Office de la Propriété Intellectuelle du Canada
Un organisme d'Industrie Canada

Canadian Intellectual Property Office
An agency of Industry Canada

CA 2191624 C 2006/08/08

(11)(21) 2 191 624

(12) BREVET CANADIEN
CANADIAN PATENT

(13) C

(86) Date de dépôt PCT/PCT Filing Date: 1996/03/28
(87) Date publication PCT/PCT Publication Date: 1996/10/03
(45) Date de délivrance/Issue Date: 2006/08/08
(85) Entrée phase nationale/National Entry: 1996/11/28
(86) N° demande PCT/PCT Application No.: JP 1996/000812
(87) N° publication PCT/PCT Publication No.: 1996/030510
(30) Priorité/Priority: 1995/03/29 (JP7/96126)

(51) Cl.Int./Int.Cl. *C12N 15/82* (2006.01),
C12N 9/16 (2006.01)

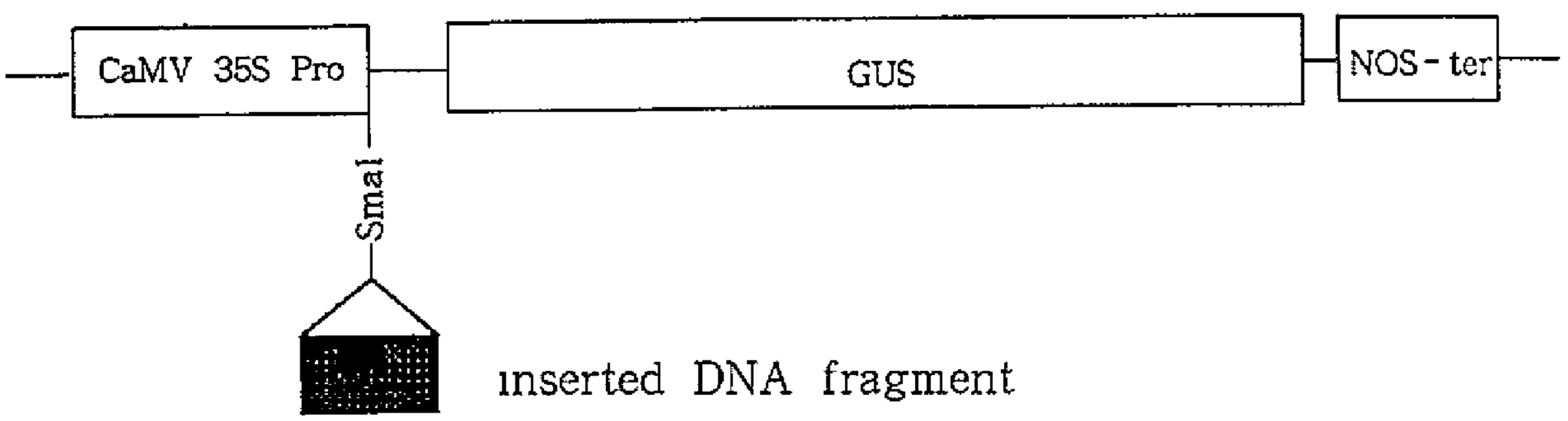
(72) Inventeurs/Inventors:
MORIOKA, SHINJI, JP;
UEKI, JUN, JP

(73) Propriétaire/Owner:
JAPAN TOBACCO INC., JP

(74) Agent: ROBIC

(54) Titre : FRAGMENT D'ADN, VECTEUR DE RECOMBINAISON LE CONTENANT ET PROCEDE D'EXPRESSION D'UN GENE ETRANGER L'UTILISANT

(54) Title: DNA FRAGMENT, RECOMBINANT VECTOR CONTAINING THE SAME AND METHOD FOR EXPRESSING FOREIGN GENES USING THE SAME



(57) Abrégé/Abstract:
A novel DNA which has a sequence different from those of publicly known DNAs capable of promoting the expression of an alien gene and can remarkably promote the expression of an alien gene. An isolated DNA fragment having a base sequence represented by SEQ ID NO:1 in the Sequence Listing; and an isolated DNA fragment represented by this base sequence wherein one or more nucleotides have been added, inserted, deleted or replaced and having the effect of promoting the expression of a gene located in the downstream thereof.

世界知的所有権機関
国際事務局

PCT

特許協力条約に基づいて公開された国際出願

(51) 国際特許分類6 C12N 15/11, 15/63, 15/82, 9/16	A1	(11) 国際公開番号 WO96/30510 (43) 国際公開日 1996年10月3日(03.10.96)

(21) 国際出願番号 PCT/JP96/00812
(22) 国際出願日 1996年3月28日(28.03.96)

(30) 優先権データ
特願平7/96126 1995年3月29日(29.03.95) JP

(71) 出願人（米国を除くすべての指定国について）
日本たばこ産業株式会社(JAPAN TOBACCO INC.)[JP/JP]
〒105 東京都港区虎ノ門二丁目2番1号 Tokyo, (JP)
(72) 発明者；および
(75) 発明者／出願人（米国についてのみ）
森岡真二(MORIOKA, Shinji)[JP/JP]
植木 潤(UEKI, Jun)[JP/JP]
〒438 静岡県磐田郡豊田町東原700 日本たばこ産業株式会社
遺伝育種研究所内 Shizuoka, (JP)
(74) 代理人
弁理士 谷川英次郎(TANIGAWA, Hidejiro)
〒102 東京都千代田区飯田橋4丁目5番12号 岩田ビル6階
谷川国際特許事務所 Tokyo, (JP)

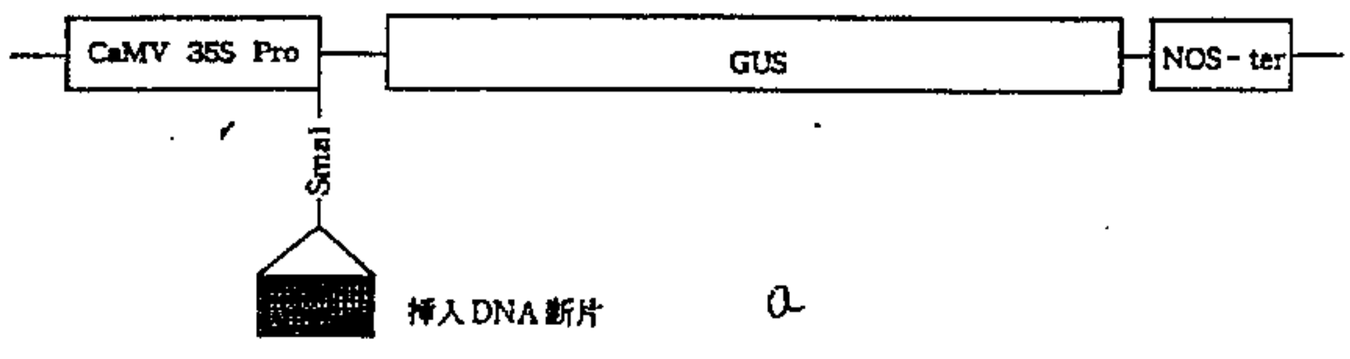
(81) 指定国
AU, CA, CN, JP, US, 欧州特許(AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

添付公開書類 国際調査報告書

2191624

(54) Title : DNA FRAGMENT, RECOMBINATION VECTOR CONTAINING THE SAME, AND METHOD OF THE EXPRESSION OF ALIEN GENE WITH THE USE OF THE SAME

(54) 発明の名称 DNA断片、それを含む組換えベクター及びそれを用いた外来遺伝子の発現方法



a ... DNA insert

(57) Abstract

A novel DNA which has a sequence different from those of publicly known DNAs capable of promoting the expression of an alien gene and can remarkably promote the expression of an alien gene. An isolated DNA fragment having a base sequence represented by SEQ ID NO:1 in the Sequence Listing; and an isolated DNA fragment represented by this base sequence wherein one or more nucleotides have been added, inserted, deleted or replaced and having the effect of promoting the expression of a gene located in the downstream thereof.

2191624

SPECIFICATION

DNA Fragment, Recombinant Vector Containing the Same and Method for Expressing Foreign Genes Using the Same

TECHNICAL FIELD

The present invention relates to a novel DNA fragment having function to promote expression of genes, a vector containing the same and a method for expressing foreign genes using the same.

BACKGROUND ART

Promotion of expression of foreign genes is one of the most important techniques in applying genetic engineering processes to plants. One of the methods therefor is utilization of a DNA having a nucleotide sequence which promotes expression of a gene.

Known nucleotide sequences which promote expression of foreign genes include the intron of the catalase gene of castor bean (Japanese Laid-open Patent Application (Kokai) No. 3-103182; Tanaka et al., Nucleic Acids Res. 18, 6767-6770 (1990)). However, since there are wide varieties of plants to be manipulated and since promotion of expression of genes is required in each of the desired growth stages or tissues of organs, it is desired that wide varieties of DNAs which promote expression of genes can be utilized.

