmid library may then be introduced into plant cells as part of a functional RNA virus which post-transcriptionally silences the homologous target gene. The EST/DNA sequences may be introduced into a plant viral vector in either the plus or minus sense orientation, and the orientation can be either directed or random based on the cloning strategy. A high-throughput, automated cloning scheme based on robotics may be used to assemble and characterize the library. In addition, double stranded RNA may also be an effective stimulator of gene silencing in transgenic plant. Gene silencing of plant genes may be induced by delivering an RNA capable of base pairing with itself to form double stranded regions. This approach could be used with any plant gene to assist in the identification of the function of a particular gene sequence.

[0079] The present invention provides a method to produce transfected plants containing viral-derived antisense RNA in the cytoplasm. Such method is much faster than the time required to obtain genetically engineered antisense transgenic plants. Systemic infection and expression of viral antisense RNA occurs as short as four days post inoculation, whereas it takes several months or longer to create a single transgenic plant. The invention provides a method to identify genes involved in the regulation of plant growth by inhibiting the expression of specific endogenous genes using viral vectors, which replicate solely in the cytoplasm. This invention provides a method to characterize specific genes and biochemical pathways in donor plants or in host plants using an RNA viral vector.

[0080] The invention is also directed to a method of determining a nucleic acid sequence in a donor plant, which has the same function as that in a genetically different host plant, by transfecting a nucleic acid sequence derived from a donor plant into a plant host. In one preferred embodiment, the method comprising the steps of (a) preparing a library of eDNAs, genomic DNAs, or a pool of RNAs of the donor plant, (b) constructing recombinant viral nucleic acids comprising a nucleic acid insert derived from the library, (c) infecting each host plant with the one of the recombinant viral nucleic acids, (d) growing the infected host plant, (e) determining one or more changes in the host plant, (f) identifying the recombinant viral nucleic acid that results in changes in the host, (g) determining the sequence of the nucleic acid insert in the recombinant viral nucleic acid that results in changes in the host, and (h) determining the sequence of an entire open reading frame of the donor from which the nucleic acid insert is derived.

[0081] The invention is further directed to a method of determining a nucleic acid sequence in a host plant, which has the same function as that in a genetically different donor plant, by transfecting a nucleic acid sequence derived from a donor plant into a host plant. In one preferred embodiment, the method comprising the steps of (a) preparing a library of cDNAs, genomic DNAs, or a pool of RNAs of the donor plant, (b) constructing recombinant viral nucleic acids comprising a nucleic acid insert derived from the library, (c) infecting each host plant with one of said recombinant viral nucleic acids, (d) growing the infected host plant, (e) determining one or more changes in the host plant, (f) identifying the recombinant viral nucleic acid that results in changes in the host, (g) determining the sequence of the nucleic acid insert in the recombinant viral nucleic acid that results in changes in the host, and (h) determining the sequence of an entire open reading frame of a gene in the host plant, the expression of which is affected by the insert. The sequence of the nucleic acid insert in the cDNA clone or in the viral vector can be determined by a standard method, for example, by dideoxy termination using double stranded templates (Sanger et al., *Proc., Natl. Acad. Sci.* USA 74:5463-5467 (1977)). Once the sequence of the nucleic acid insert is obtained, the sequence of an entire open reading frame of a gene can be determined by probing filters containing full-length cDNAs from the cDNA library with the nucleic acid insert labeled with radioactive, fluorescent, or enzyme molecules. The sequence of an entire open reading frame of a gene can also be determined by RT-PCR (*Methods Mol. Biol.,* 89:333-358 (1998)).

[0082] The present invention is also directed to a method of changing the phenotype or biochemistry of a plant by expressing transiently a nucleic acid sequence from a donor plant in an antisense orientation in a host plant, which inhibits an endogenous gene expression in the meristem of the host plant. The one or more phenotypic or biochemical changes in the host plant are detected by methods as describes previously. Transient expressing a nucleic acid sequence in a host plant can affect the gene expression in meristem. Meristems are of interest in plant development because plant growth is driven by the formation and activity of meristems throughout the entire life cycle. This invention is exemplified by a nucleic acid sequence encoding ribosomal protein S 18. The activity of S 18 promoter is restricted to meristems (Lijsebettesn et al., *EMBO J.* 13:3378-3388). Transient expression of a nucleic acid sequence in a host plant can trigger a signal transmitting to meristems and affect the gene expression in meristem.

[0083] One problem with gene silencing in a plant host is that many plant genes exist in multigene families. Therefore, effective silencing of a gene function may be especially problematic. According to the present invention, however, nucleic acids may be inserted into the viral genome to effectively silence a particular gene function or to silence the function of a multigene family. It is presently believed that about 20% of plant genes exist in multigene families.

[0084] A detailed discussion of some aspects of the "gene silencing" effect is provided in the co-pending patent application, U.S. patent application Ser. No. 08/260,546 (WO95/34668 published Dec. 21, 1995) the disclosure of which is incorporated herein by reference. RNA can reduce the expression of a target gene through inhibitory RNA interactions with target mRNA that occur in the cytoplasm and/or the nucleus of a cell.

[0085] An EST/cDNA library from a plant such as *Arabidopsis thaliana* may be assembled into a plant viral transcription plasmid background. The cDNA sequences in the transcription plasmid library can then be introduced into plant cells as cytoplasmic RNA in order to post-transcriptionally silence the endogenous genes. The EST/cDNA sequences may be introduced into the plant viral transcription plasmid in either the plus or anti-sense orientation (or both), and the orientation can be either directed or random based on the cloning strategy. A high-throughput, automated cloning strategy using robotics can be used to assemble the library. The EST clones can be inserted behind a duplicated subgenomic promoter such that they are represented as subgenomic transcripts during viral replication in plant cells. Alternatively, the EST/cDNA sequences can be inserted into the genomic RNA of a plant viral vector such that they are represented as genomic RNA during the viral replication in plant cells. The library of EST clones is then transcribed into infectious RNAs and inoculated onto a host plant susceptible to viral infection. The viral RNAs containing the EST/cDNA sequences contributed from the original library are now present in a sufficiently high concentration in the cytoplasm of host plant cells such that they cause post-transcriptional gene silencing of the endogenous gene in a host plant. Since the replication mechanism of the virus produces both sense and antisense RNA sequences, the orientation of the EST/cDNA insert is normally irrelevant in terms of producing the desired phenotype in the host plant.

[0086] It is known that silencing of endogenous genes can be achieved with homologous sequences from the same family. For example, Kumagai et al., (*Proc. Natl. Acad. Sci.* USA 92:1679 (1995)) report that the *Nicotiana benthamiana* gene for phytoene desaturase (PDS) was silenced by trans-

fection with a viral RNA derived from a clone containing a partial tomato (*Lycopersicon esculentum*) cDNA encoding PDS being in an antisense orientation. This paper is incorporated here by reference. Kumagai et al. demonstrate that gene encoding PDS from one plant can be silenced by transfection a host plant with a nucleic acid of a known sequence, namely, a PDS gene, from a donor plant of the same family. The present invention provides a method of silencing a gene in a host plant by transfecting the host plant with a viral nucleic acid comprising a nucleic acid insert derived from a cDNA library, a genomic DNA library, or a pool of RNA. Different from Kumagai et al, the sequence of the nucleic acid insert in the present invention is not identified or isolated prior to the transfection. Another feature of the present invention is that it provides a method to silence a gene of a different family; the antisense transcript of one plant results in reducing expression of the endogenous gene or multigene family of a plant of a different genus, family, order, class, subdivision, division, or subkingdom. The invention is exemplified by GTP binding proteins. In eukaryotic cells, GTP-binding proteins function in a variety of cellular processes, including signal transduction, cytoskeletal organization, and protein transport. Low molecular weight (20-25 K Daltons) of GTP-binding proteins include ras and its close relatives (for example, Ran), rho and its close close relatives, the rab family, and the ADP-ribosylation factor (ARF) family. The heterotrimeric and monomeric GTP-binding proteins that may be involved in secretion and intracellular transport are divided into two structural classes: the rab and the ARF families. The ARFs from many organisms have been isolated and characterized. The ARFs share structural features with both the ras and trimeric GTP-binding protein families. The present invention demonstrates that genes of one plant, such as Nicotiana, which encode GTP binding proteins, can be silenced by transfection with infectious RNA from a clone containing GTP binding protein open reading frame in an antisense orientation, derived from a plant of a different family, such as Arabidopsis.

[**0087**] The present invention also provides a method of isolating a conserved gene such as a gene encoding a GTP binding protein, from rice, barley, corn, soybean, maize, oilseed, and other plant of commercial interest, using another gene having homology with gene being isolated. Libraries containing full-length cDNAs from a donor plant such as rice, barley, corn, soybean and other important crops can be obtained from public and private sources or can be prepared from plant mRNAs. The cDNAs are inserted in viral vectors or in small subcloning vectors such as pBluescript (Strategene), pUC18, M13, or pBR322. Transformed bacteria are then plated and individual clones selected by a standard method. The bacteria transformants or DNAs are rearrayed at high density onto membrane filters or glass slides. Full-length cDNAs encoding GTP binding proteins can be identified by probing filters or slides with labeled nucleic acid inserts which result in changes in a host plant, or labeled probes prepared from DNAs encoding GTP binding proteins from Arabidopsis. Useful labels include radioactive, fluorescent, or chemiluminecent molecules, enzymes, etc.

[**0088**] Alternatively, genomic libraries containing sequences from rice, barley, corn, soybean and other important crops can be obtained from public and private sources, or be prepared from plant genomic DNAs. BAC clones containing entire plant genomes have been constructed and organized in a minimal overlapping order. Individual BACs are sheared to fragments and directly cloned into viral vectors. Clones that completely cover an entire BAC form a BAC viral vector sublibrary. Genomic clones can be identified by probing filters containing BACs with labeled nucleic acid inserts which result in changes in a host plant, or with labeled probes prepared from DNAs encoding GTP binding proteins from Arabidopsis. Useful labels include radioactive, fluorescent, or chemiluminecent molecules, enzymes, etc. BACs that hybridize to the probe are selected and their corresponding BAC viral vectors are used to produce infectious RNAs. Plants that are transfected with the BAC sublibrary are screened for change of function, for example, change of growth rate or change of color. Once the change of function is observed, the inserts from these clones or their corresponding plasmid DNAs are characterized by dideoxy sequencing. This provides a rapid method to obtain the genomic sequence for a plant protein, for example, a GTP binding protein. Using this method, once the DNA sequence in one plant such as *Arabidopsis thaliana* is identified, it can be used to identify conserved sequences of similar function that exist in other plant libraries.

[**0089**] Large amounts of DNA sequence information are being generated in the public domain and may be entered into a relational database. Links may be made between sequences from various species predicted to carry out similar biochemical or regulatory functions. Links may also be generated between predicted enzymatic activities and visually displayed biochemical and regulatory pathways. Likewise, links may be generated between predicted enzymatic or regulatory activity and known small molecule inhibitors, activators, substrates or substrate analogs. Phenotypic data from expression libraries expressed in transfected hosts may be automatically linked within such a relational database. Genes with similar predicted roles of interest in other crop plants may be rapidly discovered.

[**0090**] A functional genomics screen is set up using a tobacco mosaic virus TMV-O coat protein capsid for infection of *Nicotiana benthamiana,* a plant related to the common tobacco plant. For *Arabidopsis thaliana* cDNA librariers are obtained from the Arabidopsis Biological Resource Center, Bluescript® phagemid vectors are recovered by Not 1 digestion. cDNA is transformed into a plasmid. The plasmid is transcribed into viral vector RNA. The inserts are in the antisense orientation as in Figure until all of the cDNA from each cDNA library is represented on viral vectors. Each viral vector is sprayed onto the leaf of a two-week old *N. benthamiana* plant host with sufficient force to cause tissue injury and localized viral infection. Each infected

plant is grown side by side with an uninfected plant and a plant infected with a null insert vector as controls. All plants are grown in an artificial environment having 16 hours of light and 8 hours of dark. Lumens are approximately equal on each plant. At intervals of 2 days a visual and photographic observation of phenotype is made and recorded for each infected plant and each of its controls and a comparison is made. Data is entered into a Laboratory Information Management System database. At the end of the observation period stunted plants are grouped for analysis. The nucleic acid insert contained in the viral vector clone 740AT#120 is responsible for severe stunting of one of the plants. Clone 740AT #120 is sequenced. The homolog from the plant host is also sequenced. The 740AT #120 clone is found to have 80% hemology to plant host nucleic acid sequence. The amino acid sequence of hemology is 96%. The entire cDNA sequence of the insert is obtained by sequencing and found to code for a GTP binding protein. The host plant homolog is selected and sequenced. It also codes for a GTP binding protein. We conclude that this GTP binding protein coding sequence is highly conserved in nature. This information is useful in pharmaceutical development as well as in toxicology studies.

[0091] A complete classification scheme of gene functionality for a filly sequenced eukaryotic organism has been established for yeast. This classification scheme may be modified for plants and divided into the appropriate categories. Such organizational structure may be utilized to rapidly identify herbicide target loci which may confer dominant lethal phenotypes, and thereby is useful in helping to design rational herbicide programs.

