



(72) KAKITANI, MAKOTO, JP

(72) UMEMOTO, NAOYUKI, JP

(72) TSUKAHARA, MASAYOSHI, JP

(72) ISHIDA, ISAO, JP

(71) KIRIN BEER KABUSHIKI KAISHA, JP

(51) Int.Cl.⁷ C12N 15/29, C12N 15/82, A01H 5/00

(30) 1997/06/18 (9/161726) JP

(54) **PLANTES RESISTANTES AUX MOISSURES ET PROCEDE DE
PRODUCTION DESDITES PLANTES**

(54) **MOLD-RESISTANT PLANTS AND METHOD OF
CONSTRUCTION OF SAME**

(57) On décrit des ADN codant des protéines chimères qui présentent chacune un domaine contenant au moins le site de liaison avec l'éléciteur d'un récepteur de l'éléciteur, et un autre domaine contenant au moins le motif de transduction de signal d'une protéine capable de conférer à des plantes la résistance à la maladie. On décrit en outre des vecteurs recombinants contenant ces ADN, l'utilisation de ces vecteurs ou de leurs descendants pour préparer des plantes transformées, ainsi qu'une méthode de production de plantes résistantes à la maladie qui se caractérise par un transfert desdits vecteurs dans des plantes hôtes.

(57) DNAs encoding chimeric proteins each having a domain containing at least the elicitor-binding site of an elicitor receptor and another domain containing at least the signal transduction motif of a protein capable of imparting disease resistance to plants; recombinant vectors containing these DNAs; transformed plants prepared by using these vectors or offsprings thereof; and a method for the construction of disease resistant plants characterized by transferring the above vectors into plant hosts.

ABSTRACT

A DNA coding for a chimeric protein comprising a domain containing at least an elicitor binding site of an elicitor receptor and a domain containing at least a signal transduction motif of a protein which confers plant disease resistance; a recombinant vector comprising the above DNA; a transformed plant transformed with the recombinant vector or the progeny of the plant; and a method for creating a plant disease resistant plant comprising transferring the recombinant vector into a host plant.

~~FILE, PIN IN THIS AMENDED~~
~~TEXT TRANSLATION~~

DESCRIPTION

FUNGUS-RESISTANT PLANTS AND METHOD FOR CREATING THE SAME

TECHNICAL FIELD

The present invention relates to a DNA coding for a protein which confers plant disease resistance, plant disease-resistant plants and a method for creating such plants. More specifically, the invention relates to a DNA coding for a chimeric protein (chimeric receptor) comprising a domain containing at least an elicitor binding site of an elicitor receptor and a domain containing a signal transduction motif of a plant-derived resistance gene; plant disease-resistant plants into which the above DNA is transferred; and a method for creating such plants.

BACKGROUND ART

It is known that plants synthesize and accumulate antibiotic agents called phytoalexins in response to infection with pathogenic microorganisms [M. Yoshikawa (1978) Nature 257: 546]. Substances which induce such resistance reactions have been found in plant *Phytophthora* [N.T. Keen (1975) Science 187: 74] and are called "elicitors".

Various elicitors have been known, including degradation products generated by plant- or fungus-derived hydrolytic enzymes, proteinaceous elicitors produced by fungi and Avr (avirulence) gene products generated by fungi. As examples of the plant-derived degradation products, polygalacturonates produced by polygalacturonases may be given. As examples of the fungus-derived degradation products, oligosaccharides of chitin

(N-acetylchitooligosaccharides) produced by chitinases or glucans produced by glucanases may be given. As examples of the proteinaceous elicitors, elicitors derived from fungi belonging to the genus *Phytophthora* may be given. As examples of the Avr gene products, Avr4 and Avr9 derived from *Cladosporium fulvum* or *nip1* derived from *Rhynchosporium secalis* may be given. These elicitors perform specific interactions with plant resistance gene products including elicitor receptors on plant plasma membranes. Through such interactions, invasion of fungi, etc. is recognized and then signals are generated. These signals are transmitted to nuclei to cause various resistance reactions for preventing the invasion of fungus into plants. These reactions are called the hypersensitive response, and include cell death, production of hydrogen peroxide, expression of PR proteins, expression of lipoxygenases and production of resistant agents called phytoalexins [W. Knogge (1996) Plant Cell 8: 1711; A.F. Bent (1996) Plant Cell 8: 1757; K.E. Hammond-Kosack and J.D.G. Jones (1996) Plant Cell 8: 1773].

The biochemical process from the infection of plants with pathogens to the induction of resistance in the plants is believed to proceed as follows, taking the synthesis and accumulation of phytoalexin for example. First, when mycelia of pathogens invade plant cells, glucanases in plant cells work so as to cleave polysaccharides on the surface of the mycelial walls of pathogens, thereby liberating elicitors. Elicitors are polysaccharides composed of glucose. They are β -1, 6-linked or β -1, 3-linked glucans [J.K. Sharp et al. (1984) J. Biol. Chem. 259: 11321; M. Yoshikawa (1990) Plant Cell Engineering, 1.2: 695; Y. Okinaka et al. (1995) Plant Physiol. 109: 839]. If such

an elicitor binds to a receptor in plant cells, a second messenger which performs the role of signal transduction is produced. This signal transducing substance is incorporated into the nuclei of plant cells, and activates the transcription of those genes coding for phytoalexin synthesis enzymes to thereby induce phytoalexin synthesis. At the same time, the phytoalexin degradation system is inhibited. Thus, phytoalexins are efficiently accumulated in plant cells. In the case of soybean, for example, a phytoalexin playing an important role in the resistance of this plant is called glyceollin, the structure of which has been determined [M. Yoshikawa et al. (1978) *Physiol. Plant. Pathol.* 12: 73].