DISCLOSURE OF THE INVENTION

Accordingly, an object of the present invention is to provide a novel DNA, which can promote expression of

foreign genes and which has a nucleotide sequence different from those of known DNAs that promote expression of foreign genes.

The present inventors intensively studied to discover introns of rice phospholipase D (hereinafter also referred to as "PLD") gene by comparing a rice cDNA and a rice genomic DNA, and discovered that one of the introns has a function to prominently promote expression of the gene downstream thereof, thereby completing the present invention.

That is, the present invention provides an isolated DNA fragment having a nucleotide sequence shown in SEQ ID NO. 1 in Sequence Listing or having a nucleotide sequence which is the same as the nucleotide sequence shown in SEQ ID NO. 1 in Sequence Listing except that one or a plurality of nucleotides are added, inserted, deleted or substituted, the latter nucleotide sequence having a function to promote expression of a gene downstream thereof. More specifically, the present invention provides for an isolated DNA fragment having a nucleotide sequence shown in SEQ ID NO. 1 in Sequence Listing or having a nucleotide sequence, which hybridizes with said nucleotide sequence shown in SEQ ID NO. 1 when the hybridization is carried out in a hybridization solution containing 0.5M sodium phosphate buffer (pH7.2), 7% SDS, 1 mM EDTA and 100 µg/ml of salmon sperm DNA, at 65°C for 16 hours, and then washing the product twice with a washing solution containing 0.5xSSC and 0.1% SDS at 65°C for 20 minutes

each, the latter nucleotide sequence having a function to promote expression of a gene downstream thereof in a plant. The present invention also provides a recombinant vector comprising the above-mentioned DNA fragment according to the present invention and foreign gene to be expressed, which is operably linked to the DNA fragment at a downstream region of the DNA fragment. The present invention further provides a method for expressing a foreign gene comprising introducing the recombinant vector according to the present invention into host cells and expressing the foreign gene.

As experimentally confirmed in the Example described below, the DNA fragment according to the present invention largely promotes expression of the gene downstream of the DNA fragment. Therefore, it is expected that the present invention will largely contribute to expression of foreign genes by genetic engineering processes.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 shows the important part of a genetic map of pBI221 into which the DNA fragment according to the present invention is inserted, which was prepared in the Example of the present invention.

BEST MODE FOR CARRING OUT THE INVENTION

As mentioned above, the DNA fragment according to the present invention has a nucleotide sequence shown in SEQ ID NO. 1 in the Sequence Listing. As will be described in detail in the Example below, introns located upstream of rice PLD gene were identified by comparing the nucleotide sequence of the cDNA of rice PLD gene and that of the rice genomic DNA. A fragment containing one of these intron sequences having a size of 173 bp located at the 5'-flanking region was prepared by PCR and the DNA fragment was inserted into an upstream site of a reporter gene of an expression vector containing the reporter gene. By checking the expression activity of the reporter gene, it was confirmed that the DNA fragment has a function to promote expression of the gene downstream thereof. The

nucleotide sequence of the DNA fragment according to the present invention corresponds to 1661nt to 1843nt of the nucleotide sequence of the rice genomic PLD gene, which nucleotide sequence is shown in SEQ ID NO. 3 of the Sequence Listing.

The nucleotide sequence of the above-mentioned intron sequence having a size of 173 bp, which is located upstream of the rice PLD gene, is shown in SEQ ID NO. 4 in the Sequence Listing. Needless to say, the sequence shown in SEQ ID NO. 4 also has a function to promote expression of the gene downstream thereof. The nucleotide sequence shown in SEQ ID NO. 4 corresponds to 1666nt to 1838nt of the nucleotide sequence of the rice genomic PLD gene, which is shown in SEQ ID NO. 3 in the Sequence Listing.

Since the DNA fragment according to the present invention is an intron existing upstream of the rice PLD gene, and since its nucleotide sequence was determined according to the present invention, the DNA fragment may easily be prepared by PCR using the rice genomic DNA as a template. PCR is a conventional technique widely used in the field of genetic engineering and a kit therefor is commercially available, so that those skilled in the art can easily perform the PCR. One concrete example thereof is described in detail in the Example below.

It is well-known in the art that there are cases wherein the physiological activity of a physiologically

active DNA sequence is retained even if the nucleotide sequence of the DNA is modified to a small extent, that is, even if one or more nucleotides are added, inserted, deleted or substituted. Therefore, DNA fragments having the same nucleotide sequence as shown in SEQ ID NO. 1 except that the DNA fragments have such modifications, which have the function to promote expression of the gene downstream thereof, are included within the scope of the present invention. That is, the DNA fragments having the same nucleotide sequence as shown in SEQ ID NO. 1 except that one or more nucleotides are added, deleted or substituted, which have the function to promote expression of the gene downstream thereof, are included within the scope of the present invention. Particularly, in the nucleotide sequence shown in SEQ ID NO. 1, the 5 nucleotides at the 5'-end and the 6 nucleotides at the 3'-end are the nucleotides in the exon regions, so that it is thought that the nucleotide sequences which do not have these regions also have the function to promote gene expression. Thus, these DNA fragments are within the scope of the present invention.

Modification of DNA which brings about addition, deletion or substitution of the amino acid sequence encoded thereby can be attained by the site-specific mutagenesis which is well-known in the art (e.g., Nucleic Acid Research, Vol. 10, No. 20, p6487-6500, 1982). In the present specification, "one or a plurality of

nucleotides" means the number of nucleotides which can be added, deleted or substituted by the site-specific mutagenesis.

Site-specific mutagenesis may be carried out by, for example, using a synthetic oligonucleotide primer complementary to a single-stranded phage DNA except that the desired mutation as follows. That is, using the above-mentioned synthetic oligonucleotide as a primer, a complementary chain is produced by a phage, and host bacterial cells are transformed with the obtained double-stranded DNA. The culture of the transformed bacterial cells is plated on agar and plaques are formed from a single cell containing the phage. Theoretically, 50% of the new colonies contain the phage having a single-stranded chain carrying the mutation and remaining 50% of the colonies contain the phage having the original sequence. The obtained plaques are then subjected to hybridization with a kinase-treated synthetic probe at a temperature at which the probe is hybridized with the DNA having exactly the same sequence as the DNA having the desired mutation but not with the original DNA sequence that is not completely complementary with the probe. Then the plaques in which the hybridization was observed are picked up, cultured and the DNA is collected.

In addition to the above-mentioned site-specific mutagenesis, the methods for substituting, deleting or adding one or more amino acids without losing the

function include a method in which the gene is treated with a mutagen and a method in which the gene is selectively cleaved, a selected nucleotide is removed, added or substituted and then the gene is ligated.

The DNA fragment according to the present invention has a function to promote expression of the gene downstream thereof. Therefore, by inserting the DNA fragment according to the present invention into the transcriptional region of a desired foreign gene to be expressed, preferably into the 5'-end region of the transcriptional region, expression of the foreign gene is promoted. The method for expressing a foreign gene has already been established in the field of genetic engineering. That is, by inserting the desired foreign gene into a cloning site of an expression vector, introducing the resulting vector into host cells and expressing it, the foreign gene may be expressed. According to the method of the present invention, the DNA fragment according to the present invention is inserted at a site upstream of the foreign gene in a manner such that the DNA fragment is operably linked to the foreign gene, and the foreign gene is expressed. The term that the DNA fragment according to the present invention is "operably linked" to the foreign gene means that expression of the foreign gene is detectably increased by inserting the DNA fragment according to the present invention when compared with the case wherein the DNA

fragment according to the present invention is not inserted. The DNA according to the present invention may be inserted into the site immediately upstream of the foreign gene. Alternatively, another sequence may be located between the DNA according to the present invention and the foreign gene. Although the size of this intervening sequence is not restricted, it usually has a size of 0 - 1000 bp. A promoter sequence is located upstream of the DNA fragment according to the present invention. The DNA fragment according to the present invention may be inserted into the site immediately downstream of the promoter, or another sequence may be located between the promoter and the DNA according to the present invention. Although the size of this intervening sequence is not restricted, it is usually 0 - 1000 bp. In summary, all recombinant vectors with which the expression of the foreign gene is detectably increased by inserting the DNA fragment according to the present invention when compared with the case wherein the DNA fragment is not inserted, are within the scope of the present invention.

Since the nucleotide sequence of the cloning site of an expression vector is known, the DNA fragment according to the present invention may easily be inserted into the vector.

Wide varieties of such an expression vector are well-known in the art and are commercially available.