[0092] The present invention is also directed to a method of increasing yield of a grain crop. In *Rice Biotechnology Quarterly* (37:4, (1999)), it is reported that a transgenic rice plant transformed with a rgp1 gene, which encodes a small GTP binding protein from rice, was shorter than a control plant, but it produced more seeds than the control plant. To increase the yield of a grain crop, the present method comprises expressing transiently a nucleic acid sequence of a donor plant in an antisense orientation in the grain crop, wherein said expressing results in stunted growth and increased seed production of said grain crop. A preferred method comprises the steps of cloning the nucleic acid sequence into a plant viral vector and infecting the grain crop with a recombinant viral nucleic acid comprising said nucleic acid sequence. Preferred plant viral vector is derived from a Brome Mosaic virus, a Rice Necrosis virus, or a geminivirus. Preferred grain crops include rice, wheat, and barley. The nucleic acid expressed in the host plant, for example, comprises a GTP binding protein open reading frame having an antisense orientation. The present method provides a transiently expression of a gene to obtain a stunted plant. Because less energy is put into plant growth, more energy is available for production of seed, which results in increase yield of a grain crop. The present method has an advantage over other method using a trangenic plant, because it does not have an effect on the genome of a host plant.

[0093] In order to provide an even clearer and more consistent understanding of the specification and the claims, including the scope given herein to such terms, the following definitions are provided:

[0094] Adjacent: A position in a nucleotide sequence proximate to and 5' or 3' to a defined sequence. Generally, adjacent means within 2 or 3 nucleotides of the site of reference.

[0095] Anti-Sense Inhibition: A type of gene regulation based on cytoplasmic, nuclear or organelle inhibition of gene expression due to the presence in a cell of an RNA molecule complementary to at least a portion of the mRNA being translated. It is specifically contemplated that RNA molecules may be from either an RNA virus or mRNA from the host cells genome or from a DNA virus.

[0096] Cell Culture: A proliferating group of cells which may be in either an undifferentiated or differentiated state, growing contiguously or non-contiguously.

[0097] Chimeric Sequence or Gene: A nucleotide sequence derived from at least two heterologous parts. The sequence may comprise DNA or RNA.

[0098] Coding Sequence: A deoxyribonucleotide or ribonucleotide sequence which, when either transcribed and translated or simply translated, results in the formation of a cellular polypeptide or a ribonucleotide sequence which, when translated, results in the formation of a cellular polypeptide.

[0099] Compatible: The capability of operating with other components of a system. A vector or plant or animal viral nucleic acid which is compatible with a host is one which is capable of replicating in that host. A coat protein which is compatible with a viral nucleotide sequence is one capable of encapsidating that viral sequence.

[0100] Complementation Analysis: As used herein, this term refers to observing the changes produced in an organism when a nucleic acid sequence is introduced into that organism after a selected gene has been deleted or mutated so that it no longer functions fully in its normal role. A complementary gene to the deleted or mutated gene can restore the genetic phenotype of the selected gene.

[0101] Dual Heterologous Subgenomic Promoter Expression System (DHSPES): a plus stranded RNA vector having a dual heterologous subgenomic promoter expression system to increase, decrease, or change the expression of proteins, peptides or RNAs, preferably those described in U.S. Pat. Nos. 5,316,931, 5,811,653, 5,589,367, and 5,866,785, the disclosure of which is incorporated herein by reference.

[0102] Expressed sequence tags (ESTs): Relatively short single-pass DNA sequences obtained from one or more ends of cDNA clones and RNA derived therefrom. They may be present in either the 5' or the 3' orientation. ESTs have been shown useful for identifying particular genes.

[0103] Expression: The term as used herein is meant to incorporate one or more of transcription, reverse transcription and translation.

[0104] A functional Gene Profile: The collection of genes of an organism which code for a biochemical or phenotypic trait. The functional gene profile of an organism is found by screening nucleic acid sequences from a donor organism by over expression or suppression of a gene in a host organism. A functional gene profile requires a collection or library of nucleic acid sequences from a donor organism. A functional gene profile will depend on the ability of the collection or library of donor nucleic acids to cause over-expression or suppression in the host organism. Therefore, a functional gene profile will depend upon the quantity of donor genes capable of causing over-expression or suppression of host genes or of being expressed in the host organism in the absence of a homologous host gene.

[0105] Gene: A discrete nucleic acid sequence responsible for producing one or more cellular products and/or performing one or more intercellular or intracellular functions.

[0106] Gene silencing: A reduction in gene expression. A viral vector expressing gene sequences from a host may induce gene silencing of homologous gene sequences.

[0107] Homology: A high degree of nucleic acid similarity in all or some portions of a gene sequence sufficient to result in gene suppression if the nucleic acid sequence is delivered in the antisense orientation.

[0108] Host: A cell, tissue or organism capable of replicating a nucleic acid such as a vector or plant viral nucleic acid and which is capable of being infected by a virus containing the viral vector or viral nucleic acid. This term is intended to include prokaryotic and eukaryotic cells, organs, tissues or organisms, where appropriate. Bacteria, fungi, yeast, animal (cell, tissues, or organisms), and plant (cell, tissues, or organisms) are examples of a host.

[0109] Infection: The ability of a virus to transfer its nucleic acid to a host or introduce a viral nucleic acid into a host, wherein the viral nucleic acid is replicated, viral proteins are synthesized, and new viral particles assembled. In this context, the terms "transmissible" and "infective" are used interchangeably herein. The term is also meant to include the ability of a selected nucleic acid sequence to integrate into a genome, chromosome or gene of a target organism.

[0110] Multigene family: A set of genes descended by duplication and variation from some ancestral gene. Such genes may be clustered together on the same chromosome or dispersed on different chromosomes. Examples of multigene families include those which encode the histones, hemoglobins, immunoglobulins, histocompatibility antigens, actins, tubulins, keratins, collagens, heat shock proteins, salivary glue proteins, chorion proteins, cuticle proteins, yolk proteins, and phaseolins.

[0111] Non-Native: Any RNA or DNA sequence that does not normally occur in the cell or organism in which it is placed. Examples include recombinant plant viral nucleic acids and genes or ESTs contained therein. That is, an RNA or DNA sequence may be non-native with respect to a viral nucleic acid. Such an RNA or DNA sequence would not naturally occur in the viral nucleic acid. Also, an RNA or DNA sequence may be non-native with respect to a host organism. That is, such a RNA or DNA sequence would not naturally occur in the host organism.

[0112] Nucleic acid: As used herein the term is meant to include any DNA or RNA sequence from the size of one or more nucleotides up to and including a complete gene sequence. The term is intended to encompass all nucleic acids whether naturally occurring in a particular cell or organism or non-naturally occurring in a particular cell or organism.

[0113] Nucleic acid of interest: The term is intended to refer to the nucleic acid sequence whose function is to be determined. The sequence will normally be non-native to a viral vector but may be native or non-native to a host organism.

[0114] Phenotypic Trait: An observable, measurable or detectable property resulting from the expression or suppression of a gene or genes.

[0115] Plant Cell: The structural and physiological unit of plants, consisting of a protoplast and the cell wall.

[0116] Plant Organ: A distinct and visibly differentiated part of a plant, such as root, stem, leaf or embryo.

[0117] Plant Tissue: Any tissue of a plant in planta or in culture. This term is intended to include a whole plant, plant cell, plant organ, protoplast, cell culture, or any group of plant cells organized into a structural and functional unit.

[0118] Positive-sense inhibition: A type of gene regulation based on cytoplasmic inhibition of gene expression due to the presence in a cell of an RNA molecule substantially homologous to at least a portion of the mRNA being translated.

[0119] Promoter: The 5'-flanking, non-coding sequence substantially adjacent a coding sequence which is involved in the initiation of transcription of the coding sequence.

[0120] Protoplast: An isolated plant or bacterial cell without some or all of its cell wall.

[0121] Recombinant Plant Viral Nucleic Acid: Plant viral nucleic acid which has been modified to contain non-native nucleic acid sequences. These non-native nucleic acid sequences may be from any organism or purely synthetic, however, they may also include nucleic acid sequences naturally occurring in the organism into which the recombinant plant viral nucleic acid is to be introduced.

[0122] Recombinant Plant Virus: A plant virus containing the recombinant plant viral nucleic acid.

[0123] Subgenomic Promoter: A promoter of a subgenomic mRNA of a viral nucleic acid.

[0124] Substantial Sequence Homology: Denotes nucleotide sequences that are substantially functionally equivalent to one another. Nucleotide differences between such

sequences having substantial sequence homology are insignificant in affecting function of the gene products or an RNA coded for by such sequence.

[0125] Systemic Infection: Denotes infection throughout a substantial part of an organism including mechanisms of spread other than mere direct cell inoculation but rather including transport from one infected cell to additional cells either nearby or distant.

[0126] Transposon: A nucleotide sequence such as a DNA or RNA sequence which is capable of transferring location or moving within a gene, a chromosome or a genome.

[0127] Transgenic plant: A plant which contains a foreign nucleotide sequence inserted into either its nuclear genome or organellar genome.

[0128] Transcription: Production of an RNA molecule by RNA polymerase as a complementary copy of a DNA sequence or subgenomic mRNA.

[0129] Transient Expression: expression of a nucleic acid sequence in a host without insertion of the nucleic acid sequence into the host genome, such as by way of a viral vector.

[0130] Vector: A self-replicating RNA or DNA molecule which transfers an RNA or DNA segment between cells, such as bacteria, yeast, plant, or animal cells.

[0131] Virus: An infectious agent composed of a nucleic acid which may or may not be encapsidated in a protein. A virus may be a mono-, di-, tri-, or multi-partite virus, as described above.

EXAMPLES

[0132] The following examples further illustrate the present invention. These examples are intended merely to be illustrative of the present invention and are not to be construed as being limiting.

EXAMPLE 1

[0133] Gene Silencing/Co-Supression of Genes Induced by Delivering an RNA Capable of Base Pairing with Itself to Form Double Stranded Regions

[0134] Gene silencing has been used to down regulate gene expression in transgenic plants. Recent experimental evidence suggests that double stranded RNA may be an effective stimulator of gene silencing/co-suppression phenomenon in transgenic plant. For example, Waterhouse et al. (*Proc. Natl. Acad. Sci.* USA 95:13959-13964 (1998), incorporated herein by reference) described that virus resistance and gene silencing in plants could be induced by simultaneous expression of sense and antisense RNA. Gene silencing/co-suppression of plant genes may be induced by delivering an RNA capable of base pairing with itself to form double stranded regions.

[0135] This example shows: (1) a novel method for generating an RNA virus vector capable of producing an RNA capable of forming double stranded regions, and (2) a process to silence plant genes by using such a viral vector.

[0136] Step **1**: Construction of a DNA sequence which after it is transcribed would generate an RNA molecule capable of base pairing with itself. Two identical, or nearly identical, ds DNA sequences are ligated together in an inverted orientation to each other (i.e., in either a head to tail or tail to head orientation) with or without a linking nucleotide sequence between the homologous sequences. The resulting DNA sequence is then be cloned into a cDNA copy of a plant viral vector genome.

[0137] Step **2**: Cloning, screening, transcription of clones of interest using known methods in the art.

[0138] Step **3**: Infect plant cells with transcripts from clones.

[0139] As virus expresses foreign gene sequence, RNA from foreign gene forms base pair upon itself, forming double-stranded RNA regions. This approach is used with any plant or non-plant gene and used to silence plant gene homologous to assist in identification of the function of a particular gene sequence.

EXAMPLE 2

[0140] Cytoplasmic Inhibition of Phytoene Desaturase in Transfected Plant Confirms That the Partial Tomato PDS Sequence Encodes Phytoene Desaturase

[0141] Isolation of tomato mosaic virus cDNA. An 861 base pair fragment (5524-6384) from the tomato mosaic virus (fruit necrosis strain F; tom-F) containing the putative coat protein subgenomic promoter, coat protein gene, and the 3'-end was isolated by PCR using primers

```
5'-CTCGCAAAGTTTCGAACCAAATCCTC-3' (upstream)            (SEQ ID NO: 1)
                                                       and

5'-CGGGGTACCTGGGCCCCAACCGGGGGTTCCGGGGG-3' (downstream) (SEQ ID NO: 2)
```

[0142] and subcloned into the HincII site of pBluescript KS-. A hybrid virus consisting of TMV-U1 and ToMV-F was constructed by swapping an 874-bp BamHI-KpnI ToMV fragment into pBGC152, creating plasmid TTO1. The inserted fragment was verified by dideoxynucleotide sequencing. A unique AvrII site was inserted downstream of the XhoI site in TTO1 by PCR mutagenesis, creating plasmid TTO1A, using the following oligonucleotides:

5'-TCCTCGAGCCTAGGCTCGCAAAGTTTCGAACCAAATCCTCA-3' (upstream) (SEQ ID NO: 3),

5'-CGGGGTACCTGGGCCCCAACCGGGGGTTCCGGGGG-3' (downstream) (SEQ ID NO: 4).

[0143] Isolation of a partial cDNA encoding tomato phytoene desaturase. Partial cDNAs were isolated from ripening tomato fruit RNA by polymerase chain reaction (PCR) using the following oligonucleotides: PDS, 5'-TGCTCGAGTGTGTTCTTCAGTTTTCTGTCA-3' (SEQ ID NO: 5) (upstream), 5'-AACTCGAGCGCTTTGATTTCTCCGAAGCTT-3' (downstream) (SEQ ID NO:

[0144] 6). Approximately 3×10^4 colonies from a Lycopersicon esculentum cDNA library were screened by colony hybridization using a ^{32}P labeled tomato phytoene desaturase PCR product. Hybridization was carried out at 42° C. for 48 hours in 50% formamide, 5×SSC, 0.02M phosphate buffer, 5×Denhart's solution, and 0.1 mg/ml sheared calf thymus DNA. Filters were washed at 65° C. in 0.1×SSC, 0.1% SDS prior to autoradiography. PCR products and the phytoene desaturase cDNA clones were verified by dideoxynucleotide sequencing.