Therefore, a receptor specific to a glucan elicitor derived from a soybean pathogenic filamentous fungus *Phytophthora megasperma* f. sp. *glycinea* is believed to be a receptor protein which plays an important role in the synthesis and accumulation of the antitumor agent glyceollin. This glucan elicitor specific receptor has been isolated and purified from a plasma membrane fraction of soybean roots by Kakitani et al. (Japanese Unexamined Patent Publication No. 6-321995). Further, Umemoto et al. [N. Umemoto et al. (1997) *Proc. Natl. Acad. Sci. USA* 94: 1029] have cloned a gene coding for this glucan elicitor receptor.

Recently, a number of reports have been made on the cloning of plant-derived resistance genes. In these genes, characteristic motifs or domains involved in signal transduction (such as leucine-rich repeats, leucine zippers, nucleotide binding sites (P-loop) and serine/threonine kinase domains) are found in part or in their entirety. As plant-derived disease

resistance genes which are expected to contain such motifs or domains involved in signal transduction, the following genes have been reported for example: *Arabidopsis thaliana*-derived resistance gene *RPS2* [M. Mindrinos et al. (1994) Cell 78: 1089; A.F. Bent et al. (1994) Science 265: 1856], *Arabidopsis thaliana*-derived resistance gene *RPM1* [M.R. Grant et al. (1995) Science 269: 843], tobacco-derived resistance gene *N* [S. Whitham et al. (1994) Cell 78: 1101], flax-derived resistance gene *L6* [G.J. Lawrence et al. (1995) Plant Cell 7: 1195], tomato-derived resistance gene *Pto* [G.B. Martin et al. (1993) Science 262: 1432], tomato-derived resistance gene *Prf* [J.M. Salmeron et al. (1996) Cell 86: 123], tomato-derived resistance gene *Cf-9* [D.A. Jones et al. (1994) Science 266: 789], tomato-derived resistance gene *Cf-2* [M.S. Dixon et al. (1996) Cell 84: 451], and rice-derived resistance gene *Xa21* [W.-Y. Song et al. (1995) Science 270: 1804].

On the other hand, in contrast to plant cell-derived receptors, a large number of reports have been made on the cloning of animal cell-derived receptors. Also, a large number of researches into chimeric receptors composed of an extracellular domain and an intracellular domain derived from heterologous receptors have been made. For example, an extracellular domain which recognizes a ligand of a certain receptor is replaced with an extracellular domain of a receptor which recognizes a ligand other than the above ligand, and the latter, new extracellular domain is fused to an intracellular domain which transmits a signal to nuclei, thereby preparing a chimeric receptor. When this receptor is expressed in cells, it is reported that the chimeric receptor recognizes the ligand

which the new extracellular domain recognizes, to thereby perform signal transduction in cells [H. Riedel et al. (1986) Nature 324: 68, J. Lee et al. (1989) EMBO J. 8: 167; H. Lehtvaslainen et al. (1989) EMBO J. 8: 159, S. Lev et al. (1990) Mol. Cell. Biol. 10: 6064]. Further, it is also reported that even if a gene coding for a receptor of an animal cell is transferred into a heterologous animal cell and expressed therein, the gene functions in a similar manner; for example, it is reported that when human-derived GM-CSF receptor is expressed in mouse 3T3 cells, the transmission of growth promoting signal is observed [S. Watanabe et al. (1993) Mol. Cell. Biol. 13: 1440].

However, as far as the present inventors know, no report has been ever made concerning chimeric reporters in plants. Besides, no report has suggested that such chimeric receptors function in plants.

DISCLOSURE OF THE INVENTION

The glucan elicitor receptor described above has no characteristic motifs or domains involved in signal transduction as found in plant-derived resistance genes. Thus, a possibility has been suggested that this glucan elicitor receptor may lack signal transduction domains. From researches into chimeric receptors using animal cell-derived receptors, the present inventors have considered as follows: if a chimeric receptor obtained by fusing a domain recognizing a certain ligand to a domain involved in signal transduction derived from a different protein (including receptor) is transferred into plant cells and expressed therein, the ligand may be recognized and signal

transduction may be performed. The inventors have also considered that even if the cell in which a chimeric receptor is to be expressed and the cell from which a domain involved in signal transduction has been derived do not belong to the same plant species, signal transduction may be performed in a similar manner. In other words, the inventors have considered that if a gene coding for a chimeric receptor in which an elicitor receptor is bound to an intracellular signal transduction domain of a plant-derived receptor; or a gene coding for a chimeric receptor in which an elicitor receptor is bound to a domain that is encoded by a plant-derived resistance gene (e.g. *Arabidopsis thaliana*-derived resistance gene *RPS2* or *RPM1*, tobacco-derived resistance gene *N*, flax-derived resistance gene *L6*, tomato-derived resistance gene *Pto*, rice-derived resistance gene *Xa21*) and that is expected to be involved intracellular signal transduction can be synthesized and expressed in plant regardless of plant species, the efficiency in signal transduction may be increased compared to those cases where an elicitor receptor alone is expressed. As a result, it would become possible to create plants having clearly strong resistance to pathogenic fungi, and the productivity of agricultural products would be improved. Further, economical effects from reduced use of agricultural chemicals and reduction of environmental pollution could be expected.