These expression vectors contain at least a replication origin for replication in the host cells, a promoter, a cloning site giving a restriction site for inserting the foreign gene, and a selection marker such as drug resistance, and usually contain a terminator which stably terminates transcription, and an SD sequence when the host is a bacterium. In the method of the present invention, any of these known expression vectors may be employed.

Example

The present invention will now be described in more detail by way of examples thereof. However, the present invention is not restricted to the examples.

1. Purification of PLD of Rice Bran

For purification, a reference (Takano et al., Journal of Japan Food Industry Association, 34, 8-13 (1987) was referred. The enzyme activity was measured by employing phosphatidylcholine as a substrate and quantifying the choline generated by the enzyme reaction (Imamura et al., J. Biochem. 83, 677-680 (1978)). It should be noted, however, the enzyme reaction was stopped by heat treatment at 95°C for 5 minutes.

That is, to 100 g of bran of rice (*Oryza sativa*), variety "KOSHIHIKARI", one liter of hexane was added and the mixture was stirred for a whole day and night, thereby defatting the rice bran. To the resultant, 10 g of Polycral AT (trademark, polyvinylpyrrolidone,

commercially available from GAF Chemical) and 500 ml of 10 mM Tris-HCl buffer (pH7) containing 1 mM $CaCl_2$ and 5 mM 2-mercaptoethanol were added, and the resulting mixture was stirred for 1 hour to extract the enzyme. The extract was filtered through an 8-layered gauze and the filtrate was centrifuged at 15,000 x g for 20 minutes, followed by recovering the middle layer as a crude extract. The crude extract was treated with ammonium sulfate (65% saturation) and the generated precipitates were collected by centrifugation (15,000 x g, 20 minutes), followed by dialyzing the precipitates after dissolution against the above-mentioned buffer. After the dialysis, precipitates were eliminated by filtration to obtain ammonium sulfate fraction.

The ammonium sulfate fraction was applied to a column (2.0 x 10 cm) of DEAE-Cellulose (commercially available from Whattman) equilibrated with buffer A (10 mM Tris-HCl, pH 7, 1 mM $CaCl_2$, 1 mM 2-mercaptoethanol). After washing the column with about 100 ml of buffer A containing 50 mM NaCl, elution was carried out with 120 ml of buffer A having a linear gradient of NaCl concentration from 50 mM to 350 mM. PLD was eluted at a NaCl concentration of about 0.2 M. The fraction having PLD activity was collected as an eluted solution (DEAE-cellulose).

To the eluted solution (DEAE-cellulose), 3 M ammonium sulfate was added in an amount attaining the

final concentration of ammonium sulfate of 1 M, and the resulting mixture was applied to a Phenyl Sepharose column (commercially available from Pharmacia, 2.6 x 10 cm) equilibrated with buffer A containing 1 M ammonium sulfate. Elution was performed using 240 ml of buffer A having a linear gradient of ammonium sulfate concentration from 1.0 M to 0 M. PLD was eluted at a concentration of ammonium sulfate of about 0.1 M. The fraction having the activity was recovered and dialyzed against buffer A to obtain an eluted solution (Phenyl Sepharose).

The eluted solution (Phenyl Sepharose*) was applied to Mono Q column (anion-exchange column commercially available from Pharmacia, 16 x 10 cm) equilibrated with buffer A, and elution was performed using 150 ml of buffer A having a gradient of NaCl concentration from 50 mM to 350 mM. PLD was eluted at NaCl concentration from 210 mM to 235 mM. The fraction having PLD activity was recovered and dialyzed against buffer A to obtain an eluted solution (Mono Q* 1st).

The eluted solution (Mono Q 1st) was concentrated to 0.5 ml by ultrafiltration and applied to Superose 6 column (commercially available from Pharmacia, 1.0 x 30 cm equilibrated with buffer A containing 0.1 M NaCl and elution was performed using the same buffer. The molecular weight of PLD was estimated to be 78 kDa. The fraction having PLD activity was recovered as an eluted

*** Trademarks**

solution (Superose 6).

To the eluted solution (Superose 6), 2.5 ml of 40% Carrier Ampholite (commercially available from Pharmacia, pH4.0-6.0) and distilled water were added to attain a final volume of 50 ml and isoelectric electrophoresis was carried out using Rotofore (commercially available from Biorad). Electrophoresis was performed at 2°C with a constant power of 12W for 4 hours. PLD activity was observed at about pH 4.9. The fraction having PLD activity was collected and dialyzed against buffer A to obtain an isoelectric electrophoresis fraction.

The isoelectric electrophoresis fraction was applied to Mono Q column (commercially available from Pharmacia, 0.5 x 5 cm) and eluted with NaCl having a linear gradient of concentration of 50 mM to 350 mM. PLD was eluted at NaCl concentrations of about 210 mM and about 235 mM. The two fractions having PLD activity were recovered as eluted solutions (Mono Q 2nd-I, II).

Purities of the eluted solutions (Mono Q 2nd-I, II) were checked by SDS-polyacrylamide electrophoresis (Laemmli (1970)) using 7.5% acrylamide. After the electrophoresis, the gel was stained with Coomassie brilliant blue R-250. With either eluted solution, a main band was observed at a position corresponding to a molecular weight of 82 kDa. With the eluted solution (Mono Q 2nd-II), only a single band was observed.

By the purification described above, the

purification magnifications of the eluted solutions (Mono Q 2nd-I, II) were 380 times and 760 times, respectively, with respect to the crude extract.

Properties of the enzymes contained in the two fractions were determined. The results are shown in Table 1. The buffer solutions used for the measurement of the optimum pH were sodium acetate (pH 4-6), MES-NaOH (pH 5.5-7.0) and Tris-HCl (pH 7-9), all of which have a concentration of 100 mM in all of the buffer solutions. The pH stability means the pH range in which decrease in the enzyme activity is not observed after leaving the enzyme at the respective pH at 25°C for 30 minutes. The temperature stability was measured by measuring the remaining activity after leaving the enzyme to stand at 4°C, 25°C, 37°C or 50°C for 30 minutes. The substrate specificity was measured at a substrate concentration of 5 mM and expressed in terms of the relative activity taking the enzyme activity to phosphatidylcholine as 100.

Table 1

	Mono Q 2nd-I	Mono Q 2nd-II
Km Value	0.29 mM	0.29 mM
Optimum pH	6	6
pH Stability	7-8	7-8
Temperature Stability	4-37°C	4-37°C
Ca^{2+} Dependency	not less than 20 mM	not less than 20 mM
Substrate Specificity		
Phosphatidylcholine	100	100
Lysophosphatidylcholine	13	12
Sphingomyelin	6	4

2. Proof that Purified Protein is PLD

Each of the eluted solutions (Mono Q 2nd-I, II) was subjected to SDS-polyacrylamide gel electrophoresis in the same manner as in the purity test, and the obtained patterns were transferred to PVDF membranes (commercially available from Millipore), followed by staining the membranes. The band of the protein having the molecular weight of 82 kDa was cut out and the amino acid sequence of the N-terminal region of the protein was determined by a protein sequencer (commercially available from Shimazu Seisakusho, PSQ-1). For both proteins, amino acid sequence up to 10 residues from the N-terminal could be determined, and the determined sequences were identical. The sequence was as follows.

Val Gly Lys Gly Ala Thr Lys Val Tyr Ser

Although the relationship between the proteins

having the molecular weight of 82 kDa contained in the two fractions having the enzyme activity is not clear, it is thought that they have high homology in their amino acid sequences, so that it was judged that there would be no problem even if a mixture of the fractions is used as an antigen for preparing an antibody.

A mixture of the eluted solutions (Mono Q 2nd-I, II) was subjected to SDS-polyacrylamide gel electrophoresis using 7.5% acrylamide, and the gel was stained with Coomassie brilliant blue R-250. The band of the protein having the molecular weight of 82 kDa was cut out and recovered by electroelution (25 mM Tris, 192 mM glycine, 0.025% SDS, 100V, 10 hours). Then SDS was removed by electrodialysis (15 mM ammonium bicarbonate, 200 V, 5 hours) and the resultant was lyophilized. For the electroelution and electrodialysis, BIOTRAP* (commercially available from Schleicher & Schuell) was used.

The protein having the molecular weight of 82 kDa highly purified by the above-described method was administered to a rabbit in an amount of 50 µg per time at 7 days' intervals. Immunological titration test was performed for the sera before the immunization and after the third immunization. To the PLD solution containing 8.6×10^{-3} units of PLD, were added 0 - 50 µl of the serum before the immunization or after the third immunization, 50 µl of 250 mM Tris-HCl (pH7.0), 5 µl of 50 mM $CaCl_2$, 50 µl of 0.2% Triton X-100 (trademark) and

*** Trademark**

water to a total volume of 250 µl, and the mixture was left to stand at room temperature for 2.5 hours. To the resultant, 200 µl of Protein A Sepharose (commercially available from Pharmacia) was added and the resulting mixture was gently shaken at room temperature for 2 hours. The mixture was then centrifuged (500 x g, 5 minutes) and the enzyme activity in the supernatant was measured. Taking the measured enzyme activity in the case where the serum was not added as 100%, the enzyme activities in cases where 20 µl and 50 µl of the serum before immunization were added were 95% and 88%, respectively, and the enzyme activities in cases where 20 µl and 50 µl of the serum after the third immunization were added were 75% and 30%, respectively. These results prove that the protein having the molecular weight of 82 kDa is PLD.