[0145] DNA sequencing and computer analysis. A PstI, BamHI fragment containing the phytoene synthase cDNA and the partial phytoene desaturase cDNA was subcloned into pBluescript® KS+ (Stratagene, La Jolla, Calif.). The nucleotide sequencing of KS+/PDS #38 was carried out by dideoxy termination using single-stranded templates (Maniatis, *Molecular Cloning,* 1st Ed.) Nucleotide sequence analysis and amino acid sequence comparisons were performed using PCGENE® and DNA Inspector® IIE programs.

[0146] Construction of a viral vector containing a partial tomato phytoene desaturase cDNA. A XhoI fragment containing the partial tomato phytoene desaturase cDNA was subcloned into TTO1. The vector TTOlA/PDS +(**FIG. 1**) contains the phytoene desaturase cDNA in the positive orientation under the control of the TMV-U1 coat protein subgenomic promoter; while the vector TTO1A/PDS− contains the phytoene desaturase cDNA in the antisense orientation.

[0147] Analysis of *N. benthamiana* transfected by TTO1/PDS+, and TTO1/PDS−. Infectious RNAs from TTO1/PDS+, TTO1/PDS+ were prepared by in vitro transcription using SP6 DNA-dependent RNA polymerase as described previously (Dawson et al, *Proc. Natl. Acad. Sci.* USA 83:1832 (1986)) and were used to mechanically inoculate *N. benthamiana.* The hybrid viruses spread throughout all the non-inoculated upper leaves as verified by transmission electron microscopy, local lesion infectivity assay, and polymerase chain reaction (PCR) amplification. The viral symptoms resulting from the infection consisted of distortion of systemic leaves and plant stunting with mild chlorosis. The leaves from plants transfected with TTO1/PDS+ and TTO1/PDS− turned white. Agarose gel eletrophoresis of PCR cDNA isolated from virion RNA and Northern blot analysis of virion RNA indicate that the vectors are maintained in an extrachromosomal state and have not undergone any detectable intramolecular rearrangements.

[0148] Purification and analysis of carotenoids from transfected plants. The carotenoids were isolated from systemically infected tissue and analyzed by HPLC chromatography. Carotenoids were extracted in ethanol and identified by their peak retention time and absorption spectra on a 25-cm Spherisorb® ODS-15-m column using acetonitrile/methanol/2-propanol (85:10:5) as a developing solvent at a flow rate of 1 ml/min. They had identical retention time to a synthetic phytoene standard and β-carotene standards from carrot and tomato. The expression of sense and antisense RNA to a partial phytoene desaturase in transfected plants inhibited the synthesis of colored carotenoids and caused the systemically infected leaves to turn white. HPLC analysis of these plants revealed that they also accumulated phytoene. The white leaf phenotype was also observed in plants treated with the herbicide norflurazon which specifically inhibits phytoene desaturase.

[0149] Our results that phytoene accumulated in plants transfected with antisense phytoene desaturase suggests that viral vectors can be used as a potent tool to manipulate pathways in the production of secondary metabolites through cytoplasmic antisense inhibition. Leaves from systemically infected TTO1A/PDS+ plants also accumulated phytoene and developed a bleaching white phenotype; the actual mechanism of inhibition is not clear. These data are presented by Kumagai et al., *Proc. Natl. Acad. Sci.* USA 92:1679-1683 (1995).

EXAMPLE 3

[0150] Cytoplasmic Inhibition of 5-Enolpyruvylshikimate-3-Phosphate Synthase (EPSPS) Genes in Plants Blocks Aromatic Amino Acid Biosynthesis

[0151] Cytoplasmic inhibition of 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) genes in plants blocks aromatic amino acid biosynthesis and causes a systemic bleaching phenotype similar to Roundup® herbicide. See also della-Cioppa, et al., "Genetic Engineering of herbicide resistance in plants,"*Frontiers of Chemistry: Biotechnology,* Chemical Abstract Service, ACS, Columbus, Ohio, pp. 665-70 (1989). A dual subgenomic promoter vector encoding 1097 base pairs of an antisense EPSPS gene from *Nicotiana tabacum* (Class I EPSPS) is shown in plasmid pBS712. **FIG. 2** shows plasmid pBS712. Systemic expression of the *Nicotiana tabacum* Class I EPSPS gene in the antisense orientation causes a systemic bleaching phenotype similar to Roundup® herbicide.

EXAMPLE 4

[0152] Exemplary Complementation Analysis

[0153] A transgenic plant or naturally occurring plant mutant may have a non-functional gene such as the one which produces EPSPS. A plant deficient or lacking in the EPSPS gene could grow only in the presence of added aromatic amino acids. Transfection of plants with a viral vector containing a functional EPSPS gene or cDNA sequence encoding the same would cause the plant to produce a functional EPSPS gene product. A plant so

transfected would then be able to grow normally without added aromatic amino acids to its environment. In this transfected plant, the EPSPS mutation in the plant would be complemented in trans by the viral nucleic acid sequence containing the native or foreign EPSPS cDNA sequence.

EXAMPLE 5

[0154] Construction of a Tobamoviral Vector for Expression of Heterologous Genes in *A. thaliana*

[0155] Virions that were prepared as a crude aqueous extract of tissue from turnip infected with Ribgrass mosaic virus (RMV) were used to inoculate *N. benthamiana, N. tabacum, A. thaliana,* and oilseed rape (canola). Two to three weeks after transfection, systemically infected plants were analyzed by immunoblotting, using purified RMV as a standard. Total soluble plant protein concentrations were determined using bovine serum albumin as a standard. The proteins were analyzed on a 0.1% SDS/12.5% polyacrylamide gel and transferred by electroblotting for 1 hr to a nitrocellulose membrane. The blotted membrane was incubated for 1 hr with a 2000-fold dilution of anti-ribgrass mosaic virus coat antiserum. Using standard protocols, the antisera was raised in rabbits against purified RMV coat protein. The enhanced chemiluminescence horseradish peroxidase-linked, goat anti-rabbit IgG assay (Cappel Laboratories) was performed according to the manufacturer's (Amersham) specifications. The blotted membrane was subjected to film exposure times of up to 10 sec. No detectable cross-reacting protein was observed in the noninfected *N. benthamiana* control plant extracts. A 18 kDa protein cross-reacted to the anti-RMV coat antibody from systemically infected *N. benthamiana, N. tabacum, A. thaliana,* and oilseed rape (canola). This result demonstrates that RMV can systemically infect *N. benthamiana, N. tabacum, A. thaliana,* and oilseed rape (canola).

[0156] Plasmid Constructions

[0157] Ribgrass mosaic virus (RMV) is a member of the tobamovirus group that infects crucifers. A partial RMV cDNA containing the 30K subgenomic promoter, 30K ORF, coat subgenomic promoter, coat ORF, and 3'-end was isolated by RT-PCR using by using oligonucleotides TVCV183X, 5'-TAC TCG AGG TTC ATA AGA CCG CGG TAG GCG G-3' (upstream) (SEQ ID NO: 7) and TVCV KpnI, 5'-CGG GGT ACC TGG GCC CCT ACC CGG GGT TTA GGG AGG-3' (downstream) (SEQ ID NO: 8), and subcloned into the EcoRV site of KS+, creating plasmid KS+ TVCV #23 (**FIG. 3**). The RMV cDNA was characterized by restriction mapping and dideoxy nucleotide sequencing. The partial nucleotide sequence is as follows:

5'-CTCGAGGTTCATAAGACCGCGGTAGGCGGAGCGTTTGTTTACTGTAG

TATAATTAAATATTTGTCAGATAAAAGGTTGTTTAAAGATTTGTTTTTTG

TTTGACTGAGTCGATAATGTCTTACGAGCCTAAAGTTAGTGACTTCCTTG

CTCTTACGAAAAAGGAGGAAATTTTACCCAAGGCTTTGACGAGATTAAAG

ACTGTCTCTATTAGTACTAAGGATGTTATATCTGTTAAGGAGTCTGAGTC

CCTGTGTGATATTGATTTGTTAGTGAATGTGCCATTAGATAAGTATAGGT

ATGTGGGTGTTTTGGGTGTTGTTTTCACCGGTGAATGGCTGGTACCGGAT

-continued

TTCGTTAAAGGTGGGGTAACAGTGAGCGTGATTGACAAACGGCTTGAAAA

TTCCAGAGAGTGCATAATTGGTACGTACCGAGCTGCTGTAAAGGACAGAA

GGTTCCAGTTCAAGCTGGTTCCAAATTACTTCGTATCCATTGCGGATGCC

AAGCGAAAACCGTGGCAGGTTCATGTGCGAATTCAAAATCTGAAGATCGA

AGCTGGATGGCAACCTCTAGCTCTAGAGGTGGTTTCTGTTGCCATGGTTA

CTAATAACGTGGTTGTTAAAGGTTTGAGGGAAAAGGTCATCGCAGTGAAT

GATCCGAACGTCGAAGGTTTCGAAGGTGTGGTTGACGATTTCGTCGATTC

GGTTGCTGCATTCAAGGCGATTGACAGTTTCCGAAAGAAAAAGAAAAAGA

TTGGAGGAAGGGATGTAAATAATAATAAGTATAGATATAGACCGGAGAGA

TACGCCGGTCCTGATTCGTTACAATATAAAGAAGAAAATGGTTTACAACA

TCACGAGCTCGAATCAGTACCAGTATTTCGCAGCGATGTGGGCAGAGCCC

ACAGCGATGCTTAACCAGTGCGTGTCTGCGTTGTCGCAATCGTATCAAAC

TCAGGCGGCAAGAGATACTGTTAGACAGCAGTTCTCTAACCTTCTGAGTG

CGATTGTGACACCGAACCAGCGGTTTCCAGAAACAGGATACCGGGTGTAT

ATTAATTCAGCAGTTCTAAAACCGTTGTACGAGTCTCTCATGAAGTCCTT

TGATACTAGAAATAGGATCATTGAAACTGAAGAAGAGTCGCGTCCATCGG

CTTCCGAAGTATCTAATGCAACACAACGTGTTGATGATGCGACCGTGGCC

ATCAGGAGTCAAATTCAGCTTTTGCTGAACGAGCTCTCCAACGGACATGG

TCTGATGAACAGGGCAGAGTTCGAGGTTTTATTACCTTGGGCTACTGCGC

CAGCTACATAGGCGTGGTGCACACGATAGTGCATAGTGTTTTTCTCTCCA

CTTAAATCGAAGAGATATACTTACGGTGTAATTCCGCAAGGGTGGCGTAA

ACCAAATTACGCAATGTTTTAGGTTCCATTTAAATCGAAACCTGTTATTT

CCTGGATCACCTGTTAACGTACGCGTGGCGTATATTACAGTGGGAATAAC

TAAAAGTGAGAGGTTCGAATCCTCCCTAACCCCGGGTAGGGGCCCA-3'

(SEQ ID NO: 9).

[0158] The 1543 base pair from the partial RMV cDNA was compared (PCGENE) to oilseed rape mosaic virus (ORMV). The nucleotide sequence identity was 97.8%. The RMV 30K and coat ORF were compared to ORMV and the amino acid identity was 98.11% (30K) and 98.73% (coat), respectively. A partial RMV cDNA containing the 5'-end and part of the replicase was isolated by RT-PCR from RMV RNA using oligonucleotides RGMV1 5'-GAT GGC GCC TTA ATA CGA CTC ACT ATA GTT TTA TTT TTG TTG CAA CAA CAA CAA C-3' (upstream) (SEQ ID NO: 10) and RGR 132 5'-CTT GTG CCC TTC ATG ACG AGC TAT ATC ACG-3' (downstream) (SEQ ID NO: 11). The RMV cDNA was characterized by dideoxy nucleotide sequencing. The partial nucleotide sequence containing the T7 RNA polymerase promoter and part of the RMV cDNA is as follows:

[0159] 5'-ccttaatacgactcactataGTTT-TATTTTTGTTGCAACAACAACAACAAAT-TACAATAACAAC AAAACAAATACAAACAA-CAACAACATGGCACAATTTCAACAAACAGTA AACATGCA AACATTGCAGGCTGCCG-CAGGGCGCAA-

CAGCCTGGTGAATGATTTAGCCTCACGAC GTGTTTATGACAATGCTGTCGAG-GAGCTAAATGCACGCTCGAGACGC-CCTAAGGTTC ATTACTCCAAATCAGTGTC-TACGGAACAGACGCTGTTAGCTTCAAACG CTTATCCGG AGTTTGAGATTTCCTTTACT-CATACCCAACATGCCGTACACTCCCT-TGCGGGTGGCCT AAGGACTCTTGAGTTA-GAGTATCTCATGATGCAAGTTCCGTTCGGTT CTCTGACGTA CGACATCGGTGGTAACTTTG-CAGCGCACCTTTTCAAAGGACGCGAC-TACGTTCACTG CTGTATGCCAAACTTGGAT-GTACGTGATATAGCT-3' (SEQ ID NO: 12). The uppercase letters are nucleotide sequences from RMV cDNA. The lower case letters are nucleotide sequences from T7 RNA polymerase promoter. The nucleotide sequences from the 5' and 3' oligonucleotides are underlined.