It is the object of the invention to provide a DNA coding for a chimeric protein comprising a domain containing at least an elicitor binding site of an elicitor receptor and a domain containing a signal transduction motif of a plant-derived resistance gene; plant disease-resistant plants into which the

above DNA is transferred; and a method for creating such plants.

As a result of intensive and extensive researches toward the solution of the above-described problem, the present inventors have succeeded in conferring disease resistance on plants by synthesizing a gene coding for a chimeric receptor composed of a domain which is encoded by a plant-derived resistance gene and expected to be involved in signal transduction and a domain containing at least an elicitor binding site of an elicitor receptor, transferring the gene into *Arabidopsis* plants and tobacco plants, and expressing the gene in these plants. Thus, the present invention has been achieved.

The present invention relates to a DNA coding for a chimeric protein comprising a domain containing at least an elicitor binding site of an elicitor receptor and a domain containing at least a signal transduction motif of a protein which confers plant disease resistance. This chimeric protein functions as a receptor, thereby conferring disease resistance on plants.

Specific examples of the above elicitor include a glucan, polygalacturonate, N-acetylchitosaccharide, elicitin, *Cladosporium fulvum*-derived Avr gene product, or *Rhynchosporium secalis*-derived niple gene product. Specific examples of the above elicitor receptor include receptors of the above-enumerated elicitors. As the elicitor binding site of a glucan elicitor receptor, a protein comprising at least the amino acid sequence shown in SEQ ID NO: 27, or a protein which comprises the amino acid sequence shown in SEQ ID NO: 27 having deletion, substitution, addition or insertion of at least one amino acid and which has ability to bind to the relevant glucan elicitor

may be given. As the domain containing a signal transduction motif, a domain containing at least one motif selected from the group consisting of leucine-rich repeats, leucine zippers, nucleotide binding sites and serine/threonine kinase domains may be given. Specific examples of such a domain include a signal transduction domain of an expression product of tomato-derived *Pto*, *Prf*, *Cf-2* or *Cf-9* gene; a signal transduction domain of an expression product of rice-derived *Xa21* gene; a signal transduction domain of an expression product of *Arabidopsis thaliana*-derived *RPS2* or *RPM1* gene (e.g. the domain represented by SEQ ID NO: 29); a signal transduction domain of an expression product of flax-derived *L6* gene; and a signal transduction domain of an expression product of tobacco-derived *N* gene.

The present invention further relates to a DNA coding for a chimeric protein comprising the amino acid sequence shown in SEQ ID NO: 5, 7 or 9; or a chimeric protein which comprises the amino acid sequence shown in SEQ ID NO: 5, 7 or 9 having deletion, substitution, addition or insertion of at least one amino acid and which confers plant disease resistance. Specific examples of the DNA include DNAs comprising the nucleotide sequence shown in SEQ ID NO: 6, 8 or 10.

The present invention further relates to a recombinant vector comprising the above-described DNA.

The present invention further relates to a transformed plant transformed with the above recombinant vector, or the progeny of the plant.

The present invention further relates to a plant having plant disease resistance, the plant having the above DNA and expressing the DNA, or the progeny of the plant.

The present invention further relates to a method for creating a plant disease resistant plant, comprising transferring the above recombinant vector into a plant.

The present invention further relates to a method for creating a plant having plant disease resistance or the progeny of this plant, comprising incorporating the above DNA into a chromosome of a plant and expressing the DNA.

Hereinbelow, the present invention will be described in detail.

The DNA of the invention can be obtained by ligating a DNA coding for a domain containing at least an elicitor binding site of an elicitor receptor to a DNA coding for a domain containing at least a signal transduction motif of a protein which confers plant disease resistance. The DNA of the invention is sometimes called "the DNA coding for a chimeric receptor" or "the chimeric DNA".

Specific examples of the elicitor receptor include receptors which bind to plant degradation product elicitors such as polygalacturonates or fungal cell wall degradation product elicitors such as glucans, N-acetylchitooligosaccharides, and receptors which bind to proteinaceous elicitors such as elicitors produced by fungi, and receptors which bind to such elicitors as *Cladosporium fulvum*-derived Avr gene products or *Rhynchosporium secalis*-derived niple gene products.

As a method for cloning cDNAs coding for these elicitor receptors (hereinafter, referred to as "ERDNAs", the method described below may be used.

Briefly, the above elicitor is labelled with an isotope.

Then, an elicitor receptor which specifically binds to the elicitor is purified from plant plasma membrane fractions by determining binding activities of the fractions with the labelled elicitor. Based on the partial amino acid sequence information obtained from the purified elicitor receptor, PCR primers are designed. Using these primers and a plant cell-derived DNA as a template, a PCR is performed to thereby obtain a probe. With this probe, a cDNA of interest is cloned from a cDNA library.

On the other hand, the DNA coding for a signal transduction motif in the present invention is already known as described above. Thus, the DNA can be cloned by PCR, etc.

Subsequently, ERDNA is ligated to the DNA coding for a signal transduction motif. The ligation of these DNAs are performed by digesting these DNAs with appropriate restriction enzymes and then ligating with ligase. Alternatively, the ligation is performed by binding ERDNA to the DNA coding for a signal transduction motif with an appropriate oligonucleotide. Alternatively, the above DNAs may be ligated by insertion into an appropriate plasmid. The relative locations of ERDNA and the DNA coding for a signal transduction motif at the time of ligation are not particularly limited, i.e. ERDNA may be located upstream of the DNA coding for a signal transduction motif or vice versa.