3. Determination of Amino Acid Sequence of Internal Regions

The PLD protein was fragmentated in a gel (Cleveland et al., J. Biol. Chem., 252, 1102(1977)). The cut out gel containing the PLD protein was inserted into a stacking gel well on a 15% acrylamide gel prepared for separation of peptides, and *Staphylococcus aureus* V8 protease (commercially available from Wako Pure Chemical Industries, Ltd) in an amount of 1/10 volume of the PLD protein was overlaid, followed by starting electrophoresis. The electrophoresis was stopped at the time point at which the bromophenol blue reached the

center of the stacking gel and then restarted 30 minutes later. After the electrophoresis, the pattern was transferred to a PVDF membrane and the membrane was stained. Clear bands were observed at the positions corresponding to molecular weights of 20, 14, 13, 11 and 10 kDa. Each of the bands of the peptide fragments having molecular weights of 20, 14 and 13 kDa were cut out and their amino acid sequences were determined by a protein sequencer. The determined sequences are as follows.

20 kDa Asn Tyr Phe His Gly Ser Asp Val Asn ? Val
Leu ? Pro Arg Asn Pro Asp Asp(Asp) ? ?
Ile

14 kDa Thr ? Asn Val Gln Leu Phe Arg Ser Ile Asp
Gly Gly Ala Ala Phe Gly Phe Pro Asp Thr Pro
Glu Glu Ala Ala Lys ? Gly Leu Val Ser Gly

13 kDa Ile Ala Met Gly Gly Tyr Gln Phe Tyr His Leu
Ala Thr Arg Gln Pro Ala Arg Gly Gln Ile His
Gly Phe Arg Met Ala Leu ? Tyr Glu His Leu
Gly Met Leu ? Asp Val Phe

(In the sequences, "?" means the amino acid residue which could not be identified, and the amino acid residue in parentheses means that the amino acid residue could not be identified confidentially.

4. <u>Preparation of cDNA Library of Rice Immature Seeds</u>

Total RNAs were extracted from immature seeds obtained after 5 days from flowering by the SDS-phenol

method, and prepared by the lithium chloride precipitation. Poly(A)$^+$RNA was prepared using Oligotex-dT30* (commercially available from Takara Shuzo) according to the instructions provided by the manufacturer. For the cDNA cloning, cDNA synthesis System Plus (commercially available from Amersham) and cDNA Cloning System λgt10* (commercially available from Amersham) were used. However, λZAPII* vector (commercially available from Stratagene) was used as the cloning vector and XL1-Blue was used as the host cells.

5. Preparation of Probes

Oligonucleotides corresponding to the amino acid sequences of PLD were synthesized by a DNA synthesizer (commercially available from Applied Biosystems). The sequences thereof as well as the corresponding amino acid sequences are as follows.

20KF 5' AAYTAYTTYCAYGG 3'

20KR1 5' RTCRTCRTCNGGRTT 3'

(In these sequences, "R" represents a purine base A or G; "Y" represents a pyrimidine base T or G; and N represents G, A, T or C.)

The 20KF is a mixture of 32 kinds of oligonucleotides containing the DNA sequences encoding the amino acid sequence of

Asn Tyr Phe His Gly

found in a peptide having a molecular weight of 20 kDa, and the 20KR1 is a mixture of 128 kinds of

*** Trademarks**

oligonucleotides containing complementary chains of the DNA sequences encoding the amino acid sequence of

Asn Pro Asp Asp(Asp)

found in the same peptide.

The cDNA synthesis was carried out using 10 ng of Poly(A)$^+$RNA, 0.3 µg of random hexamer (N6), 10 U of an RNase inhibitor (RNA Guard*, commercially available from Pharmacia), 1 mM each of dATP, dCTP, dGTP and dTTP, 1 x PCR buffer (commercially available from Takara Shuzo), 50 mM of magnesium chloride and 100 U of a reverse transcriptase (M-MuLV RTase, commercially available from BRL) in a total volume of 10 µl. The reaction was carried out at 37°C for 30 minutes and the reaction mixture was then heated at 95°C for 5 minutes, followed by retaining the reaction mixture in ice.

Polymerase chain reaction (PCR) was performed using the above-described cDNA as a template and 20KF and 20KR1 as primers. The reaction was performed using 10 µl of the cDNA synthesis reaction mixture, a mixture of 50 pmol each of the primers, 200 µM each of dATP, dCTP, dGTP and dTTP, 1 x PCR buffer (commercially available from TAKARA SHUZO), and 2.5 U of AmpliTaq DNA polymerase (commercially available from TAKARA SHUZO) in a total volume of 50 µl. A cycle of temperature conditions of 94°C for 1 minute, 40°C for 1 minute and 72°C for 2.5 minutes was repeated 30 times in a DNA Thermocycler

*** Trademark**

(commercially available from Perkin Elmer Cetus).

The PCR product was separated on 2% agarose gel. A small number of fragments were detected by the ethidium bromide staining method. One of them had a size of 94 bp as expected.

The PCR fragment was cut out from the gel and subcloned into pUC19 plasmid. The DNA sequence of the subcloned PCR fragment was determined by the dideoxy method using T7 sequencing kit (commercially available from Pharmacia). Between the two primers, a DNA sequence encoding the expected amino acid sequence was observed. The nucleotide sequence of the DNA between the primers and the amino acid sequence encoded thereby are as follows.

```
C TCT GAC GTG AAC TGT GTT CTA TGC CCT CGC
  Ser Asp Val Asn Cys Val Leu Cys Pro Arg
```

Isotope ^{32}P (commercially available from Amersham) was incorporated into the oligonucleotide using a DNA 5'-end labelling kit MEGALABEL* (commercially available from Takara Shuzo) to obtain a radioactive oligonucleotide probe.

6. Screening of PLD Gene-containing Clones

Using the radioactive oligonucleotide as a probe, a cDNA library was screened. Hybridization solution contained 0.5 M sodium phosphate buffer (pH 7.2), 7% SDS, 1 mM EDTA and 100 µg/ml of salmon sperm DNA, and hybridization was performed after adding the probe to the

*** Trademark**

hybridization solution at 45°C for 16 hours. The washing solution contained 0.3 M NaCl and 30 mM sodium citrate, and washing was performed twice at 45°C for 20 minutes each. Positive plaques were isolated and subcloned *in vivo* into pBluescript* plasmid (commercially available from Stratagene) in accordance with the instructions provided by the manufacturer of λZAPII cloning vector. The nucleotide sequence was determined by the dideoxy method. As a result, a region encoding the internal amino acid sequence determined in the "Section 3" existed.

7. Determination of Nucleotide Sequence of 5'-end Region

Since a clone containing the full length of cDNA could not be isolated, a DNA fragment having the 5'-end region was prepared by RACE method (Edwards et al., Nucleic Acids Res., 19, 5227-5232 (1991)). 5'-AmpliFINDER RACE Kit (commercially available from Clonetech) was used in accordance with the manual attached to the product. An oligoDNA was synthesized based on the nucleotide sequence of the cDNA determined in "Section 6", and PCR was performed using the mRNA prepared by the method described in "Section 4" as a template. The PCR product was subcloned into a PCRII* vector (commercially available from Invitrogen) and the nucleotide sequence was determined by the dideoxy method. The thus determined nucleotide sequence of the cDNA of rice PLD as well as the deduced amino acid sequence

*** Trademarks**

encoded thereby is shown in SEQ ID NO. 2 in the Sequence Listing. It is thought that translation is initiated from the 182nd nucleotide shown in SEQ ID NO. 2 since a termination codon exists at 36 bases upstream thereof.

8. Isolation of PLD Genomic Clone Corresponding to PLD cDNA and Identification of Promoter Region

To isolate a genomic DNA clone having the regulatory sequence of the PLD gene corresponding to the PLD cDNA determined in "section 6", which was cloned into pBluescript plasmid, a genomic library of rice, variety "KOSHIHIKARI" was prepared. This was carried out by partially digesting DNAs from live leaves of KOSHIHIKARI with *Mbo* I, purifying a fraction having a size of 16 - 20 kb by sucrose gradient centrifugation, and using lambda DASH II* (commercially available from Stratagene) and GigapackII*Gold (commercially available from Stratagene). The genomic library was screened with the PLD cDNA clone as a probe. The screening was carried out in the same manner as in "Section 6" except that hybridization was performed at 65°C for 16 hours, the washing solution contained 0.5 x SSC and 0.1% SDS, and that the washing was performed twice at 65°C for 20 minutes each. The nucleotide sequence of the hybridized genomic clone was determined by the dideoxy method. As a result, a region homologous to the cDNA sequence determined in "Section 6" existed.