[0160] Full length infectious RMV cDNA clones were obtained by RT-PCR from RMV RNA using oligonucleotides RGMV1, 5'-GAT GGC GCC TTA ATA CGA CTC ACT ATA GTT TTA TTT TTG TTG CAA CAA CAA CAA C-3' (upstream) (SEQ ID NO: 13) and RG1 APE, 5'-ATC GTT TAA ACT GGG CCC CTA CCC GGG GTT AGG GAG G-3' (downstream) (SEQ ID NO: 14). The RMV cDNA was characterized by dideoxy nucleotide sequencing. The partial nucleotide sequence containing the T7 RNA polymerase promoter and part of the RMV cDNA is as follows:

5'-CCTTAATACGACTCACTATAGTTTTATTTTTGTTGCAACAACAACAA

CAAATTACAATAACAACAAAACAAATACAAACAACAACAACATGGCACAA

TTTCAACAAACAGTAAACATGCAAACATTCCAGGCTGCCGCAGGGCGCAA

CAGCCTGGTGAATGATTTAGCCTCACGACGTGTTTATGACAATGCTGTCG

AGGAGCTAAATGCACGCTCGAGACGCCCTAAGGTTCATTACTCCAAATCA

GTGTCTACGGAACAGACGCTGTTAGCTTCAAACGCTTATCCGGAGTTTGA

GATTTCCTTTACTCATACCCAAACATGCCGTACACTCCCTTGCGGGTGGC

CTAAGGACTCTTGAGTTAGAGTATCTCATGATGCAAGTTCCGTTCGGTTC

TCTGACGTACGACATCGGTGGTAACTTTGCAGCGCACCTTTTCAAAGGAC

GCGACTACGTTCACTGCTGTATGCCAAACTTGGATGTACTGATATAGCT-

3'

(SEQ. ID NO: 15).

[0161] The uppercase letters are nucleotide sequences from RMV cDNA. The nucleotide sequences from the 5' and 3' oligonucleotides are underlined. Full length infectious RMV cDNA clones were obtained by RT-PCR from RMV RNA using oligonucleotides RGMV1, 5'-gat ggc gcc tta ata cga ctc act ata gtt tta ttt ttg ttg caa caa caa caa c-3' (upstream) (SEQ ID NO: 16) and RG1 APE, 5'-ATC GTT TAA ACT GGG CCC CTA CCC GGG GTT AGG GAG G-3' (downstream) (SEQ ID NO: 17).

EXAMPLE 6

[0162] *Arabidopsis thaliana* cDNA Library Construction in a Dual Subgenomic Promoter Vector

[0163] Arabidopsis thaliana cDNA libraries obtained from the Arabidopsis Biological Resource Center (ABRC). The four libraries from ABRC were size-fractionated with inserts of 0.5-1 kb (CD4-13), 1-2 kb (CD4-14), 2-3 kb (CD4-15), and 3-6 kb (CD4-16). All libraries are of high quality and have been used by several dozen groups to isolate genes. The pBluescript® phagemids from the Lambda ZAP II vector were subjected to mass excision and the libraries were recovered as plasmids according to standard procedures.

[0164] Alternatively, the cDNA inserts in the CD4-13 (Lambda ZAP II vector) were recovered by digestion with NotI. Digestion with NotI in most cases liberated the entire *Arabidopsis thaliana* cDNA insert because the original library was assembled with NotI adapters. NotI is an 8-base cutter that infrequently cleaves plant DNA. In order to insert the NotI fragments into a transcription plasmid, the pBS735 transcription plasmid (**FIG. 4**) was digested with PacI/XhoI and ligated to an adapter DNA sequence created from the oligonucleotides 5'-TCGAGCGGCCGCAT-3' (SEQ ID NO: 18) and 5'-GCGGCCGC-3' (SEQ ID NO: 19). The resulting plasmid pBS740 (**FIG. 5**) contains a unique NotI restriction site for bidirectional insertion of NotI fragments from the CD4-13 library. Recovered colonies were prepared from these for plasmid minipreps with a Qiagen BioRobot 9600®. The plasmid DNA preps performed on the BioRobot 9600® are done in 96-well format and yield transcription quality DNA. An Arabidopsis cDNA library was transformed into the plasmid and analyzed by agarose gel electrophoresis to identify clones with inserts. Clones with inserts are transcribed in vitro and inoculated onto *N. benthamiana* or *Arabidopsis thaliana*. Selected leaf disks from transfected plants are then taken for biochemical analysis.

EXAMPLE 7

[0165] High Throughput Robotics

[0166] The efficiency of inoculation of subject organisms such as plants is improved by using means of high throughput robotics. For example, host plants such as *Arabidopsis thaliana* were grown in microtiter plates such as the standard 96-well and 384-well microtiter plates. A robotic handling arm then moved the plates containing the organism to a colony picker or other robot that delivered inoculations to each plant in the well. By this procedure, inoculation was performed in a very high speed and high throughput manner. It is preferable that the plant is a germinating seed or at least in the development cycle to enable access to the cells to be transfected. Equipment used for automated robotic production line include, but not be limited to, robots of these types: electronic multichannel pipetmen, Qiagen BioRobot 9600®, Robbins Hydra liquid handler, Flexys Colony Picker, New Brunswick automated plate pourer, GeneMachines HiGro shaker incubator, New Brunswick floor shaker, three Qiagen BioRobots, MJ Research PCR machines (PTC-200, Tetrad), ABI 377 sequencer and Tecan Genesis RSP200 liquid handler.

EXAMPLE 8

[0167] Genomic DNA Library Construction in a Recombinant Viral Nucleic Acid Vector

[0168] Genomic DNAs represented in BAC (bacterial artificial chromosome) or YAC (yeast artificial chromosome) libraries are obtained from the Arabidopsis Biological

Resource Center (ABRC). The BAC/YAC DNAs are mechanically size-fractionated, ligated to adapters with cohesive ends, and shotgun-cloned into recombinant viral nucleic acid vectors. Alternatively, mechanically size-fractionated genomic DNAs are blunt-end ligated into a recombinant viral nucleic acid vector. Recovered colonies are prepared for plasmid minipreps with a Qiagen BioRobot 9600®. The plasmid DNA preps done on the BioRobot 9600® are assembled in 96-well format and yield transcription quality DNA. The recombinant viral nucleic acid/Arabidopsis genomic DNA library is analyzed by agarose gel electrophoresis (template quality control step) to identify clones with inserts. Clones with inserts are then transcribed in vitro and inoculated onto *N. benthamiana* and/or *Arabidopsis thaliana.* Selected leaf disks from transfected plants are then be taken for biochemical analysis.

[0169] Genomic DNA from Arabidopsis typically contains a gene every 2.5 kb (kilobases) on average. Genomic DNA fragments of 0.5 to 2.5 kb obtained by random shearing of DNA were shotgun assembled in a recombinant viral nucleic acid expression/knockout vector library. Given a genome size of Arabidopsis of approximately 120,000 kb, a random recombinant viral nucleic acid genomic DNA library would need to contain minimally 48,000 independent inserts of 2.5 kb in size to achieve IX coverage of the Arabidopsis genome. Alternatively, a random recombinant viral nucleic acid genomic DNA library would need to contain minimally 240,000 independent inserts of 0.5 kb in size to achieve 1X coverage of the Arabidopsis genome. Assembling recombinant viral nucleic acid expression/knockout vector libraries from genomic DNA rather than cDNA has the potential to overcome known difficulties encountered when attempting to clone rare, low-abundance mRNA's in a cDNA library. A recombinant viral nucleic acid expression/knockout vector library made with genomic DNA would be especially useful as a gene silencing knockout library. In addition, the Dual Heterologous Subgenomic Promoter Expression System (DHSPES) expression/knockout vector library made with genomic DNA would be especially useful for expression of genes lacking introns. Furthermore, other plant species with moderate to small genomes (e.g. rose, approximately 80,000 kb) would be especially useful for recombinant viral nucleic acid expression/knockout vector libraries made with genomic DNA. A recombinant viral nucleic acid expression/knockout vector library can be made from existing BAC/YAC genomic DNA or from newly-prepared genomic DNAs for any plant species.

EXAMPLE 9

[0170] Genomic DNA or cDNA Library Construction in a DHSPES Vector, and Transfection of Individual Clones From Said Vector Library Onto T-DNA Tagged or Transposon Tagged or Mutated Plants

[0171] Genomic DNA or cDNA library construction in a recombinant viral nucleic acid vector, and transfection of individual clones from the vector library onto T-DNA tagged or transposon tagged or mutated plants may be performed according to the procedure set forth in Example 7. Such a protocol may be easily designed to complement mutations introduced by random insertional mutagenesis of T-DNA sequences or transposon sequences.

EXAMPLE 10

[0172] Identification of Nucleotide Sequences Involved in the Regulation of Plant Growth By Cytoplasmic Inhibition of Gene Expression Using Viral Derived RNA (GTP Binding Proteins)

[0173] In the following examples, we show: (1) a method for producing antisense RNA using an RNA viral vector, (2) a method to produce viral-derived antisense RNA in the cytoplasm, (3) a method to inhibit the expression of endogenous plant proteins in the cytoplasm by viral antisense RNA, and (4) a method to produce transfected plants containing viral antisense RNA, such method is much faster than the time required to obtain genetically engineered antisense transgenic plants. Systemic infection and expression of viral antisense RNA occurs as short as four days post inoculation, whereas it takes several months or longer to create a single transgenic plant. These examples demonstrates that novel positive strand viral vectors, which replicate solely in the cytoplasm, can be used to identify genes involved in the regulation of plant growth by inhibiting the expression of specific endogenous genes. These examples enable one to characterize specific genes and biochemical pathways in transfected plants using an RNA viral vector.

[0174] Tobamoviral vectors have been developed for the heterologous expression of uncharacterized nucleotide sequences in transfected plants. A partial *Arabidopsis thaliana* cDNA library was placed under the transcriptional control of a tobamovirus subgenomic promoter in a RNA viral vector. Colonies from transformed *E. coli* were automatically picked using a Flexys robot and transferred to a 96 well flat bottom block containing terrific broth (TB) Amp 50 ug/ml. Approximately 2000 plasmid DNAs were isolated from overnight cultures using a BioRobot and infectious RNAs from 430 independent clones were directly applied to plants. One to two weeks after inoculation, transfected *Nicotiana benthamiana* plants were visually monitored for changes in growth rates, morphology, and color. One set of plants transfected with 740 AT #120 were severely stunted. DNA sequence analysis revealed that this clone contained an Arabidopsis GTP binding protein open reading frame (ORF) in the antisense orientation. This demonstrates that an episomal RNA viral vector can be used to deliberately alter the metabolic pathway and cause plant stunting. In addition, our results suggest that the Arabidopsis antisense transcript can turn off the expression of the *N. benthamiana* gene.

[0175] Construction of an *Arabidopsis thaliana* cDNA Library in an RNA Viral Vector

[0176] An *Arabidopsis thaliana* CD4-13 cDNA library was digested with NotI. DNA fragments between 500 and 1000 bp were isolated by trough elution and subcloned into the NotI site of pBS740. *E. coli* C600 competent cells were transformed with the pBS740 AT library and colonies containing Arabidopsis cDNA sequences were selected on LB Amp 50 ug/ml. Recombinant C600 cells were automatically picked using a Flexys robot and then transferred to a 96 well flat bottom block containing terrific broth (TB) Amp 50 ug/ml. Approximately 2000 plasmid DNAs were isolated from overnight cultures using a BioRobot (Qiagen) and infectious RNAs from 430 independent clones were directly applied to plants.

[0177] Isolation of a Gene Encoding a GTP Binding Protein

[0178] One to two weeks after inoculation, transfected *Nicotiana benthamiana* plants were visually monitored for changes in growth rates, morphology, and color. Plants transfected with 740 AT #120 (**FIG. 6**) were severely stunted. Plasmid 740 AT #120 contains the TMV-U1 126-, 183-, and 30-kDa ORFs, the TMV-U5 coat protein gene (U5 cp), the T7 promoter, an Arabidopsis thaliana CD4-13 NotI fragment, and part of the pUC19 plasmid. The TMV-U1 subgenomic promoter located within the minus strand of the 30-kDa ORF controls the synthesis of the CD4-13 antisense subgenomic RNA.

[0179] DNA Sequencing and Computer Analysis

[0180] A 782 bp NotI fragment of 740 AT #120 containing the ADP-ribosylation factor (ARF) cDNA was characterized. DNA sequence of NotI fragment of 740 AT #120 (774 base pairs) is as follows:

```
5'-CCGAAACATTCTTCGTAGTGAAGCAAAATGGGGTTGAGTTTCGCCAA
GCTGTTTAGCAGGCTTTTTGCCAAGAAGGAGATGCGAATTCTGATGGTTG
GTCTTGATGCTGCTGGTAAGACCACAATCTTGTACAAGCTCAAGCTCGGA
GAGATTGTCACCACCATCCCTACTATTGGTTTCAATGTGGAAACTGTGGA
ATACAAGAACATTAGTTTCACCGTGTGGGATGTCGGGGGTCAGGACAAGA
TCCGTCCCTTGTGAGACACTACTTCCAGAACACTCAAGGTCTAATCTTTG
TTGTTGATAGCAATGACAGAGACAGAGTTGTTGAGGCTCGAGATGAACTC
CACAGGATGCTGAATGAGGACGAGCTGCGTGATGCTGTGTTGCTTGTGTT
TGCCAACAAGCAAGATCTTCCAAATGCTATGAACGCTGCTGAAATCACAG
ATAAGCTTGGCCTTCACTCCCTCCGTCAGCGTCATTGGTATATCCAGAGC
ACATGTGCCACTTCAGGTGAAGGGCTTTATGAAGGTCTGGACTGGCTCTC
CAACAACATCGCTGGCAAGGCATGATGAGGGAGAAATTGCGTTGCATCGA
GATGATTCTGTCTGCTGTGTTGGGATCTCTCTCTGTCTTGATGCAAGAGA
GATTATAAATATTATCTGAACCTTTTTGCTTTTTTGGGTATGTGAATGTT
TCTTATTGTGCAAGTAGATGGTCTTGTACCTAAAAATTTACTAGAAGAAC
CCTTTTAAATAGCTTTCGTGTATTGT-3'
(SEQ. ID NO: 20).
```

[0181] The nucleotide sequencing of 740 AT #120 was carried out by dideoxy termination using double stranded templates (Sanger et al. 1977). Nucleotide sequence analysis and amino acid sequence comparisons were performed using DNA Strider, PCGENE and NCBI Blast programs. 740 AT #120 contained an open reading frame (ORF) in the antisense orientation that encodes a protein of 181 amino acids with an apparent molecular weight of 20,579 Daltons. **FIG. 7** shows a nucleotide sequence comparison of *A. thaliana* 740 AT #120 and *A. thaliana* est AA042085, SEQ ID NOs: 21 and 22. The nucleotide sequence from 740 AT #120 is also compared with a rice (*Oryza sativa*) ADP ribosylation factor AF012896, SEQ ID NOs: 23 and 24 (**FIG. 8**); which shows 82% (456/550) positives and identities. The nucleotide sequence from 740 AT #120 exhibits a high degree of homology (81-84% identity and positive) to rice, barley, carrot, corn and *A. thaliana* DNA encoding ARFs (Table 1). The amino acid sequence derived from 740 AT #120 exhibits an even higher degree of homology (96-98% identity and 97-98% positive) to ARFs from rice, carrot, corn and *A. thaliana*. (Table 2).