Finally, the nucleotide sequence of the resultant chimeric DNA is determined by conventional methods (e.g. the dideoxy method, Maxam-Gilbert method). Usually, the nucleotide sequence is analyzed with an automated DNA sequencer.

Fig. 1 shows schematic representations of the chimeric DNA

of the invention which is composed of, by way of example, a DNA coding for a glucan elicitor receptor (ER) and *RPS2* gene from *Arabidopsis thaliana*. In Fig. 1, CER is a chimeric DNA composed of a 5' end region of *RPS2* gene corresponding to about one third (1/3) of *RPS2* (SEQ ID NO: 29) and the entire ERDNA (SEQ ID NO: 2). IER and ISER in Fig. 1 have a truncated ERDNA. In these chimeric DNAs, ERDNA is located upstream of *RPS2* gene (i.e. their locations are exchanged as compared to the locations in CER).

The elicitor binding site of the glucan elicitor receptor corresponds to the partial sequence from position 239 to position 442 of the amino acid sequence shown in SEQ ID NO: 1, and has the amino acid sequence shown in SEQ ID NO: 27. As long as the elicitor binding site is a protein comprising at least the amino acid sequence shown in SEQ ID NO: 27, or a protein which comprises the amino acid sequence shown in SEQ ID NO: 27 having deletion, substitution, addition or insertion of at least one amino acid and which has ability to bind to the relevant glucan elicitor, the length of the elicitor receptor is not particularly limited; the elicitor receptor may be in a truncated form (see "2-11", "2-16" and "1-5" in Fig. 1). As a DNA coding for the amino acid sequence of the elicitor binding site represented by SEQ ID NO: 27, the nucleotide sequence shown in SEQ ID NO: 28 may be given.

The nucleotide sequence of the above chimeric DNA and the amino acid sequence of a chimeric protein encoded by the DNA are shown in SEQ ID NOS: 5 and 6, respectively, for CER; in SEQ ID NOS: 7 and 8, respectively, for IER; and in SEQ ID NOS: 9 and 10, respectively, for ISER. The amino acid sequences (SEQ ID NOS: 5,

7 and 8) may have mutation such as deletion, substitution, addition or insertion of at least 1 amino acid, as long as these chimeric proteins have binding function as elicitor receptors and confer disease resistance activity on plants. In addition to those nucleotide sequences coding for the amino acid sequences comprised in the chimeric proteins of the invention, DNAs coding for the same polypeptides differing only in degenerate codons (termed "degenerate isomers") are also included in the DNA of the invention. For example, a DNA in which a codon AAC corresponding to Asn is changed to a degenerate codon (e.g. AAT) is called herein a degenerate isomer.

The introduction of the above mutation may be performed by site specific mutagenesis (see, for example, D.F. Mark et al., Proc. Natl. Acad. Sci. USA 81: 5662-5666, 1984; PCT WO 85/00817 published on Feb. 28, 1985; R.P. Wharton et al., Nature 316: 601-605, Aug. 15, 1985) which was a well-known technique before the filing of the present application. Once the above nucleotide sequence is determined, the DNA of the invention can be obtained by chemical synthesis, PCR or hybridization using a DNA fragment having the nucleotide sequence as a probe.

The DNA sequence coding for a chimeric receptor used in the invention preferably has at least one stop codon (e.g. TAG) adjacent to its 3' end. Also, if desired, the DNA sequence may optionally have an ATG codon for initiation methionine upstream of its 5' end, the reading frame of which codon coincides with the reading frame of the DNA sequence. Further, the DNA sequence may have other DNAs of appropriate lengths upstream of its 5' end and downstream of its 3' end as non-translated regions.

Typically, the DNA sequence coding for a chimeric receptor

used in the invention is present being inserted into plasmid or phage DNA as a part of the constituents thereof; or is present in microorganisms (in particular, bacteria), phage particles or plants while being inserted into plasmid, phage or genomic DNA as a part of the constituents thereof. Specific examples of the above bacteria include *E. coli* and *Agrobacterium*. For stable expression of the DNA sequence coding for a chimeric receptor in plants, a promoter, DNA coding for translation initiation codon (ATG) and terminator may be added to the DNA sequence in an appropriate combination.

The DNA coding for a chimeric receptor of the invention (hereinafter sometimes referred to as the "chimeric DNA") may be incorporated into an appropriate vector to construct the recombinant vector of the invention.

The vector into which the chimeric DNA of the invention is to be incorporated is not particularly limited as long as it is replicable in a host. For example, plasmid DNA, phage DNA and the like may be employed.

A binary vector, which is one of plasmid DNA, can be prepared from *E. coli* or *Agrobacterium* by, for example, alkali extraction (Birnboim, H.C. & Doly, J. (1979) *Nucleic Acid Res.* 7: 1513). Specific examples of binary vectors include pBI121 and pBI101. Alternatively, a commercial plasmid such as pUC118 (Takara Shuzo), pUC119 (Takara Shuzo), pBluescript SK+ (Stratagene) or pGEM-T (Promega) may be used.

Specific examples of phage DNA include M13mp18, M13mp19 and M13tv18.

In order to ligate a chimeric DNA to a vector, a method may be used in which a purified chimeric DNA of the invention is

digested with appropriate restriction enzymes and inserted into the relevant restriction site or multicloning site of an appropriate vector DNA for ligation.