The transcription initiation site was determined by

*** Trademarks**

the method described in "Section 7". In the vicinity of the transcription initiation site, a "TATA" consensus sequence box was observed. The ATG translation initiation site was determined based on the determined DNA sequence as the most upstream ATG codon in the translation open reading frame of the clone and as the ATG codon which is first accessible in the mRNA synthesized in rice.

The DNA sequence of a part of the genomic clone hybridized with the cDNA clone is shown in SEQ ID NO. 3. In the genomic DNA sequence, an open reading frame starting from the ATG translation initiation codon, which overlaps with the corresponding cDNA sequence has been identified. The promoter region exists upstream of the ATG translation initiation codon and starts from the site immediately upstream thereof.

9. <u>Identification of Introns and Analysis of Functions thereof on Expression of Genes</u>

From comparison between the cDNA (SEQ ID NO. 2) and the genomic DNA (SEQ ID NO. 3), it was proved that 3 introns exist in PLD gene. Among these, the intron having a size of 173 bp located at the 5'-flanking region of the mRNA (i.e., the nucleotide sequence between 1666nt and 1838nt of the nucleotide sequence shown in SEQ ID NO. 3, the sequence being shown in SEQ ID NO. 4) was tested for its influence on expression of a gene in plant cells. Primers of 15mer each of which contains 5 bases of exon

region (5'-ACCCGGTAAGCCCAG-3', 3'-CCCCCGCGTCCATCC-5') were synthesized and PCR was carried out using the genomic clone as a template according to the method described in the Section of "5. Preparation of Probes". The PCR product was subcloned into PCRII vector and a fragment was cut out with *Eco* RI. The fragment was blunted and inserted into the *Sma* I site of a plasmid pBI221 (commercially available from Toyobo) (see Fig. 1). The obtained recombinant plasmid was introduced into rice cultured cells (Baba et al., Plant Cell Physiol. 27, 463-471 (1986)) in accordance with the reported method (Shimamoto et al., Nature, 338, 274-276 (1989)) and β-glucuronidase (GUS) activity was measured. As shown in Table 2, by introducing the intron, the GUS activity was increased. Further, as shown in Table 3, increase in the GUS activity was also observed in the case where the intron was inserted in the reverse direction. The direction of the intron was determined based on the sizes of the fragments cut out with *Bgl* II and *Bam* HI, utilizing the *Bgl* II site existing in the intron sequence and the *Bam* HI site existing in pBI221.

Table 2

Plasmid	GUS Activity
pBI221	10.4
pBI221 + intron	105.7

(pmol MU/min./mg protein)

Table 3

Plasmid	GUS Activity
pBI221	8.8
pBI221 + intron	79.4
pBI221 + intron (reverse direction)	54.2

(pmol MU/min./mg protein)