[0182] The protein encoded by 740 AT #120, 120P, contained three conserved domains: the phosphate binding loop motif, GLDAAGKT, SEQ ID NO:25 (consensus GXXXXGKS/T); the G'motif, DVGGQ, SEQ ID NO:25 (consensus DXXGQ), a sequence which interacts with the gamma-phosphate of GTP; and the G motif NKQD, SEQ ID NO:26 (consensus NKXD), which is specific for guanidinyl binding. The 120P contains a putative glycine-myristoylation site at position 2, a potential N-glycosylation site (NXS) at position 60, and several putative serine/threonine phosphorylations sites.

TABLE 1

740 AT #120 Nucleotide sequence comparison

	Score	Expect	Identities	Positives
barley E10542	540.8 bits (1957)	1.4e-157	461/548 (84%)	461/548 (84%)
A. thaliana M95166	538.5 bits (1949)	7.4e-157	461/550 (83%)	461/550 (83%)
rice AF012896	537.7 bits (1946)	1.3e-156	462/553 (83%)	462/553 (83%)
carrot D45420	531.4 bits (1923)	9.8e-155	471/579 (81%)	471/579 (81%)
corn X80042	512.3 bits (1854)	6.8e-149	450/549 (81%)	450/549 (81%)

[0183]

TABLE 2

Amino acid sequence comparison of 740 AT #120 with ARFs from other organisms.

	Score	Expect	Identities	Positives
A. thaliana ARF1 g543841	365 bits (928)	e-101	179/181 (98%)	179/181 (98%)
rice g1703380	363 bits (921)	e-100	177/181 (97%)	179/181 (98%)
corn G1351974	356 bits (905)	3e-98	174/181 (96%)	179/181 (98%)
carrot g1703375	362 bits (919)	e-100	177/181 (97%)	178/181 (97%)

EXAMPLE 11

[0184] Isolation of an *Arabidopsis thaliana* ARF Genomic Clone

[0185] A genomic clone encoding *A. thaliana* ARF can be isolated by probing filters containing *A. thaliana* BAC clones using a 32p labeled 740 AT #120 NotI insert. Other members of the *A. thaliana* ARF multigene family have been identified using programs of the University of Wisconsin Genetic Computer Group. The BAC clone T0813 located on chromosome II has a high degree of homology to 740 AT #120 (78% to 86% identity).

EXAMPLE 12

[0186] Construction of a *Nicotiana benthamiana* cDNA Library

[0187] Vegetative *N. benthamiana* plants were harvested 3.3 weeks after sowing and sliced up into three groups of tissue: leaves, stems and roots. Each group of tissue was flash frozen in liquid nitrogen and total RNA was isolated from each group separately using the following hot borate method (Larry Smart and Thea Wilkins, 1995). Frozen tissue was ground to a fine powder with a pre-chilled mortar and pestle, and then further homogenized in pre-chilled glass tissue grinder. Immediately thereafter, 2.5 ml/g tissue hot (~82° C.) XT Buffer (0.2M borate decahydrate, 30 mM EGTA, 1 % (w/v) SDS. Adjusted pH to 9.0 with 5N NaOH, treated with 0.1% DEPC and autoclaved. Before use, added 1% deoxycholate (sodium salt), 10 mM dithiothreitol, 15 Nonidet P-40 (NP-40) and 2% (w/v) polyvinylpyrrilidone, MW 40,000 (PVP-40)) was added to the ground tissue. The tissue was homogenized 1-2 minutes and quickly decanted to a pre-chilled Oak Ridge centrifuge tube containing 105 μl of 20 mg/ml proteinase K in DEPC treated water. The tissue grinder was rinsed with an additional 1 ml hot XT Buffer per g tissue, which was then added to rest of the homogenate. The homogenate was incubated at 42° C. at 100 rpm for 1.5 h. 2M KCl was added to the homogenate to a final concentration of 160 mM, and the mixture was incubated on ice for 1 h to precipitate out proteins. The homogenate was centrifuged at 12,000×g for 20 min at 4° C., and the supernatant was filtered through sterile miracloth into a clean 50 ml Oak Ridge centrifuge tube. 8M LiCl was added to a final concentration of 2M LiCl and incubated on ice overnight. Precipitated RNA was collected by centrifugation at 12,000×g for 20 min at 4° C. The pellet was washed three times in 3-5 ml 4° C. 2M LiCl. Each time the pellet was resuspended with a glass rod and then spun at 12,000×g for 20 min at 4° C. The RNA pellet was suspended in 2 ml 10 mM Tris-HCl (pH 7.5), and purified from insoluble cellular components by spinning at 12,000×g for 20 min at 4° C. The RNA containing supernatant was transferred to a 15 ml Corex tube and precipitated overnight at −20° C. with 2.5 volumes of 100 % ethanol. The RNA was pelleted by centrifugation at 9,800×g for 30 min at 4° C. The RNA pellet was washed in 1-2 ml cold 70° C. ethanol and centrifuged at 9,800×g for 5 min at 4° C. Residual ethanol was removed from the RNA pellet under vacuum, and the RNA was resuspended in 200 μl DEPC treated dd-water and transferred to a 1.5 ml microfuge tube. The Corex tube was rinsed in 100 μl DEPC-treated dd-water, which was then added to the rest of the RNA. The RNA was then precipitated with 1/10 volume of 3M sodium acetate, pH 6.0 and 2.5 volumes of cold 100% ethanol at −20° C. for 1-2 h. The tube was centrifuged for 20 min at 16,000×g, and the RNA pellet washed with cold 70% ethanol, and centrifuged for 5 min at 16,000×g. After drying the pellet under vacuum, the RNA was resuspended in DEPC-treated water. This is the total RNA.

[0188] Messenger RNA was purified from total RNA using an Poly(A)Pure kit (Ambion, Austin Tex.), following the manufacturer's instructions. A reverse transcription reaction was used to synthesize cDNA from the mRNA template, using the Stratagene kit. The cDNA was subcloned into the bacteriophage at EcoRI/XhoI by ligating to arms using the Gigapack III Gold kit (Stratagene, La Jolla, Calif.), following the manufacturer's instructions.

EXAMPLE 13

[0189] Isolation and Characterization of a cDNA Encoding *Nicotiana benthamiana* ADP-Ribosylation Factor

[0190] A 488 bp cDNA from *N. benthamiana* stem cDNA library was isolated by polymerase chain reaction (PCR) using the following oligonucleotides: ATARFK15, 5' AAG AAG GAG ATG CGA ATT CTG ATG GT 3' (upstream-)(SEQ ID NO:27), ATARFN176, 5' ATG TTG TTG GAG AGC CAG TCC AGA CC 3' (downstream)(SEQ ID NO: 28). The vent polymerase in the reaction was inactivated using phenol/chloroform, and the PCR product was directly cloned into the HincII site in Bluescript KS+(Stratagene). The plasmid map of KS+ Nb ARF #3, which contains the *N. benthamiaca* ARF ORF in pBluescript KS+is shown in **FIG. 9**. The nucleotide sequence of *N. benthamiana* KS+Nb ARF#3, which contains partial ADP-ribosylation factor ORF, was determined by dideoxynucleotide sequencing. The nucleotide sequence from KS+ Nb ARF#3, SEQ ID NO: 29, had a strong similarity to other plant ADP-ribosylation factor sequences (82 to 87% identity at the nucleotide level). The nucleotide sequence comparison of *N. benthamiana* KS+ Nb ARF#3 and *A. thaliana* 740 AT #120 is shown in **FIG. 10** (SEQ ID NOs: 26 and 27).

[0191] A full-length cDNA encoding ARF is isolated by screening the *N. benthamiana* cDNA library by colony hybridization using a ^{32}P-labeled *N. benthamiana* KS+/Nb ARF #3 probe. Hybridization is carried out at 42° C. for 48 hours in 50% formamide, 5X SSC, 0.02M phosphate buffer, 5×Denhart's solution, and 0.1 mg/ml sheared calf thymus DNA. Filters are washed at 65° C. in 0.1×SSC, 0.1% SDS prior to autoradiography.

EXAMPLE 14

[0192] Rapid Isolation of cDNAs Encoding Rice, Barley, Corn, Soybean, and Other ADP-Ribosylation Factor

[0193] Libraries containing full-length cDNAs from rice, barley, corn, soybean and other important crops are obtained from public and private sources or can be prepared from plant mRNAs. The cDNAs are inserted in viral vectors or in small subcloning vectors such as pBluescript (Stratagene), pUC18, M13, or pBR322. Transformed bacteria (*E. coli*) are then plated on large petri plates or bioassay plates containing the appropriate media and antibiotic. Individual clones are selected using a robotic colony picker and arrayed into 96 well microtiter plates. The cultures are incubated at 37° C. until the transformed cells reach log phase. Aliquots are removed to prepare glycerol stocks for long term storage at −80° C. The remainder of the culture is used to inoculate an additional 96 well microtiter plate containing selective media and grown overnight. DNAs are isolated from the cultures and stored at −20° C. Using a robotic unit such as the Qbot (Genetix), the *E. coli* transformants or DNAs are rearrayed at high density on nylon filters or glass slides. Full-length cDNAs encoding ARFs from rice, barley, corn, soybean and other important crops are isolated by screening the various filters of slides by hybridization using a ^{32}P-labeled or fluorescent *N. benthamiana* KS+/Nb ARF #3 probe, or ^{32}P-labeled Arabidopsis 740 AT #120 NotI insert.

EXAMPLE 15

[0194] Rapid Isolation of Genomic Clones Encoding ADP Ribosylation Factor From Rice, Barley, Corn, Soybean, and Other Plant

[0195] Genomic libraries containing sequences from rice, barley, corn, soybean and other important crops are obtained from public and private sources, or are prepared from plant genomic DNAs. BAC clones containing entire plant genomes have been constructed and organized in minimal overlapping order. Individual BACs are sheared to 500-1000 bp fragments and directly cloned into viral vectors. Approximate 200-500 clones that completely cover an entire BAC will form a BAC viral vector sublibrary. These libraries can be stored as bacterial glycerol stocks at −80 C. and as DNA at −20 C. Genomic clones are identified by first probing filters of BACs with a ^{32}P-labeled or fluorescent *N. benthamiana* KS+/Nb ARF #3 probe, or ^{32}P-labeled Arabidopsis 740 AT #120 NotI insert. BACs that hybridize to the probe are selected and their corresponding BAC viral vector sublibrary is used to produce infectious RNA. Plants that are transfected with the BAC sublibrary are screened for loss of function (for example, stunted plants). The inserts from these clones or their corresponding plasmid DNAs are characterized by dideoxy sequencing. This provides a rapid method to obtain the genomic sequence for the plant ARFs.

EXAMPLE 16

[0196] Identification of Nucleotide Sequences Involved in the Regulation of Plant Development By Cytoplasmic Inhibition of Gene Expression Using Viral Derived RNA (G-Protein Coupled Receptor)

[0197] This example again demonstrates that an episomal RNA viral vector can be used to deliberately manipulate a signal transduction pathway in plants. In addition, our results suggest that the Arabidopsis antisense transcript can turn off the expression of the *N. benthamiana* gene.

[0198] A partial *Arabidopsis thaliana* cDNA library was placed under the transcriptional control of a tobamovirus subgenomic promoter in a RNA viral vector. (see EXAMPLE 9). Colonies from transformed *E. coli* were automatically picked using a Flexys robot and transferred to a 96 well flat bottom block containing terrific broth (TB) Amp 50 ug/ml. Approximately 2000 plasmid DNAs were isolated from overnight cultures using a BioRobot and infectious RNAs from 430 independent clones were directly applied to plants. One to two weeks after inoculation, transfected *Nicotiana benthamiana* plants were visually monitored for changes in growth rates, morphology, and color. One set of plants transfected with 740 AT #88 (**FIG. 11**) developed a white phenotype on the infected leaf tissue. DNA sequence analysis revealed that this clone contained an Arabidopsis G-protein coupled receptor open reading frame (ORF) in the antisense orientation.