The chimeric DNA to be expressed should be incorporated into the vector so that the function of the DNA can be revealed. For this purpose, signal peptide genes, terminators, drug resistance genes, etc. may be incorporated into the recombinant vector of the invention in addition to promoters and the above-described gene. As signal peptide genes useful in the invention, the signal peptide gene of β -1,3 glucanase derived from tomato (*Lycopersicon esculentum*) or the like may be given. As enhancers useful in the invention, tobacco mosaic virus Ω sequence or the like may be given. As drug resistance genes useful in the invention, kanamycin resistance gene, hygromycin resistance gene, or the like may be given.

Specific examples of the above promoter include the promoter of the gene encoding ribulose-1,5-biphosphate carboxylase small subunit (R. Fluhr et al., Proc. Natl. Acad. Sci. USA (1986) 83: 2358), the promoter of nopaline synthase (NOS) gene (W.H.R. Langridge et al., Plant Cell Rep. (1985) 4: 355), the promoter generating cauliflower mosaic virus 19S-RNA (H. Guilley et al., Cell (1982) 30: 763), and the promoter generating cauliflower mosaic virus 35S-RNA (J.T. Odell et al., Nature (1985) 313: 810). Specific examples of the above terminator include the terminator of nopaline synthase gene [A. Depicker et al., J. Mol. Appl. Gen. (1982) 1: 561] and the terminator of octopine synthase gene [J. Gielen et al., EMBO J. (1984) 3: 835].

Specifically, as shown in Fig. 2, cauliflower mosaic virus

35S promoter (p35S) is ligated upstream of the chimeric DNA, and nopaline synthase gene terminator (NosT) is ligated downstream of the chimeric DNA (this construct is called "ER cassette"). Further, a cassette in which kanamycin resistance gene (NPT II) is ligated between nopaline synthase gene promoter and terminator ("NPT II cassette") and a cassette in which hygromycin resistance gene is ligated between p35S and NosT ("Hm cassette") are constructed. Then, NPT II cassette is ligated upstream of ER cassette, followed by ligation of Hm cassette downstream of ER cassette to thereby obtain the recombinant vector of the invention (Fig. 2).

From various researches into elicitors and their receptors, it has been suggested that, in the case of glucan elicitors and their receptors for example, glucan elicitor receptors play an important role in plant resistance to a very wide range of fungi containing glucans in their cell wall components. Further, a gene coding for a glucan elicitor receptor has been cloned and, as a result of analysis of its primary sequence, it has been suggested that the elicitor receptor lacks domains involved in signal transduction. On the other hand, plant-derived resistance genes so far reported are roughly classified into the following 4 classes based on their structures [M.S. Dixon et al. (1996) Cell 84: 451].

Class 1: Genes of this class have leucine-rich repeats and nucleotide binding sites. Some of them also have leucine zippers. These genes are believed to be localized probably in cytoplasm. Specific examples of resistance genes falling under this class include tobacco-derived resistance gene *N*, *Arabidopsis thaliana*-

derived resistance genes *RPS2* and *RPM1*, flax-derived resistance gene *L6*, and tomato-derived resistance gene *Prf*.

Class 2: Genes of this class have leucine-rich repeats and membrane binding sites. These genes are believed to be localized probably in cell surfaces. Tomato-derived resistance genes *Cf-2* and *Cf-9*, and the like fall under this class.

Class 3: Genes of this class have serine/threonine kinase domains. Tomato-derived resistance gene *Pto*, and the like fall under this class.

Class 4: Genes of this class have leucine-rich repeats, transmembrane domains and serine/threonine-like kinase domains. Probably, it is believed that the leucine-rich repeats recognize elicitors to thereby activate serine/threonine kinases located in cells. Rice-derived resistance gene *Xa-21* and the like fall under this class.

Thus, plant-derived resistance genes contain one or more motifs or domains which are expected to be involved in signal transduction, i.e. leucine-rich repeats, leucine zippers, nucleotide binding sites and serine/threonine kinase domains.

Therefore, in the present invention, a DNA coding for a chimeric protein is prepared by ligating a domain containing at least an elicitor binding site of an elicitor receptor to a domain involved in signal transduction encoded by a plant-derived disease resistance gene (e.g. *RPS2* gene, *RPM1* gene, *N* gene, *L6* gene, *Pto* gene, *Xa21* gene, *Cf-2* gene, *Cf-9* gene, *Prf* gene), and then expressed.

For example, a chimeric DNA composed of a DNA coding for a domain containing at least an elicitor binding site of a glucan elicitor receptor and a DNA coding for a domain containing a

leucine zipper and a nucleotide binding site of *Arabidopsis thaliana*-derived resistance gene (*RPS2*), or a fragment of the chimeric DNA, is transferred into higher plants and expressed therein by known methods.

Consequently, it is possible to confer on plants still stronger resistance to fungi, as compared to those cases where a glucan elicitor receptor alone is introduced and expressed.

A theory has been proposed that those fungi infectious to plants generally have suppressors and, thus, have acquired ability to suppress plants' inherent resistance to fungi. Even in these cases, however, it is possible to create a plant having stronger resistance to fungi by transferring the DNA of the invention coding for a chimeric receptor or a fragment of the DNA into the plant and expressing it therein so that the chimeric receptor functions, or by modifying the DNA or regulating the amount of expression of the DNA.

Further, by transferring the DNA of the invention or a fragment thereof into higher plants in combination with a gene or character which enhance resistance to fungi (e.g. a glucanase which exhibits resistance to fungi), it is possible to confer on plants still stronger resistance to fungi than the resistance of those plants into which a glucanase or the like alone has been introduced.