SEQUENCE LISTING

SEQ ID NO.: 1

SEQUENCE LENGTH: 183

SEQUENCE TYPE: Nucleic Acid

MOLECULE TYPE: Genomic DNA

ORIGINAL SOURCE

ORGANISM: *Oryza sativa*

SEQUENCE DESCRIPTION

ACCCGGTAAG CCCAGTGTGC TTAGGCTAAG CGCACTAGAG CTTCTTGCTC GCTTGCTTCT 60

TCTCCGCTCA GATCTGCTTG CTTGCTTGCT TCGCTAGAAC CCTACTCTGT GCTGCGAGTG 120

TCGCTGCTTC GTCTTCCTTC CTCAAGTTCG ATCTGATTGT GTGTGTGGGG GGGCGCAGGT 180

AGG 183

SEQ ID NO.: 2

SEQUENCE LENGTH: 3040

SEQUENCE TYPE: Nucleic Acid

MOLECULE TYPE: cDNA to mRNA

ORIGINAL SOURCE

ORGANISM: *Oryza sativa*

SEQUENCE DESCRIPTION

AGTCTCTCTT CTCCCGCAAT TTTATAATCT CGATCGATCC AATCTGCTCC CCTTCTTCTT 60

CTACTCTCCC CATCTCGGCT CTCGCCATCG CCATCCTCCT CTCCCTTCCC GGAGAAGACG 120

CCTCCCTCCG CCGATCACCA CCCGGTAGGG CGAGGAGGGA GCCAAATCCA AATCAGCAGC 180

C ATG GCG CAG ATG CTG CTC CAT GGG ACG CTG CAC GCC ACC ATC TTC GAG 229

Met Ala Gln Met Leu Leu His Gly Thr Leu His Ala Thr Ile Phe Glu

1 5 10 15

GCG GCG TCG CTC TCC AAC CCG CAC CGC GCC AGC GGA AGC GCC CCC AAG 277

Ala Ala Ser Leu Ser Asn Pro His Arg Ala Ser Gly Ser Ala Pro Lys

```
                20                  25                  30

TTC ATC CGC AAG TTT GTG GAG GGG ATT GAG GAC ACT GTG GGT GTC GGC     325
Phe Ile Arg Lys Phe Val Glu Gly Ile Glu Asp Thr Val Gly Val Gly

        35                  40                  45

AAA GGC GCC ACC AAG GTG TAT TCT ACC ATT GAT CTG GAG AAA GCT CGT     373
Lys Gly Ala Thr Lys Val Tyr Ser Thr Ile Asp Leu Glu Lys Ala Arg

    50                  55                  60

GTA GGG CGA ACT AGG ATG ATA ACC AAT GAG CCC ATC AAC CCT CGC TGG     421
Val Gly Arg Thr Arg Met Ile Thr Asn Glu Pro Ile Asn Pro Arg Trp
65                  70                  75                  80

TAT GAG TCG TTC CAC ATC TAT TGC GCT CAT ATG GCT TCC AAT GTG ATC     469
Tyr Glu Ser Phe His Ile Tyr Cys Ala His Met Ala Ser Asn Val Ile

                85                  90                  95

TTC ACT GTC AAG ATT GAT AAC CCT ATT GGG GCA ACG AAT ATT GGG AGG     517
Phe Thr Val Lys Ile Asp Asn Pro Ile Gly Ala Thr Asn Ile Gly Arg

            100                 105                 110

GCT TAC CTG CCT GTC CAA GAG CTT CTC AAT GGA GAG GAG ATT GAC AGA     565
Ala Tyr Leu Pro Val Gln Glu Leu Leu Asn Gly Glu Glu Ile Asp Arg

        115                 120                 125

TGG CTC GAT ATC TGT GAT AAT AAC CGC GAG TCT GTT GGT GAG AGC AAG     613
Trp Leu Asp Ile Cys Asp Asn Asn Arg Glu Ser Val Gly Glu Ser Lys

    130                 135                 140

ATC CAT GTG AAG CTT CAG TAC TTC GAT GTT TCC AAG GAT CGC AAT TGG     661
Ile His Val Lys Leu Gln Tyr Phe Asp Val Ser Lys Asp Arg Asn Trp
145                 150                 155                 160

GCG AGG GGT GTC CGC AGT ACC AAG TAT CCA GGT GTT CCT TAC ACC TTC     709
Ala Arg Gly Val Arg Ser Thr Lys Tyr Pro Gly Val Pro Tyr Thr Phe

                165                 170                 175
```

```
TTC TCT CAG AGG CAA GGG TGC AAA GTT ACC TTG TAC CAA GAT GCT CAT     757
Phe Ser Gln Arg Gln Gly Cys Lys Val Thr Leu Tyr Gln Asp Ala His
                180                 185                 190
GTC CCA GAC AAC TTC ATT CCA AAG ATT CCG CTT GCC GAT GGC AAG AAT     805
Val Pro Asp Asn Phe Ile Pro Lys Ile Pro Leu Ala Asp Gly Lys Asn
        195                 200                 205
TAT GAA CCC CAC AGA TGC TGG GAG GAT ATC TTT GAT GCT ATA AGC AAT     853
Tyr Glu Pro His Arg Cys Trp Glu Asp Ile Phe Asp Ala Ile Ser Asn
    210                 215                 220
GCT CAA CAT TTG ATT TAC ATC ACT GGC TGG TCT GTA TAC ACT GAG ATC     901
Ala Gln His Leu Ile Tyr Ile Thr Gly Trp Ser Val Tyr Thr Glu Ile
225                 230                 235                 240
ACC TTG GTT AGG GAC TCC AAT CGT CCA AAA CCT GGA GGG GAT GTC ACC     949
Thr Leu Val Arg Asp Ser Asn Arg Pro Lys Pro Gly Gly Asp Val Thr
                245                 250                 255
CTT GGG GAG TTG CTC AAG AAG AAG GCC AGT GAA GGT GTT CGG GTC CTC     997
Leu Gly Glu Leu Leu Lys Lys Lys Ala Ser Glu Gly Val Arg Val Leu
            260                 265                 270
ATG CTT GTG TGG GAT GAC AGG ACT TCA GTT GGT TTG CTA AAG AGG GAT    1045
Met Leu Val Trp Asp Asp Arg Thr Ser Val Gly Leu Leu Lys Arg Asp
        275                 280                 285
GGC TTG ATG GCA ACA CAT GAT GAG GAA ACT GAA AAT TAC TTC CAT GGC    1093
Gly Leu Met Ala Thr His Asp Glu Glu Thr Glu Asn Tyr Phe His Gly
    290                 295                 300
TCT GAC GTG AAC TGT GTT CTA TGC CCT CGC AAC CCT GAT GAC TCA GGC    1141
Ser Asp Val Asn Cys Val Leu Cys Pro Arg Asn Pro Asp Asp Ser Gly
305                 310                 315                 320
AGC ATT GTT CAG GAT CTG TCG ATC TCA ACT ATG TTT ACA CAC CAT CAG    1189
```

```
Ser Ile Val Gln Asp Leu Ser Ile Ser Thr Met Phe Thr His His Gln
                325                 330                 335
AAG ATA GTA GTT GTT GAC CAT GAG TTG CCA AAC CAG GGC TCC CAA CAA   1237
Lys Ile Val Val Val Asp His Glu Leu Pro Asn Gln Gly Ser Gln Gln
            340                 345                 350
AGG AGG ATA GTC AGT TTC GTT GGT GGC CTT GAT CTC TGT GAT GGA AGG   1285
Arg Arg Ile Val Ser Phe Val Gly Gly Leu Asp Leu Cys Asp Gly Arg
        355                 360                 365
TAT GAC ACT CAG TAC CAT TCT TTG TTT AGG ACA CTC GAC AGT ACC CAT   1333
Tyr Asp Thr Gln Tyr His Ser Leu Phe Arg Thr Leu Asp Ser Thr His
    370                 375                 380
CAT GAT GAC TTC CAC CAG CCA AAC TTT GCC ACT GCA TCA ATC AAA AAG   1381
His Asp Asp Phe His Gln Pro Asn Phe Ala Thr Ala Ser Ile Lys Lys
385                 390                 395                 400
GGT GGA CCT AGA GAG CCA TGG CAT GAT ATT CAC TCA CGG CTG GAA GGG   1429
Gly Gly Pro Arg Glu Pro Trp His Asp Ile His Ser Arg Leu Glu Gly
                405                 410                 415
CCA ATC GCA TGG GAT GTT CTT TAC AAT TTC GAG CAG AGA TGG AGA AAG   1477
Pro Ile Ala Trp Asp Val Leu Tyr Asn Phe Glu Gln Arg Trp Arg Lys
            420                 425                 430
CAG GGT GGT AAG GAT CTC CTT CTG CAG CTC AGG GAT CTG TCT GAC ACT   1525
Gln Gly Gly Lys Asp Leu Leu Leu Gln Leu Arg Asp Leu Ser Asp Thr
        435                 440                 445
ATT ATT CCA CCT TCT CCT GTT ATG TTT CCA GAG GAC AGA GAA ACA TGG   1573
Ile Ile Pro Pro Ser Pro Val Met Phe Pro Glu Asp Arg Glu Thr Trp
    450                 455                 460
AAT GTT CAG CTA TTT AGA TCC ATT GAT GGT GGT GCT GCT TTT GGG TTC   1621
Asn Val Gln Leu Phe Arg Ser Ile Asp Gly Gly Ala Ala Phe Gly Phe
```

```
465                 470                 475                 480
CCT GAT ACC CCT GAG GAG GCT GCA AAA GCT GGG CTT GTA AGC GGA AAG   1669
Pro Asp Thr Pro Glu Glu Ala Ala Lys Ala Gly Leu Val Ser Gly Lys
                485                 490                 495
GAT CAA ATC ATT GAC AGG AGC ATC CAG GAT GCA TAC ATA CAT GCC ATC   1717
Asp Gln Ile Ile Asp Arg Ser Ile Gln Asp Ala Tyr Ile His Ala Ile
            500                 505                 510
CGG AGG GCA AAG AAC TTC ATC TAT ATA GAG AAC CAA TAC TTC CTT GGA   1765
Arg Arg Ala Lys Asn Phe Ile Tyr Ile Glu Asn Gln Tyr Phe Leu Gly
        515                 520                 525
AGT TCC TAT GCC TGG AAA CCC GAG GGC ATC AAG CCT GAA GAC ATT GGT   1813
Ser Ser Tyr Ala Trp Lys Pro Glu Gly Ile Lys Pro Glu Asp Ile Gly
    530                 535                 540
GCC CTG CAT TTG ATT CCT AAG GAG CTT GCA CTG AAA GTT GTC AGT AAG   1861
Ala Leu His Leu Ile Pro Lys Glu Leu Ala Leu Lys Val Val Ser Lys
545                 550                 555                 560
ATT GAA GCC GGG GAA CGG TTC ACT GTT TAT GTT GTG GTG CCA ATG TGG   1909
Ile Glu Ala Gly Glu Arg Phe Thr Val Tyr Val Val Val Pro Met Trp
                565                 570                 575
CCT GAG GGT GTT CCA GAG AGT GGA TCT GTT CAG GCA ATC CTG GAC TGG   1957
Pro Glu Gly Val Pro Glu Ser Gly Ser Val Gln Ala Ile Leu Asp Trp
            580                 585                 590
CAA AGG AGA ACA ATG GAG ATG ATG TAC ACT GAC ATT ACA GAG GCT CTC   2005
Gln Arg Arg Thr Met Glu Met Met Tyr Thr Asp Ile Thr Glu Ala Leu
        595                 600                 605
CAA GCC AAG GGA ATT GAA GCG AAC CCC AAG GAC TAC CTC ACT TTC TTC   2053
Gln Ala Lys Gly Ile Glu Ala Asn Pro Lys Asp Tyr Leu Thr Phe Phe
    610                 615                 620
```

```
TGC TTG GGT AAC CGT GAG GTG AAG CAG GCT GGG GAA TAT CAG CCT GAA    2101
Cys Leu Gly Asn Arg Glu Val Lys Gln Ala Gly Glu Tyr Gln Pro Glu
625                 630                 635                 640
GAA CAA CCA GAA GCT GAC ACT GAT TAC AGC CGA GCT CAG GAA GCT AGG    2149
Glu Gln Pro Glu Ala Asp Thr Asp Tyr Ser Arg Ala Gln Glu Ala Arg
                645                 650                 655
AGG TTC ATG ATC TAT GTC CAC ACC AAA ATG ATG ATA GTT GAC GAT GAG    2197
Arg Phe Met Ile Tyr Val His Thr Lys Met Met Ile Val Asp Asp Glu
            660                 665                 670
TAC ATC ATC ATC GGT TCT GCA AAC ATC AAC CAG AGG TCG ATG GAC GGC    2245
Tyr Ile Ile Ile Gly Ser Ala Asn Ile Asn Gln Arg Ser Met Asp Gly
        675                 680                 685
GCT AGG GAC TCT GAG ATC GCC ATG GGC GGG TAC CAG CCA TAC CAT CTG    2293
Ala Arg Asp Ser Glu Ile Ala Met Gly Gly Tyr Gln Pro Tyr His Leu
    690                 695                 700
GCG ACC AGG CAA CCA GCC CGT GGC CAG ATC CAT GGC TTC CGG ATG GCG    2341
Ala Thr Arg Gln Pro Ala Arg Gly Gln Ile His Gly Phe Arg Met Ala
705                 710                 715                 720
CTG TGG TAC GAG CAC CTG GGA ATG CTG GAT GAT GTG TTC CAG CGC CCC    2389
Leu Trp Tyr Glu His Leu Gly Met Leu Asp Asp Val Phe Gln Arg Pro
                725                 730                 735
GAG AGC CTG GAG TGT GTG CAG AAG GTG AAC AGG ATC GCG GAG AAG TAC    2437
Glu Ser Leu Glu Cys Val Gln Lys Val Asn Arg Ile Ala Glu Lys Tyr
            740                 745                 750
TGG GAC ATG TAC TCC AGC GAC GAC CTC CAG CAG GAC CTC CCT GGC CAC    2485
Trp Asp Met Tyr Ser Ser Asp Asp Leu Gln Gln Asp Leu Pro Gly His
        755                 760                 765
CTC CTC AGC TAC CCC ATT GGC GTC GCC AGC GAT GGT GTG GTG ACT GAG    2533
```