[0199] DNA Sequencing and Computer Analysis

[0200] A 750 bp NotI fragment of 740 AT #88 containing the G-protein coupled receptor cDNA was characterized. DNA sequence of NotI fragment of 740 AT #88 (750 bp) is as follows:

[0201] 5 '-TTTCGATCTAAGGTTCGTGATCTC-CTTCTTCTCTACGAAGTTTA-CACTTTTTCTTCA AAGGAAACAATGAGC-CAGTACAATCAACCTCCCGTTGGTGTTCC TCCTCCTCAAGGT TATCCACCGGAGG-GATATCCAAAAGATGCTTATCCACCA-CAAGGATATCCTCCTCAG GGATATCCTCAG-CAAGGCTATCCACCTCAGGGATATCCTCAAC AAGGTTATCCTCAG CAAGGATATCCTCCAC-CGTACGCGCCTCAATATCCTCCACCAC-CGCAAGCATCAGCA ACAACAGAGCAAGTC-CTGGCTTTCTAGAAGGATGTCTTGCTGCTCT GTGTTGTTGCT GTCTCTTGGATGCTTGCT-TCTGATTGGAGTCTCTCTCTCTCTGCAT-AAAGCTTCGGGA TTTATTTGTAA-GAGGGTTTTTGGGTTAAACAAAAACCTTA ATTGATTTGTGGGGCATT AAAAAT-GAATCTCTCGATGATTCTCTTCGTTTAT-GTGGTAATGTTCTTCGGTTATAAC ATTTAA-CATTGCTATCGACGTTCTGCCTAGTTGGATTT GATTATTGGGAATGTAAATT GGTTGGGAA-GACACCGGGCCGTTAATGACAGAAC-CCGAACTGAGATGGAGTATGAT CTGAAATATTTAAAACAATCCTCGCGA-CATAGCCTCCAATCTCATCGTAAATATTCT TTTTAAACTATTCCCAATCTTAACTTT-TATAGTCTGGTCGACTGACCACTACTCTTTTT CCTT-3' (SEQ ID NO: 30) The nucleotide sequencing of 740 AT #88 was carried out by dideoxy termination using double stranded templates (Sanger et al., *Proc. Natl. Acad. Sci.* USA 74(12):5463-5467 (1977)). Nucleotide sequence analysis and amino acid sequence comparisons were performed using DNA Strider, PCGENE and NCBI Blast programs.

EXAMPLE 17

[0202] Identification of Nucleotide Sequences Containing an Arabidopsis S18 Ribosomal Protein Open Reading Frame

[0203] One to two weeks after inoculation, transfected *Nicotiana benthamiana* plants were visually monitored for changes in growth rates, morphology, and color. One set of plants transfected with 740 AT #377 (**FIG. 12**) were severely stunted. DNA sequence analysis (**FIG. 13**, SEQ ID NO: 31) revealed that this clone contained an Arabidopsis S18 ribosomal protein open reading frame (ORF) in the antisense orientation.

EXAMPLE 18

[0204] Identification of L19 Ribosomal Protein Gene Involved in the Regulation of Plant Growth by Cytoplasmic Inhibition of Gene Expression Using Viral Derived RNA

[0205] One to two weeks after inoculation, transfected *Nicotiana benthamiana* plants were visually monitored for changes in growth rates, morphology, and color. One set of plants transfected with 740 AT #2483 (**FIG. 14**) were severely stunted. DNA sequence analysis (**FIG. 15**, SEQ ID NO: 32) revealed that this clone contained an Arabidopsis L19 ribosomal protein open reading frame (ORF) in the antisense orientation. The 740 AT #2483 nucelotide sequence exhibited a high degree of homology (77-78% identities and positives) to plant, L19 ribosomal proteins genes (Table 3).

TABLE 3

740 AT #2483 Nucleotide sequence comparison

	Score	pValue	Identities	Positives
A. thaliana AF075597	389 (107.5 bits) 101/130 (77%)	2.60E-38	101/130 (77%)	
Rice mRNA for ribosomal protein L19 D21304 (78%)	198 (54.7 bits)	2.20E-10	50/64 (78%)	50/64
N. tabacum L19 mRNA Z31720	194 (53.6 bits)	3.50E-05	50/64 (78%)	50/64

EXAMPLE 19

[0206] Novel Requirements for Production of Infectious Viral Vector in Vitro Derived RNA Transcripts

[0207] This example demonstrates the production of highly infectious viral vector transcripts containing 5' nucleotides with reference to the virus vector.

[0208] Construction of a library of subgenomic cDNA clones of TMV and BMV has been described in Dawson et al., *Proc. Natl. Acad. Sci.* USA 83:1832-1836 (1986) and Ahlquist et al., *Proc. Natl. Acad. Sci.* USA 81:7066-7070 (1984). Nucleotides were added between the transcriptional start site of the promoter for in vitro transcription, in this case T7, and the start of the cDNA of TMV in order to maximize transcription product yield and possibly obviate the need to cap virus transcripts to insure infectivity. The relevant sequence is the T7 promoter . . . TATAG^ TATTTT . . . (SEQ ID NO: 33) where the ^ indicates the base preceding is the start site for transcription and the bold letter is the first base of the TMV cDNA. Three approaches were taken: 1) addition of G, GG or GGG between the start site of transcription and the TMV cDNA(. . . TATAGGTATTT, (SEQ ID NO: 34), . . . and associated sequences); 2) addition of G and a random base (GN or N2) or a G and two random bases (GNN or N3) between the start site of transcription and the TMV cDNA (. . . TATAGNTATTT (SEQ ID NO: 35), . . . and associated sequences), and the addition of a GT and a single random base between the start site of transcription and the TMV cDNA (. . . TATAGTNGTATTT (SEQ ID NO: 36), . . . and associated sequences). The use of random bases was based on the hypothesis that a particular base may be 2D: best suited for an additional nucleotide attached to the cDNA, since it will be complementary to the normal nontemplated base incorporated at the 3'-end of the TMV (−) strand RNA. This allows for more ready mis-initiation and restoration of wild type sequence. The GTN would allow the mimicking of two potential sites for initiation, the added and the native sequence, and facilitate more ready mis-initiation of transcription in vivo to restore the native TMV cDNA sequence. Approaches included cloning GFP expressing TMV vector sequences into vectors containing extra G, GG or GGG bases using standard molecular biology techniques. Likewise, full length PCR of TMV expression clone 1056 was done to add N2, N3 and GTN bases between the T7 promoter and the TMV cDNA. Subsequently, these PCR products were cloned into pUC based vectors. Capped and uncapped transcripts were made in vitro and inoculated to tobacco protoplasts or *Nicotiana benthamiana* plants, wild type and 30k expressing transgenics. The results are that an extra G, . . . TATAGGTATTTT (SEQ ID NO: 37), . . . , or a GTC, . . . TATAGTCGTATTTT (SEQ ID NO: 38), . . . , were found to be well tolerated as additional 5' nucleotides on the 5' of TMV vector RNA transcripts and were quite infectious on both plant types and protoplasts as capped or non-capped transcripts. Other sequences may be screened to find other options. Clearly, infectious transcripts may be derived with extra 5' nucleotides.

[0209] Other derivatives based on the putative mechanistic function of the GTN strategy that yielded the GTC functional vector are to use multiple GTN motifs preceeding the 5' most nt of the virus cDNA or the duplication of larger regions of the 5'-end of the TMV genome. For example: TATA^ GTNGTNGTATT, (SEQ ID NO: 39), . . . or TATA^ GTNGTNGTNGTNGTATT (SEQ ID NO: 40), . . . or TATA^ GTATTTGTATTT (SEQ ID NO: 41), . . . In this manner the replication mediated repair mechanism may be potentiated by the use of multiple recognition sequences at the 5'-end of transcribed RNA. The replicated progeny may exhibit the results of reversion events that would yield the wild type virus 5' virus sequence, but may include portions or entire sets of introduced additional base sequences. This strategy can be applied to a range of RNA viruses or RNA viral vectors of various genetic arrangements derived from wild type virus genome. This would require the use of sequences particular to that of the virus used as a vector.

EXAMPLE 20

[0210] Infectivity of Uncapped Transcripts

[0211] Two TMV-based virus expression vectors were initially used in these studies pBTI 1056 which contains the T7 promoter followed directly by the virus cDNA sequence (. . . TATAGTATT . . .), and pBTI SBS60-29 which contains the T7 promoter (underlined) followed by an extra guanine residue then the virus cDNA sequence (. . . TATAGGTATT . . .). Both expression vectors express the cycle 3 shuffled green fluorescent protein (GFPc3) in localized infection sites and systemically infected tissue of infected plants. Transcriptions of each plasmid were carried out in the absence of cap analogue (uncapped) or in the presence of 8-fold greater concentration of RNA cap analogue than rGTP (capped). Transcriptions were mixed with abrasive and inoculated on expanded older leaves of a wild type *Nicotiana benthamiana* (Nb) plant and a Nb plant expressing a TMV U1 30k movement protein transgene (Nb 30K). Four days post inoculation (dpi), long wave UV light was used to judge the number of infection sites on the inoculated leaves of the plants. Systemic, noninoculated tissues, were monitored from 4 dpi on for appearance of systemic infection indicating vascular movement of the inoculated virus. Table 4 shows data from one representative experiment.

TABLE 4

	Local infection sites		Systemic Infection	
Construct	Nb	Nb 30K	Nb	Nb 30K
pBTI1056				
Capped	5	6	yes	yes
Uncapped	0	5	no	yes
PBTI SBS60-29				
Capped	6	6	yes	yes
Uncapped	1	5	yes	yes

[0212] *Nicotiana tabacum* protoplasts were infected with either capped or uncapped transcriptions (as described above) of pBTI SBS60 which contains the T7 promoter followed directly by the virus cDNA sequence (TATAG-TATT . . .). This expression vector also expresses the GFPc3 gene in infected cells and tissues. Nicotiana tabacum protoplasts were transfected with 1 ml of each transcriptions. Approximately 36 hours post infection transfected protoplasts were viewed under UV illumination and cells showing GFPc3 expression. Approximately 80% cells transfected with the capped PBTI SBS60 transcripts showed GFP expression while 5% of cells transfected with uncapped transcripts showed GFP expression. These experiments were repeated with higher amounts of uncapped inoculum. In this case a higher proportion of cells, >30% were found to be infected at this time with uncapped transcripts, where >90% of cells infected with greater amounts of capped transcripts were scored infected.

[0213] These results indicate that, contrary to the practiced art in scientific literature and in issued patents (Ahlquist et al., U.S. Pat. No. 5,466,788), uncapped transcripts for virus expression vectors are infective on both plants and in plant cells, however with much lower specific infectivity. Therefore, capping is not a prerequisite for establishing an infection of a virus expression vector in plants; capping just increases the efficiency of infection. This reduced efficiency can be overcome, to some extent, by providing excess in vitro transcription product in an infection reaction for plants or plant cells.

[0214] The expression of the 30K movement protein of TMV in transgenic plants also has the unexpected effect of equalizing the relative specific infectivity of uncapped verses capped transcripts. The mechanism behind this effect is not fully understood, but could arise from the RNA binding activity of the movement protein stabilizing the uncapped transcript in infected cells from prereplication cytosolic degradation.

[0215] Extra guanine residues located between the T7 promoter and the first base of a virus cDNA lead to increased amount of RNA transcript as predicted by previous work with phage polymerases. These polymerases tend to initiate more efficiently at . . . TATAGG or . . . TATAGGG than . . . TATAG. This has an indirect effect on the relative infectivity of uncapped transcripts in that greater amounts are synthesized per reaction resulting in enhanced infectivity.

[0216] Data Concerning Cap Dependent Transcription of pBTI1056 GTN#28

[0217] TMV-based virus expression vector pBTI 1056 GTN#28 which contains the T7 promoter (underlined) followed GTC bases (bold) then the virus cDNA sequence (. . . TATAGTCGTATT SEQ ID NO: 42, . . .). This expression vector expresses the cycle 3 shuffled green fluorescent protein (GFPc3) in localized infection sites and systemically infected tissue of infected plants. This vector was transcribed in vitro in the presence (capped) and absence (uncapped) of cap analogue. Transcriptions were mixed with abrasive and inoculated on expanded older leaves of a wild type Nicotiana benthamiana (Nb) plant and a Nb plant expressing a TMV U1 30k movement protein transgene (Nb 30K). Four days post inoculation (dpi) long wave UV light was used to judge the number of infection sites on the inoculated leaves of the plants. Systemic, non-inoculated tissues, were monitored from 4 dpi on for appearance of systemic infection indicating vascular movement of the inoculated virus. Table 5 shows data from two representative experiments at 11 dpi.

TABLE 5

	Local infection sites		Systemic Infection	
Construct	Nb	Nb 30K	Nb	Nb 30K
Experiment 1				
pBTI1056 GTN#28				
Capped	18	25	yes	yes
Uncapped	2	4	yes	yes
Experiment 2				
pBTI1056 GTN#28				
Capped	8	12	yes	yes
Uncapped	3	7	yes	yes

[0218] These data further support the claims concerning the utility of uncapped transcripts to initiate infections by plant virus expression vectors and further demonstrates that the introduction of extra, non-viral nucleotides at the 5'-end of in vitro transcripts does not preclude infectivity of uncapped transcripts.

[0219] Although the invention has been described with reference to the presently preferred embodiments, it should be understood that various modifications can be made without departing from the spirit of the invention.

SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 42

<210> SEQ ID NO 1
<211> LENGTH: 26
<212> TYPE: DNA
<213> ORGANISM: Tomato mosaic virus

-continued

<400> SEQUENCE: 1

ctcgcaaagt ttcgaaccaa atcctc 26

<210> SEQ ID NO 2
<211> LENGTH: 35
<212> TYPE: DNA
<213> ORGANISM: Tomato mosaic virus

<400> SEQUENCE: 2

cggggtacct gggccccaac cgggggttcc ggggg 35

<210> SEQ ID NO 3
<211> LENGTH: 41
<212> TYPE: DNA
<213> ORGANISM: Tomato mosaic virus

<400> SEQUENCE: 3

tcctcgagcc taggctcgca aagtttcgaa ccaaatcctc a 41

<210> SEQ ID NO 4
<211> LENGTH: 35
<212> TYPE: DNA
<213> ORGANISM: Tomato mosaic virus

<400> SEQUENCE: 4

cggggtacct gggccccaac cgggggttcc ggggg 35

<210> SEQ ID NO 5
<211> LENGTH: 30
<212> TYPE: DNA
<213> ORGANISM: Tomato mosaic virus

<400> SEQUENCE: 5

tgctcgagtg tgttcttcag ttttctgtca 30

<210> SEQ ID NO 6
<211> LENGTH: 30
<212> TYPE: DNA
<213> ORGANISM: Tomato mosaic virus

<400> SEQUENCE: 6

aactcgagcg ctttgatttc tccgaagctt 30

<210> SEQ ID NO 7
<211> LENGTH: 31
<212> TYPE: DNA
<213> ORGANISM: Ribgrass mosaic virus

<400> SEQUENCE: 7

tactcgaggt tcataagacc gcggtaggcg g 31

<210> SEQ ID NO 8
<211> LENGTH: 36
<212> TYPE: DNA
<213> ORGANISM: Ribgrass mosaic virus

<400> SEQUENCE: 8

cggggtacct gggcccctac ccggggttta gggagg 36

<210> SEQ ID NO 9

-continued

<211> LENGTH: 1543
<212> TYPE: DNA
<213> ORGANISM: Ribgrass mosaic virus

<400> SEQUENCE: 9

ctcgaggttc ataagaccgc ggtaggcgga gcgtttgttt actgtagtat aattaaatat 60

ttgtcagata aaaggttgtt taaagatttg ttttttgttt gactgagtcg ataatgtctt 120

acgagcctaa agttagtgac ttccttgctc ttacgaaaaa ggaggaaatt ttacccaagg 180

ctttgacgag attaaagact gtctctatta gtactaagga tgttatatct gttaaggagt 240

ctgagtccct gtgtgatatt gatttgttag tgaatgtgcc attagataag tataggtatg 300

tgggtgtttt gggtgttgtt ttcaccggtg aatggctggt accggatttc gttaaaggtg 360

gggtaacagt gagcgtgatt gacaaacggc ttgaaaattc cagagagtgc ataattggta 420

cgtaccgagc tgctgtaaag gacagaaggt tccagttcaa gctggttcca aattacttcg 480

tatccattgc ggatgccaag cgaaaaccgt ggcaggttca tgtgcgaatt caaaatctga 540

agatcgaagc tggatggcaa cctctagctc tagaggtggt ttctgttgcc atggttacta 600

ataacgtggt tgttaaaggt ttgagggaaa aggtcatcgc agtgaatgat ccgaacgtcg 660

aaggtttcga aggtgtggtt gacgatttcg tcgattcggt tgctgcattc aaggcgattg 720

acagtttccg aaagaaaaag aaaaagattg gaggaaggga tgtaaataat aataagtata 780

gatatagacc ggagagatac gccggtcctg attcgttaca atataaagaa gaaaatggtt 840

tacaacatca cgagctcgaa tcagtaccag tatttcgcag cgatgtgggc agagcccaca 900

gcgatgctta accagtgcgt gtctgcgttg tcgcaatcgt atcaaactca ggcggcaaga 960

gatactgtta gacagcagtt ctctaacctt ctgagtgcga ttgtgacacc gaaccagcgg 1020

tttccagaaa caggataccg ggtgtatatt aattcagcag ttctaaaacc gttgtacgag 1080

tctctcatga agtcctttga tactagaaat aggatcattg aaactgaaga agagtcgcgt 1140

ccatcggctt ccgaagtatc taatgcaaca caacgtgttg atgatgcgac cgtggccatc 1200

aggagtcaaa ttcagctttt gctgaacgag ctctccaacg gacatggtct gatgaacagg 1260

gcagagttcg aggttttatt accttgggct actgcgccag ctacataggc gtggtgcaca 1320

cgatagtgca tagtgttttt ctctccactt aaatcgaaga gatatactta cggtgtaatt 1380

ccgcaagggt ggcgtaaacc aaattacgca atgttttagg ttccatttaa atcgaaacct 1440

gttatttcct ggatcacctg ttaacgtacg cgtggcgtat attacagtgg gaataactaa 1500

aagtgagagg ttcgaatcct ccctaacccc gggtaggggc cca 1543

<210> SEQ ID NO 10
<211> LENGTH: 55
<212> TYPE: DNA
<213> ORGANISM: Ribgrass mosaic virus

<400> SEQUENCE: 10

gatggcgcct taatacgact cactatagtt ttatttttgt tgcaacaaca acaac 55

<210> SEQ ID NO 11
<211> LENGTH: 30
<212> TYPE: DNA
<213> ORGANISM: Ribgrass mosaic virus

<400> SEQUENCE: 11

-continued

cttgtgccct tcatgacgag ctatatcacg 30

<210> SEQ ID NO 12
<211> LENGTH: 496
<212> TYPE: DNA
<213> ORGANISM: Ribgrass mosaic virus

<400> SEQUENCE: 12

ccttaatacg actcactata gttttatttt tgttgcaaca acaacaacaa attacaataa 60

caacaaaaca aatacaaaca acaacaacat ggcacaattt caacaaacag taaacatgca 120

aacattgcag gctgccgcag ggcgcaacag cctggtgaat gatttagcct cacgacgtgt 180

ttatgacaat gctgtcgagg agctaaatgc acgctcgaga cgccctaagg ttcattactc 240

caaatcagtg tctacggaac agacgctgtt agcttcaaac gcttatccgg agtttgagat 300

ttcctttact catacccaac atgccgtaca ctcccttgcg ggtggcctaa ggactcttga 360

gttagagtat ctcatgatgc aagttccgtt cggttctctg acgtacgaca tcggtggtaa 420

ctttgcagcg caccttttca aaggacgcga ctacgttcac tgctgtatgc caaacttgga 480

tgtacgtgat atagct 496

<210> SEQ ID NO 13
<211> LENGTH: 55
<212> TYPE: DNA
<213> ORGANISM: Ribgrass mosaic virus

<400> SEQUENCE: 13

gatggcgcct taatacgact cactatagtt ttatttttgt tgcaacaaca acaac 55

<210> SEQ ID NO 14
<211> LENGTH: 37
<212> TYPE: DNA
<213> ORGANISM: Ribgrass mosaic virus

<400> SEQUENCE: 14

atcgtttaaa ctgggcccct acccggggtt agggagg 37

<210> SEQ ID NO 15
<211> LENGTH: 497
<212> TYPE: DNA
<213> ORGANISM: Ribgrass mosaic virus

<400> SEQUENCE: 15

ccttaatacg actcactata gttttatttt tgttgcaaca acaacaacaa attacaataa 60

caacaaaaca aatacaaaca acaacaacat ggcacaattt caacaaacag taaacatgca 120

aacattccag gctgccgcag ggcgcaacag cctggtgaat gatttagcct cacgacgtgt 180

ttatgacaat gctgtcgagg agctaaatgc acgctcgaga cgccctaagg ttcattactc 240

caaatcagtg tctacggaac agacgctgtt agcttcaaac gcttatccgg agtttgagat 300

ttcctttact catacccaaa catgccgtac actcccttgc gggtggccta aggactcttg 360

agttagagta tctcatgatg caagttccgt tcggttctct gacgtacgac atcggtggta 420

actttgcagc gcacctttc aaaggacgcg actacgttca ctgctgtatg ccaaacttgg 480

atgtacgtga tatagct 497

<210> SEQ ID NO 16

-continued

<211> LENGTH: 55
<212> TYPE: DNA
<213> ORGANISM: Ribgrass mosaic virus

<400> SEQUENCE: 16

gatggcgcct taatacgact cactatagtt ttatttttgt tgcaacaaca acaac 55

<210> SEQ ID NO 17
<211> LENGTH: 37
<212> TYPE: DNA
<213> ORGANISM: Ribgrass mosaic virus

<400> SEQUENCE: 17

atcgtttaaa ctgggcccct acccggggtt agggagg 37

<210> SEQ ID NO 18
<211> LENGTH: 14
<212> TYPE: DNA
<213> ORGANISM: Arabidopsis thaliana

<400> SEQUENCE: 18

tcgagcggcc gcat 14

<210> SEQ ID NO 19
<211> LENGTH: 8
<212> TYPE: DNA
<213> ORGANISM: Arabidopsis thaliana

<400> SEQUENCE: 19

gcggccgc 8

<210> SEQ ID NO 20
<211> LENGTH: 773
<212> TYPE: DNA
<213> ORGANISM: Arabidopsis thaliana

<400> SEQUENCE: 20

ccgaaacatt cttcgtagtg aagcaaaatg gggttgagtt tcgccaagct gtttagcagg 60

ctttttgcca agaaggagat gcgaattctg atggttggtc ttgatgctgc tggtaagacc 120

acaatcttgt acaagctcaa gctcggagag attgtcacca ccatccctac tattggtttc 180

aatgtggaaa ctgtggaata caagaacatt agtttcaccg tgtgggatgt cgggggtcag 240

gacaagatcc gtcccttgtg agacactact tccagaacac tcaaggtcta atctttgttg 300

ttgatagcaa tgacagagac agagttgttg aggctcgaga tgaactccac aggatgctga 360

atgaggacga gctgcgtgat gctgtgttgc ttgtgtttgc caacaagcaa gatcttccaa 420

atgctatgaa cgctgctgaa atcacagata agcttggcct tcactccctc cgtcagcgtc 480

attggtatat ccagagcaca tgtgccactt caggtgaagg gctttatgaa ggtctggact 540

ggctctccaa caacatcgct ggcaaggcat gatgagggag aaattgcgtt gcatcgagat 600

gattctgtct gctgtgttgg gatctctctc tgtcttgatg caagagagat tataaatatt 660

atctgaacct ttttgctttt ttgggtatgt gaatgtttct tattgtgcaa gtagatggtc 720

ttgtacctaa aaatttacta gaagaaccct tttaaatagc tttcgtgtat tgt 773

<210> SEQ ID NO 21
<211> LENGTH: 404
<212> TYPE: DNA

-continued

<213> ORGANISM: Arabidopsis thaliana

<400> SEQUENCE: 21

tccgaaacat tcttcgtact gaagcaaaat ggggttgagt ttcgccaagc tgtttagcag 60

gctttttgcc aagaaggaga tgcgaattct gatggttggt cttgatgctg ctggtaagac 120

cacaatcttg tacaagctca agctcggaga gattgtcacc accatccctt actattggtt 180

tcaatgtgga aactgtggaa tacaagaaca ttagtttcac cgtgtggatg tcgggggtca 240

ggacaagatc cgtccgtccc ttgtggagac actacttcca gaacactcaa ggtctaatct 300

ttgttgttga tagcaatgac agagacagag ttgttgaggc tcgagatgaa ctccacagga 360

tgctgaatga ggacgagctg cgtgatgctg tgttgcttgt gttt 404

<210> SEQ ID NO 22
<211> LENGTH: 400
<212> TYPE: DNA
<213> ORGANISM: Arabidopsis thaliana
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1)...(400)
<223> OTHER INFORMATION: n= a, t, c or g

<400> SEQUENCE: 22

tccgaaacat tcttcgtagt gaagcaaaat ggggttgagt ttcgccaagc tgtttagcag 60

gctttttgcc aagaaggaga tgcgaattct gatggttggt cttgatgctg ctggtaagac 120

cacaatgttg tacaagctca agctcggaga gattgtcacc accatcccta ctattggttt 180

caatgtggaa actgtggaat acaagaacat tagtttcacc gtgtgggatg tcgggggtca 240

ggacaagatc cgtcccttgt ggagacacta cttccagaac actcaaggtc taatctttgt 300

tgttgatagc aatgacagag acagagttgt tgaggctcga gatgaactcc acaggatgct 360

gnatgagnac gagctgcgtg atgctgtgtt gcttgtgttt 400

<210> SEQ ID NO 23
<211> LENGTH: 550
<212> TYPE: DNA
<213> ORGANISM: Arabidopsis thaliana