Examples of DNA sequences coding for glucanases include a DNA comprising a nucleotide sequence coding for a protein comprising the amino acid sequence shown in SEQ ID NO: 11 or 13, or protein which comprises the amino acid sequence shown in SEQ ID NO: 11 or 13 having deletion, substitution, addition or insertion of at least one amino acid and which has glucanase

activity. The above-described DNA is intended to include all the degenerate isomers. Specific examples of such isomers include a DNA comprising the nucleotide sequence shown in SEQ ID NO: 12 or 14.

As methods for transferring a gene into plant cells or plant, conventional methods may be used. For example, methods described in "Plant Genetic Transformation and Gene Expression; A Laboratory Manual", J. Draper, et al. (eds.), Blackwell Scientific Publications (1988) may be used. Examples of methods for gene transfer include biological methods using viruses or *Agrobacterium* and physicochemical methods such as electroporation, the polyethylene glycol method, microinjection, the particle gun method and the dextran method.

When the plant to be transformed is a dicotyledonous plant, a biological method using *Agrobacterium* is generally preferable. When the plant to be transformed is a monocotyledonous plant, or a dicotyledonous plant that is not susceptible to infection with *Agrobacterium*, a physical/chemical method such as electroporation is preferable. As a plant material into which a DNA of interest is to be transferred, an appropriate material may be selected from leaves, stems, roots, petals, tubers, protoplasts, calli, pollen, seed embryos, shoot primordia, etc. according to the method of transfer.

When a DNA of interest is to be transferred into cultured plant cells, protoplasts are generally used as a material, and the DNA is transferred therein by a physical/chemical method such as electroporation, the polyethylene glycol method or the like. On the other hand, when a DNA of interest is to be transferred into plant tissues, leaves, stems, roots, petals,

tubers, protoplasts, calli, pollen, seed embryos, shoot primordia or the like are used as a plant material. Preferably, leaves or stems are used. The DNA is transferred into such plant tissues by a biological method using a virus or *Agrobacterium*, or a physical/chemical method such as the particle gun method, microinjection or the like. Preferably, a biological method using *Agrobacterium* is used.

In order to regenerate a plant from those plant tissues or plant cells into which a DNA sequence coding for a chimeric receptor has been transferred, these transformed plant tissues or cells may be cultured in a medium, such as MS medium, containing appropriate plant growth regulators. The resultant seedlings which are rooting may be transferred to soil to give grown-up plants.

By transferring and expressing a DNA sequence coding for a chimeric receptor by the procedures as described above, resistance to pathogenic fungi can be conferred on plants or enhanced in plants. Examples of such plants include those plants which are susceptible to infection with pathogenic fungi containing glucans in their cell walls. Specifically, these plants include, but are not limited to, solanaceous plants, leguminous plants, asteraceous plants, caryophyllaceous plants and gramineous plants. More specific examples include, but are not limited to, tobacco, soybean, potato, rice, chrysanthemum and carnation.

As pathogenic fungi, those containing glucan in cell walls are preferable as targets. Specific examples of such pathogenic fungi include, but are not limited to, the genera *Phytophthora*, *Rhizoctonia*, *Pyricularia*, *Puccinia*, *Fusarium*, *Uromyces*, *Botrytis*

and *Alternaria*. More specifically, the pathogenic fungi include, but are not limited to, *Phytophthora nicotianae*, *Rhizoctonia solani*, *Pyricularia oryzae*, *Puccinia horiana*, *Fusarium oxysporum*, *Fusarium roseum*, *Fusarium tricinctum*, *Uromyces dianthi*, *Botrytis cinerea* and *Alternaria dianthi*.

According to the present invention, the DNA sequence coding for a chimeric receptor transferred into a plant can be inherited to subsequent generations through seeds. Thus, the transferred DNA sequence is also present in those seeds which are formed from the pollens or ovaries of the plant of the invention, and the inherited character can be transmitted to the progeny. Accordingly, the plant of the invention into which a DNA sequence coding for a chimeric receptor has been transferred can be propagated through seeds without losing its resistance to pathogenic fungi. The plant of the invention can also be propagated by a mass propagation method using plant tissue culture or by conventional techniques such as cutting, layering, grafting, division, etc. without losing its resistance to pathogenic fungi.

Whether a transformed plant has resistance to fungi or not can be confirmed from the presence or absence of resistance reaction when a glucan elicitor is added, or examined by fungus inoculation tests. Resistance reaction tests by elicitor addition are performed by adding a glucan elicitor solution to the surface of leaves, or infiltrating a glucan elicitor solution into leaves from the back side of leaves with a syringe and, after a specific period of time, observing browning reaction appearing on leaf surfaces or examining the accumulation of fluorescent substances (phytoalexins) excited by

UV light. In fungus inoculation tests, resistance can be examined by the following inoculation methods, which are roughly divided into two groups, i.e. inoculation to the aerial part and inoculation to the subterranean part.

I. The following methods may be employed for the inoculation of a pathogen to the aerial part.

(1) Spray inoculation: A suspension of a pathogenic fungi or spores thereof is sprayed onto plants, which are then incubated in an inoculation box at an appropriate temperature under saturation humidity for one hour to examine the severity of the disease.

(2) Dusting inoculation: Spores collected from lesions are directed dusted off onto plants with a writing brush, or inoculated onto plants with a small duster. The inoculated plants are incubated a whole day and night in an inoculation box adjusted to a temperature suitable for infection. Subsequently, the same procedures as in the spray inoculation method are taken.

(3) Wound inoculation: A spore suspension or a piece of cultured mycelial tuft is inoculated into cut ends of branches or those portions where the bark is pressed through with a cork borer.