Leu Leu Ser Tyr Pro Ile Gly Val Ala Ser Asp Gly Val Val Thr Glu

770 775 780

CTG CCC GGG ATG GAG TAC TTT CCT GAC ACA CGG GCC CGC GTC CTC GGC 2581

Leu Pro Gly Met Glu Tyr Phe Pro Asp Thr Arg Ala Arg Val Leu Gly

785 790 795 800

GCC AAG TCG GAT TAC ATG CCC CCC ATC CTC ACC TCA TAGACGAGGA AGCACT 2633

Ala Lys Ser Asp Tyr Met Pro Pro Ile Leu Thr Ser

805 810

ACACTACAAT CTGCTGGCTT CTCCTGTCAG TCCTTCTGTA CTTCTTCAGT TTGGTGGCGA 2693

GATGGTATGG CCGTTGTTCA GAATTTCTTC AGAATAGCAG TTGTTACAGT TGTGAATCAT 2753

AAAGTAATAA GTGCAGTATC TGTGCATGGT TGAGTTGGGA AGAAGATCGG GGATGCAATG 2813

ATGCTTGTGA AGTTGTGATG CCGTTTGTAA GATGGGAAGT TGGGAACTAC TAAGTAATTG 2873

GCATGATTGT ACTTTGCACT ACTGTTTAGC GTTGTTGATA CTGGTTAACC GTGTGTTCAT 2933

CTGAACTTGA TTCTTGATGC AGTTTGTGGC ATTACCAGTT TATCATCGTT CTTCAGGAAA 2993

AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAA 3040

SEQ ID NO.: 3

SEQUENCE LENGTH: 2799

SEQUENCE TYPE: Nucleic Acid

MOLECULE TYPE: Genomic DNA

ORIGINAL SOURCE

ORGANISM: *Oryza sativa*

SEQUENCE DESCRIPTION

CAAGGGTGTA CATAGATTTG TCTCGTAAAA TAGTATTATA ATATTATAAA CTTATTACTC 60

TATCCGTTCT AAAATATAAG AACCTTATGA CTGGATGGAA CATTTCCTAG TACTACGAAT 120

CTGAACACAT GTCTAGATTC ATAGTACTAG GAAATGTCTC ATCGCGGTAC TAGGTTCTTA 180

TATTTTAGGA TGGAGGGAGT TTAATATAAA ACTAATGGTT AGAACTTTGA AAGTTTGATT 240

TTAAATGTCA AATATTTATG GCTGGAGGTA GTATAATATG TTTTTTTTGG GACGTAGACT 300

AGGTAGTATA ATATGTTTGG TTGTGTTTAG ATCCAATATT TGGATCCAAA CTTCAGTCAT 360

```
TTTCCATCAC ATCAACTTGT CATATACACA TAACTTTTCA GTCACATCAT CCCCAATTTC  420
AACCAAAATC AAACTTTGCG CTGAACTAAA CACAACCTTT GGGCCCGTTT AGTTCCCCAA  480
TTTTTTTCCC AAAAACATCA CATCGAATCT TTGGACACAT GCATGAAGCA TTAAATATAG  540
ATAAAAAGAA AAACTAATTG CACAGTTATG GAGGAAATCG CGAGACGAAT CTTTTAAGCC  600
TAATTAGTCC GTGATTAGCC ATAAGTGCTA CAGTAACCCA ATTGTGCTAA TGACGGCTTA  660
ATTAGTCTCC ACAAGATTCG TCTCGCAGTT TCCAGGCGAG TTCTGAAATT AGTTTTTTCA  720
TTCGTGTCCG AAAACCCCTT CCGACATCCG GTCAAACGTT CGATATGACA CCCACAAATT  780
TTCTTTTCCC CAACTAAACA CACCCTTTAT CTCTTACCCT CTGGCTCTTT CAGTAGGCAT  840
ATCCAAGACA GCTGGTAATG CAGGCTCGGA CATAATTTGA CAGTTACGTT CATGTGACCG  900
ACGGTTGATG CTAGTGCAAC TGCAACATAC TGTTCAGATG GATGTCCCAA CGAGCTCAAA  960
ACAACTTAGG TGGCGCGTCG CGATTCATCA ATAACTCAAA TGGAAGCGCA AGTGCACGTA 1020
CGAAAATGAC AGCGAGTGAG GTGGCGAGCC TCACCTTGGT GATCCCAACC GGATAAGCTA 1080
TGCATCAGCC AGTTTCGTGG GGCTGCACAT TTCGTCGAAC ACCTGGAGTC CACGCCGCCG 1140
GCGACGTCGG CACAGCGCGC CCGCCCACCG CCCACGCACG CGCTTGACTC CACCCATGTT 1200
CTCCCTTCTC GACGCCCGCG AAGCCAGCGA ACCGATCCGA GGAAGTCAAG CCCCCACCGC 1260
CACTTGGACC GACCTCGGGA CGACGACGCC CCCGCGCTCT TCTAGACGCG CGGACGACGC 1320
GGGCGCTGGC TCCGCGACGC GACGTCGCGG TCATGGAGTA ACCGCGACGG ACAGATACTT 1380
CTACCCGTTT TTAACCTCGC CTCCTCCTCC TCCCGGCTCG AGATCCGTGG CCACGACGCG 1440
TGGTGGGAAA CCGGGAACGA CGTGCACGCA CGCACACAGG GCAAGTTTCA GTAGAAAAAT 1500
CGCCGGCATC CAGATCGGGA CAGTCTCTCT TCTCCCGCAA TTTTATAATC TCGCTCGATC 1560
CAATCTGCTC CCCTTCTTCT TCTACTCTCC CCATCTCGGC TCTCGCCATC GCCATCCTCC 1620
TCTCCCTTCC CGGAGAAGAC GCCTCCCTCC GCCGATCACC ACCCGGTAAG CCCAGTGTGC 1680
TTAGGCTAAG CGCACTAGAG CTTCTTGCTC GCTTGCTTCT TCTCCGCTCA GATCTGCTTG 1740
CTTGCTTGCT TCGCTAGAAC CCTACTCTGT GCTGCGAGTG TCGCTGCTTC GTCTTCCTTC 1800
CTCAAGTTCG ATCTGATTGT GTGTGTGGGG GGGCGCAGGT AGGGCGAGGA GGGAGCCAAA 1860
TCCAAATCAG CAGCC ATG GCG CAG ATG CTG CTC CAT GGG ACG CTG CAC GCC  1911
                Met Ala Gln Met Leu Leu His Gly Thr Leu His Ala
                1               5                  10
```

```
ACC ATC TTC GAG GCG GCG TCG CTC TCC AAC CCG CAC CGC GCC AGC GGA   1959
Thr Ile Phe Glu Ala Ala Ser Leu Ser Asn Pro His Arg Ala Ser Gly
        15                  20                  25
AGC GCC CCC AAG TTC ATC CGC AAG GTTCGGACCC TTCTCCTTAA TCTACTCGTC  2013
Ser Ala Pro Lys Phe Ile Arg Lys
    30                  35
TTTGCTCTTG CTCTTTTTCT TTTGTGTCCC TTTCTTGTGT GTGCGTTTGC ATGAGCCCGA 2073
ATTTGATCTG CTAGTGCACA GTACAGTCAG ATACACTGAA ACGATCTGGA AATTCTGGAT 2133
TATTAGGAAA AATAAAGAGG TAGTAGACAA GAATTGGAGA TACTTTCTAT CAAGATTGGT 2193
CTATTATGCT TGGCCATTTC TTGTTTGACC CAAGTACTTC TTTGAATCTA GAGTTTGCTG 2253
TGTGTGATGT GGTGTGTGTT TGTGTCACCA AAAATCTTCA TTAGCTAAAA CTGAAATTTT 2313
ATTTATTAAC TGACCTACTA AAAATGTAGA GTTCTCTGTG TGTGATGTGT GCTTGTGTCA 2373
CCAAAAATCT TGATTTGATA GAGTTTTTAT TTATTTATTA ACTGACCTAC TACAAATCTA 2433
TTGCTGTATG CTATGTGTGT CTGTATCACC TGAAATGCAA TGTCTTCTTC TTTGTTGTTC 2493
TTGATCTAAC ACGTGAGCTC ATGTCAACAG TTT GTG GAG GGG ATT GAG GAC ACT  2547
                                Phe Val Glu Gly Ile Glu Asp Thr
                                        40
GTG GGT GTC GGC AAA GGC GCC ACC AAG GTG TAT TCT ACC ATT GAT CTG   2595
Val Gly Val Gly Lys Gly Ala Thr Lys Val Tyr Ser Thr Ile Asp Leu
45                  50                  55                  60
GAG AAA GCT CGT GTA GGG CGA ACT AGG ATG ATA ACC AAT GAG CCC ATC   2643
Glu Lys Ala Arg Val Gly Arg Thr Arg Met Ile Thr Asn Glu Pro Ile
65                  70                  75
AAC CCT CGC TGG TAT GAG TCG TTC CAC ATC TAT TGC GCT CAT ATG GCT   2691
Asn Pro Arg Trp Tyr Glu Ser Phe His Ile Tyr Cys Ala His Met Ala
80                  85                  90
TCC AAT GTG ATC TTC ACT GTC AAG ATT GAT AAC CCT ATT GGG GCA ACG   2739
Ser Asn Val Ile Phe Thr Val Lys Ile Asp Asn Pro Ile Gly Ala Thr
```