<400> SEQUENCE: 23

aaatggggtt gagtttcgcc aagctgttta gcaggctttt tgccaagaag gagatgcgaa 60

ttctgatggt tggtcttgat gctgctggta agaccacaat cttgtacaag ctcaagctcg 120

gagagattgt caccaccatc cctactattg gtttcaatgt ggaaactgtg gaatacaaga 180

acattagttt caccgtgtgg gatgtcgggg gtcaggacaa gatccgtccc ttgtggagac 240

actacttcca gaacactcaa ggtctaatct ttgttgttga tagcaatgac agagacagag 300

ttgttgaggc tcgagatgaa ctccacagga tgctgaatga ggacgagctg cgtgatgctg 360

tgttgcttgt gtttgccaac aagcaagatc ttccaaatgc tatgaacgct gctgaaatca 420

cagataagct tggccttcac tccctccgtc agcgtcattg gtatatccag agcacatgtg 480

ccacttcagg tgaagggctt tatgaaggtc tggactggct ctccaacaac atcgctggca 540

aggcatgatg 550

<210> SEQ ID NO 24
<211> LENGTH: 550
<212> TYPE: DNA

-continued

<213> ORGANISM: Oryza sativa

<400> SEQUENCE: 24

agatggggct cacgttcacg aagctgttca gccgcctctt cgccaagaag gagatgagga 60

tcctcatggt cggtctcgat gcggccggta aaaccaccat cctctacaag ctcaagctcg 120

gcgagatcgt caccactatc cccaccatcg gttttaatgt cgaaactgtt gagtacaaga 180

acattagctt caccgtttgg gatgttggtg gtcaggacaa gatcaggccc ctgtggaggc 240

actatttcca gaacacccag ggcctcattt ttgttgtgga cagcaatgac agagagcgtg 300

ttgttgaggc cagggatgag ctccaccgta tgctgaatga ggatgagcta cgtgatgctg 360

tgctgctggt gtttgcaaac aaacaagatc ttcctaatgc catgaacgct gctgagatca 420

ccgacaagct tggtctgcac tccttgcgcc agcggcactg gtacatccag agcacatgtg 480

ctacctctgg tgaggggttg tatgaggggc ttgactggct ttccaacaac attgccaaca 540

aggcttgaag 550

<210> SEQ ID NO 25
<211> LENGTH: 8
<212> TYPE: PRT
<213> ORGANISM: Arabidopsis thaliana

<400> SEQUENCE: 25

Gly Leu Asp Ala Ala Gly Lys Thr
1 5

<210> SEQ ID NO 26
<211> LENGTH: 389
<212> TYPE: DNA
<213> ORGANISM: Arabidopsis thaliana

<400> SEQUENCE: 26

tggtcttgat gctgctggta agaccacaat cttgtacaag ctcaagctcg gagagattgt 60

caccaccatc cctactattg gtttcaatgt ggaaactgtg gaatacaaga acattagttt 120

caccgtggga tgtcgggggt caggacaaga tccgtccctt gtggagacac tacttccaga 180

acactcaagg tctaatcttt gttgttgata gcaatgacag agacagagtt gttgaggctc 240

gagatgaact ccacaggatg ctgaatgagg acgagctgcg tgatgctgtg ttgcttgtgt 300

ttgccaacaa gcaagatctt ccaaatgcta tgaacgctgc tgaaatcaca gataagcttg 360

gccttcactc cctccgtcag cgtcattgg 389

<210> SEQ ID NO 27
<211> LENGTH: 391
<212> TYPE: DNA
<213> ORGANISM: N. benthiamiana

<400> SEQUENCE: 27

cggtcttgat gcagctggta aaaccaccat attgtacaag ctcaagctgg gagagatagt 60

taccactatt cctaccattg gattcaatgt ggagactgtt gaatacaaga acataagctt 120

cacggtctgg gatgttggtg gtcaggacaa gatccgacca ttgtggaggc attacttcca 180

aaacacacaa ggacttatct ttgtggtcga tagtaatgat cgtgatcgtg ttgttgaggc 240

tagagatgag ctgcaccgga tgttgaatga ggatgaactg agggatgctg tgctgcttgt 300

gtttgctaac aagcaagatc ttccaaatgc tatgaatgct gctgagatta ctgacaagct 360

-continued

tggtcttcat tctctccgtc aacgtcactg g 391

<210> SEQ ID NO 28
<211> LENGTH: 26
<212> TYPE: DNA
<213> ORGANISM: N. benthiamiana

<400> SEQUENCE: 28

aagaaggaga tgcgaattct gatggt 26

<210> SEQ ID NO 29
<211> LENGTH: 26
<212> TYPE: DNA
<213> ORGANISM: N. benthiamiana

<400> SEQUENCE: 29

atgttgttgg agagccagtc cagacc 26

<210> SEQ ID NO 30
<211> LENGTH: 750
<212> TYPE: DNA
<213> ORGANISM: Arabidopsis thaliana

<400> SEQUENCE: 30

tttcgatcta aggttcgtga tctccttctt ctctacgaag tttacacttt ttcttcaaag 60

gaaacaatga gccagtacaa tcaacctccc gttggtgttc ctcctcctca aggttatcca 120

ccggagggat atccaaaaga tgcttatcca ccacaaggat atcctcctca gggatatcct 180

cagcaaggct atccacctca gggatatcct caacaaggtt atcctcagca aggatatcct 240

ccaccgtacg cgcctcaata tcctccacca ccgcaagcat cagcaacaac agagcaagtc 300

ctggctttct agaaggatgt cttgctgctc tgtgttgttg ctgtctcttg gatgcttgct 360

tctgattgga gtctctctct ctctgcataa agcttcggga tttatttgta agagggtttt 420

tgggttaaac aaaaacctta attgatttgt ggggcattaa aaatgaatct ctcgatgatt 480

ctcttcgttt atgtggtaat gttcttcggt tataacattt aacattgcta tcgacgttct 540

gcctagttgg atttgattat tgggaatgta aattggttgg gaagacaccg ggccgttaat 600

gacagaaccc gaactgagat ggagtatgat ctgaaatatt taaaacaatc ctcgcgacat 660

agcctccaat ctcatcgtaa atattctttt taaactattc ccaatcttaa cttttatagt 720

ctggtcgact gaccactact ctttttcctt 750

<210> SEQ ID NO 31
<211> LENGTH: 543
<212> TYPE: DNA
<213> ORGANISM: Arabidopsis thaliana

<400> SEQUENCE: 31

cttaaaagca atatgacagt agagaagatc tctcacaaaa gacccaaaat cgagtcgtgc 60

aaaattgtac gaacaacaaa atttaaaatt cagtccttat caaagatcca atccagctgc 120

aactagcaac attggcttaa cgcttcttag acacaccaac agtctttcct ctgcgaccag 180

ttgtcttggt gtgttgtcca cgaacacgga gaccccagta atgtctcaga ccacgatggt 240

ttctgatttt cttgagacgc tcaagatcat ccctgagctt catgtcaagg gcattggaga 300

caacttgaga gtacttccca tccttgtaat ctttctgtct gttcaaaaac cagtctggaa 360

-continued

tcttgaactg tcttgggttt gcaacaatag tcatgaggtt gtcaatctca gctgcagata 420

actcaccagc cctcttgttc atgtcgacat cggctttctt gcagacaatg ttggccaatc 480

tccttccaat acctttgata gaggtaaggg caaacataat cttttgctta ccatcaacgg 540

tag 543

<210> SEQ ID NO 32
<211> LENGTH: 757
<212> TYPE: DNA
<213> ORGANISM: Arabidopsis thaliana

<400> SEQUENCE: 32

ctgacataag ttatgttctt tgcgaaaata aaagttattc cacaaacgca ttcgataaaa 60

cattcaaaac cttcttcaga gtctaatccg tgaactgatg atcgatatag cttcacacta 120

tatatcctct tcacttctta gacttcttct tcggtacagc tgcagttgga gcaggtgtag 180

cagcaggtgc tggagcagct acaggcgcaa catctccacc gggaccctta gctaaacgct 240

cctctctcct agcatgcttc ctttctcggc tagccttgtt cttcgccctc ttagcctcaa 300

actgatcaag acagagtctt ctccctagcc ttctcaagcc tttgacttgt ggatactctc 360

catcaagaca cgcttgttct tgaacacatt tacccttaac acgcatgtac atggtcatgg 420

tacatgtgct tgtcaatctt ctttcgtctc tctggatttc ttcaaacaga cgcctaagaa 480

acacgccttc ctacgcattc cacagtacct ttggttggga acctaactta cgggtacccc 540

ttccttttaa ccgattccag agtggcgacc ctttatcttg gcaatcttca ttttgcgagc 600

ttggaacaag agtgaatctt gggtggcttc tgatgatgaa acctctttaa actttctgag 660

gttttggcgg aaatggctga aacggatttg tgggaccaac caaattgcct ttcggcttaa 720

tactgatgcg accgtttgag taaaaaccgc cttcagg 757

<210> SEQ ID NO 33
<211> LENGTH: 11
<212> TYPE: DNA
<213> ORGANISM: Tobacco mosaic virus

<400> SEQUENCE: 33

tatagtattt t 11

<210> SEQ ID NO 34
<211> LENGTH: 11
<212> TYPE: DNA
<213> ORGANISM: Tobacco mosaic virus

<400> SEQUENCE: 34

tataggtatt t 11

<210> SEQ ID NO 35
<211> LENGTH: 11
<212> TYPE: DNA
<213> ORGANISM: Tobacco mosaic virus
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1)...(11)
<223> OTHER INFORMATION: n= a, t, c, or g

<400> SEQUENCE: 35

tatagntatt t 11

-continued

<210> SEQ ID NO 36
<211> LENGTH: 13
<212> TYPE: DNA
<213> ORGANISM: Tobacco mosaic virus
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1)...(12)
<223> OTHER INFORMATION: n= a, t, c or g

<400> SEQUENCE: 36

tatagtngta ttt 13

<210> SEQ ID NO 37
<211> LENGTH: 12
<212> TYPE: DNA
<213> ORGANISM: Tobacco mosaic virus

<400> SEQUENCE: 37

tataggtatt tt 12

<210> SEQ ID NO 38
<211> LENGTH: 14
<212> TYPE: DNA
<213> ORGANISM: Tobacco mosaic virus

<400> SEQUENCE: 38

tatagtcgta tttt 14

<210> SEQ ID NO 39
<211> LENGTH: 15
<212> TYPE: DNA
<213> ORGANISM: Tobacco mosaic virus
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1)...(15)
<223> OTHER INFORMATION: n= a. t, c or g

<400> SEQUENCE: 39

tatagtngtn gtatt 15

<210> SEQ ID NO 40
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Tobacco mosaic virus
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1)...(21)
<223> OTHER INFORMATION: n= a, t, c or g

<400> SEQUENCE: 40

tatagtngtn gtngtngtat t 21

<210> SEQ ID NO 41
<211> LENGTH: 16
<212> TYPE: DNA
<213> ORGANISM: Tobacco mosaic virus

<400> SEQUENCE: 41

tatagtattt gtattt 16

<210> SEQ ID NO 42
<211> LENGTH: 12

-continued

<212> TYPE: DNA
<213> ORGANISM: Tobacco mosaic virus

<400> SEQUENCE: 42

tatagtcgta tt 12

What is claimed is:

1. A method for identifying a nucleic acid sequence in a donor plant having the same function as that in a host plant that belongs to a different genus, family, order, class, subdivision, or division from said donor plant, said method comprising the steps of:

(a) preparing a library of cDNAs, genomic DNAs, or a pool of RNAs of said donor plant;

(b) constructing recombinant viral nucleic acids obtained from a tobamo virus comprising an unidentified nucleic acid insert obtained from said library in an antisense orientation relative to said DNA or RNA sequence of said donor plant;

(c) infecting a host plant with one or more said recombinant viral nucleic acids, and transiently expressing said unidentified nucleic acid in the host plant;

(d) detecting one or more changes in said host plant;

(e) identifying said recombinant viral nucleic acid that results in changes in said host plant;

(f) determining the sequence of said nucleic acid insert in said recombinant viral nucleic acid; and

(g) determining the sequence of an entire open reading frame of a gene in said donor from which said nucleic acid insert is derived.

2. The method according to claim 1, wherein said nucleic acid sequence comprising a GTP binding protein open reading frame.

3. A method for identifying a nucleic acid sequence in a host plant having the same function as that in a donor plant that belongs to a different genus, family, order, class, subdivision, or division from said donor plant, said method comprising the steps of:

(a) preparing a cDNA library, a genomic DNA library, or a pool of RNA of said donor plant;

(b) constructing recombinant viral nucleic acids obtained from a tobamo virus comprising an unidentified nucleic acid insert obtained from said library in an antisense orientation relative to said DNA or RNA sequence of said donor plant;

(c) infecting a host plant with one or more said recombinant viral nucleic acids, and transiently expressing said unidentified nucleic acid in the host plant;

(d) detecting one or more changes in said host plant;

(e) identifying said recombinant viral nucleic acid that results in changes in said host plant; and

(f) determining the sequence of said nucleic acid insert in said recombinant viral nucleic acid; and

(g) determining the sequence of an entire open reading frame of a gene in said host plant, the expression of which gene is affected by said insert.

4. The method according to claim 3, wherein said nucleic acid sequence comprising a GTP binding protein open reading frame.

5. A method for detecting the presence of a nucleic acid sequence that has homology in a donor plant and in a host plant, wherein said donor plant and said host plant belong to a different genus, family, order, class, subdivision, or division from said donor plant, said method comprising the steps of:

(a) preparing a cDNA library, a genomic DNA library, or a pool of RNA of said donor plant;

(b) constructing recombinant viral nucleic acids obtained from a tobamo virus comprising an unidentified nucleic acid insert obtained from said library in an antisense orientation relative to said DNA or RNA sequence of said donor plant;

(c) infecting a host plant with one or more said recombinant viral nucleic acids, and transiently expressing said unidentified nucleic acid in the host plant; and

(d) detecting one or more changes in said host plant.

6. The method according to claim 5, further comprising the steps of:

(e) identifying said recombinant viral nucleic acid that results in changes in said host plant;

(f) determining the sequence of said nucleic acid insert in said recombinant viral nucleic acid; and

(g) determining the sequence of an entire open reading frame of said donor from which said nucleic acid insert is derived.

7. The method according to claim 5, wherein said nucleic acid sequence comprising a GTP binding protein open reading frame having an antisense orientation.

* * * * *