(4) Needle inoculation: A needle or a bundle of needles dipped in a suspension of a pathogenic fungus is used to scratch plants for inoculation.

(5) Cell adhesion inoculation: When a pathogenic fungus which does not form spores on medium is used, cultured mycelial masses or sclerotia are adhered to the surface of plants for inoculation.

(6) Damaged plant inoculation: In stead of direct inoculation of a pathogenic fungus, plants damaged by the fungus are placed on plants to be tested, or foliage damaged by the fungus is suspended over plants to be tested so that inoculation is performed by mycelia growing from the damaged plant or spores falling off from the foliage.

II. The following methods may be employed for the inoculation of a pathogen to the subterranean part.

(1) Soil drenching inoculation: Soil is drenched with a suspension of spores or disrupted mycelia for inoculation.

(2) Soil mixing inoculation: A filamentous fungus cultured in wheat bran, crop grains, straw of rice, straw of barley, etc. is mixed with soil for inoculation.

(3) Diseased plant burying inoculation: Residue from diseased plants is buried in soil to prepare a soil polluted with the pathogenic fungus. When the residue has been rotten, the soil is mixed well. Then, seeds are sown therein or seedlings are transplanted thereto.

(4) Polluted soil inoculation: If a specific soil disease is always occurring in some area, soil is taken therefrom, mixed well and filled in pots or vats. Then, seeds are sown therein or seedlings are transplanted thereto.

(5) Root dipping inoculation: A suspension of spores from a pathogenic bacterium or filamentous fungus is prepared. Roots of healthy seedlings are dipped in this suspension for 3-5 min and then transplanted to pots.

Resistance tests may be performed according to the above-described inoculation methods. More specifically, resistance to

Phytophthora can be assayed by directly inoculating fungal mycelia into plants and observing the expansion of lesions. Alternatively, the resistance may be assayed by inoculating zoospores from *Phytophthora* and observing the vicissitudes of lesions. Resistance to soil fungi other than *Phytophthora* can be assayed by mixing cultured fungal cells with soil, sowing seeds or growing plants on the soil, and observing the phenomenon of damping-off. However, methods for testing resistance to fungi are not limited to these methods.

BRIEF EXPLANATION OF THE DRAWINGS

Fig. 1 shows chimeric DNA constructs of the invention schematically .

Fig. 2 shows an outline of the construction of plasmid pBI(Hm) - ER.

Fig. 3 shows an outline of the construction of plasmid pBI(Hm) - LER.

Fig. 4 shows an outline of a plasmid containing a DNA coding for an N-terminal region of RPS2 protein corresponding to about 1/3 of this protein.

Fig. 5 shows an outline of the construction of plasmid pBI(Hm) - CER.

Fig. 6 shows an outline of the construction of plasmid pBI(Hm) - IER.

Fig. 7 shows an outline of the construction of plasmid pBI(Hm) - ISER.

BEST MODES FOR CARRYING OUT THE INVENTION

Hereinbelow, the present invention will be described more specifically with reference to the following Examples. However, the technical scope of the invention is not limited by these Examples.

In the following Examples, a glucan elicitor receptor is abbreviated to "ER". Abbreviations used in Figs. 2-7 have the following meanings.

Restriction sites:

B: BamHI; Xb: XbaI; SI: SalI; RI: EcoRI; RV: EcoRV; Xh:XhoI;
Sph: SphI; Sp: SpeI; ScI: SacI; Sm: SmaI; Nsp: NspI; N: NotI;
Nsi: NsiI; Nc: NcoI; H3: HindIII

(Sm): SmaI recognition sequence is deleted by the insertion of DNA into SmaI site.

dScI, dB: SacI site and BamHI site are deleted, respectively.

L: Leader sequence of β -1,3-glucanase

pNOS: NOS promoter; p35S: Cauliflower mosaic virus 35S promoter;

Nos-T: NOS terminator

NPTII: Kanamycin resistance gene; Hm: Hygromycin resistance gene

RB, LB: Right border DNA sequence of Ti-plasmid, left border DNA sequence of Ti-plasmid

[EXAMPLE 1] Construction of Vector Plasmids

Individual vectors were constructed by the procedures described below. Unless otherwise specified, reagents such as restriction enzymes, linkers, etc. used in plasmid preparation were products manufactured by Takara Shuzo. For the amplification and selection in *E. coli*, DH5 α (BRL) was used.

(1) Construction of pBI(Hm)-ER (see Fig. 2)

The BamHI and SalI sites located in the multilinker in plasmid pER23-1 [N. Umemoto et al. (1997) Proc. Natl. Acad. Sci. USA 279: 141] were digested with restriction enzymes BamHI and SalI to obtain a fragment containing the full length ER gene sandwiched between the two restriction sites. This fragment was separated from agarose gel. On the other hand, a plant expression type binary plasmid pBI121 (Clontech) was digested with BamHI and SacI; and synthetic linker DNAs (SEQ ID NOS: 15 and 16) shown below were synthesized with an automated DNA synthesizer (Applied Biosystems).

These linker DNAs were annealed and ligated to the digested pBI121 to thereby produce pBI linker. Then, the above-described fragment was cloned between the BamHI and SalI sites of this pBI linker to thereby obtain pBI-ER.