95 100 105

AAT ATT GGG AGG GCT TAC CTG CCT GTC CAA GAG CTT CTC AAT GGA GAG 2787

Asn Ile Gly Arg Ala Tyr Leu Pro Val Gln Glu Leu Leu Asn Gly Glu

110 115 120

GAG ATT GAC AGA 2799

Glu Ile Asp Arg

125

SEQ ID NO.: 4

SEQUENCE LENGTH: 173

SEQUENCE TYPE: Nucleic Acid

MOLECULE TYPE: Genomic DNA

ORIGINAL SOURCE

ORGANISM: *Oryza sativa*

SEQUENCE DESCRIPTION

GTAAGCCCAG TGTGCTTAGG CTAAGCGCAC TAGAGCTTCT TGCTCGCTTG CTTCTTCTCC 60

GCTCAGATCT GCTTGCTTGC TTGCTTCGCT AGAACCCTAC TCTGTGCTGC GAGTGTCGCT 120

GCTTCGTCTT CCTTCCTCAA GTTCGATCTG ATTGTGTGTG TGGGGGGGCG CAG 173

CLAIMS

1. An isolated DNA fragment having a nucleotide sequence shown in SEQ ID NO. 1 in Sequence Listing or having a nucleotide sequence which hybridizes with said nucleotide sequence shown in SEQ ID NO. 1 when the hybridization is carried out in a hybridization solution containing 0.5M sodium phosphate buffer (pH7.2), 7% SDS, 1 mM EDTA and 100 µg/ml of salmon sperm DNA, at 65°C for 16 hours, and then washing the product twice with a washing solution containing 0.5xSSC and 0.1% SDS at 65°C for 20 minutes each, the latter nucleotide sequence having a function to promote expression of a gene downstream thereof in a plant.

2. The DNA fragment according to claim 1, which has a nucleotide sequence shown in SEQ ID NO. 1 in Sequence Listing.

3. An isolated DNA fragment having a nucleotide sequence shown in SEQ ID NO.4 in Sequence Listing or having a nucleotide sequence which hybridizes with said nucleotide sequence shown in SEQ ID NO. 4 when the hybridization solution containing 0.5M sodium phosphate buffer (pH7.2), 7% SDS, 1mM EDTA and 100 µg/ml of salmon sperm DNA, at 65°C for 16 hours, and then washing the product twice with a washing solution containing 0.5xSSC and 0.1% SDS at 65°C for 20 minutes each, the latter nucleotide sequence having a function to promote expression of a gene downstream thereof in a plant.

4. The DNA fragment according to claim 3, which has a nucleotide sequence shown in SEQ ID NO. 4 in Sequence Listing.

5. A recombinant vector comprising said DNA fragment according to claim 1 and a foreign gene to be expressed, which is operably linked to said DNA fragment at a downstream region of said DNA fragment.

6. The recombinant vector according to claim 5, wherein said DNA fragment has a nucleotide sequence shown in SEQ ID NO. 1 in Sequence Listing.

7. A recombinant vector comprising said DNA fragment according to claim 3 and a foreign gene to be expressed, which is operably linked to said DNA fragment at a downstream region of said DNA fragment.

8. The recombinant vector according to claim 7, wherein said DNA fragment has a nucleotide sequence in SEQ ID No.4 in Sequence Listing.

9. A method for expressing a foreign gene comprising introducing said recombinant vector according to claim 5, into host plant cells and expressing said foreign gene.

10. The method according to claim 9, wherein said DNA fragment has a nucleotide sequence shown in SEQ ID No.1 in Sequence Listing.

11. A method for expressing a foreign gene comprising introducing said recombinant vector according to claim 7, into host plant cells and expressing said foreign gene.

12. The method according to claim 11, wherein said DNA fragment has a nucleotide sequence shown in SEQ ID No.4 in Sequence Listing.

2191624

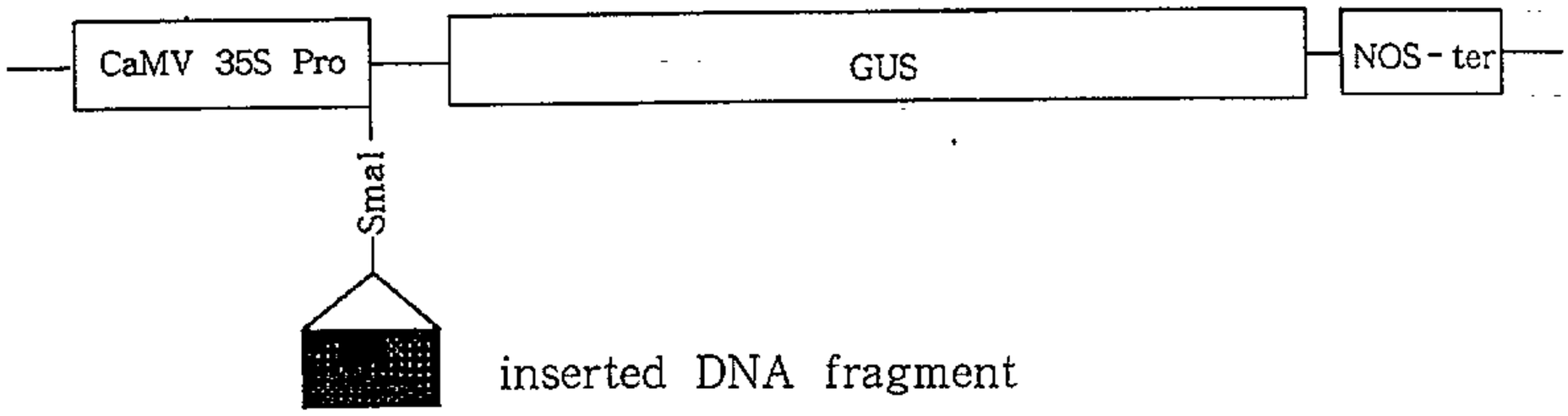
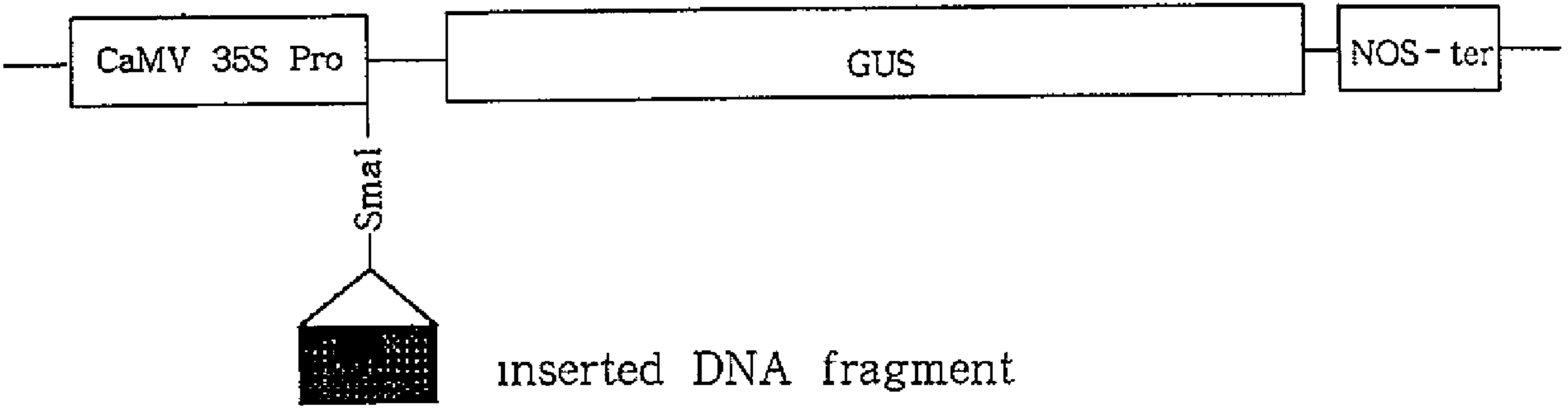


Fig. 1


CaMV 35S Pro
GUS
NOS - ter
SmaI
inserted DNA fragment