5'-CTAGAGGATCCGGTACCCCCGGGGTCGACGAGCT-3' (SEQ ID NO: 15)

5'-CGTCGACCCCCGGGGGTACCGGATCCT-3' (SEQ ID NO: 16)

Subsequently, pBI-ER was digested with BamHI and SacI to obtain a fragment containing the full length ER gene. This fragment was separated from agarose gel and ligated to the plasmid fragment obtained by digesting a plant expression type binary plasmid pIG121-Hm [Nakamura et al. (1991), Plant Biotechnology II (Modern Chemistry Suppl. Vol. 20), pp.123-132; released from Prof. Kenzo Nakamura of Nagoya University, Japan] with BamHI and SacI, to thereby obtain pBI(Hm)-ER, the vector of interest.

As a result of DNA sequencing of the resultant vector, the

nucleotide sequence shown in SEQ ID NO: 2 was obtained. The amino acid sequence encoded by the nucleotide sequence of SEQ ID NO: 2 is shown in SEQ ID NO: 1.

E. coli DH5 α EKB633 carrying plasmid pER23-1 transferred thereinto was deposited at the National Institute of Bioscience and Human-technology, Agency of Industrial Science and Technology (1-3, Higashi 1-chome, Tsukuba City, Ibaraki Pref., Japan) on June 15, 1994 under the Accession No. FERM BP-4699.

(2) Construction of pBI(Hm)-LER (see Fig. 3)

The two HindIII sites in pER23-1 were cleaved with HindIII. The resultant plasmid was subjected to self-ligation to thereby delete the portion of ER after the HindIII site located in ER. Subsequently, the remaining two NotI sites were cleaved with NotI and filled in with Klenow fragment. To the resultant plasmid, BamHI linker was introduced to thereby create a BamHI site. This plasmid was digested with BamHI and HindIII to obtain a fragment containing the portion of ER prior to the HindIII site. To this fragment, a plasmid fragment containing the portion of ER after the HindIII site obtained by digesting pER23-1 with BamHI and HindIII was ligated.

The resultant plasmid was further digested with XbaI and BamHI. Then, synthetic linker DNAs (SEQ ID NOS: 17 and 18) derived from the untranslated 5'end of soybean β -1,3-glucanase and synthesized with an automated DNA synthesizer were annealed and ligated to the digested plasmid, thereby preparing pLER.

5'-CTAGACTTCTTTCCTCAACCTTCTTCTTCTTATATATTCGAAC-3' (SEQ ID NO: 17)

5'-GATCGTTCGAATATATAAGAAGAAAGAAGGTTGAGGAAAGAAGT-3' (SEQ ID NO: 18)

Subsequently, pLER was digested with SalI and filled in with Klenow fragment. Then, SacI linker was introduced to create a SacI site. The resultant plasmid was further digested with XbaI and SacI to obtain a fragment containing the ER gene with the leader sequence (L). This fragment was separated from agarose gel and then ligated to a plasmid fragment obtained by digesting a plant expression type binary vector plasmid pIG121-Hm with XbaI and SacI, thereby obtaining pBI(Hm)-LER, the vector of interest.

As a result of DNA sequencing of the resultant vector, the nucleotide sequence shown in SEQ ID NO: 4 was obtained. The amino acid sequence encoded by the nucleotide sequence of SEQ ID NO: 4 is shown in SEQ ID NO: 3.

(3) Acquisition of cDNA Clones Coding for an N-Terminal Region of *Arabidopsis thaliana*-Derived RPS2 Corresponding to about 1/3 of This Protein

Arabidopsis thaliana cDNA was obtained by plating a cDNA library (Clontech) on ten 15 cm LB plates at 200,000 plaques/plate after infecting *E. coli* C600hfl, and recovering the phage from the plates. 5 pools of phage library were obtained, the phage recovered from 2 plates being taken as 1 pool. Phage DNA was recovered from each pool according to the method of Ishida ("Gene Expression Experiment Manual", Ishida & Ando (eds.), Kodansha Scientific Co.)

The cDNA of interest encoding the N-terminal region of pRPS2 was amplified by PCR (using EX-Taq kit from Takara Shuzo) using either of the following PCR primer sets and about 1 μ g of phage DNA (pool #1-#5) as a template, and then cloned into

pBluescript SKII(+) plasmid.

(i) Sense primer: 5'-TGGCATGCGATGGATTTTCATCTCATCTCTT-3'

(SEQ ID NO:19)

Antisense primer: 5'-GGGAATTCACTCCGCGAGCCGGCGAAT-3'

(SEQ ID NO:20)

(ii) Sense primer: 5'-ACAAGTAAAAGAAAGAGCGAGAAATCATCGAAATG-3'

(SEQ ID NO:21)

Antisense primer: 5'-AGCCATGGCTCCTCCTAAAGTGAT-3'

(SEQ ID NO:22)

PCR was performed using the above primer set (i) or (ii).

The amplified DNA was phosphorylated and inserted into the SmaI site of pBluescript SKII(+), thereby obtaining the plasmid DNAs shown in Fig. 4 (pRPS2-6, pRPS2-14 and pRPS2-16). It was confirmed by DNA sequencing that these DNAs were identical with the reported sequence [M. Mindrinos et al. (1994) Cell 78: 1089] except their primer portions.

(4) Construction of pBI(Hm)-CER (see Fig. 5)

pRPS2-14 was digested with restriction enzyme XbaI, followed by self-ligation to thereby delete the XbaI fragment. BamHI linker was inserted into the EcoRV site of the resultant plasmid. Further, SacI linker was inserted into the XhoI site. The resultant plasmid was digested with BamHI and SalI, and the full-length ER cut out from pER23-1 with BamHI and salI was cloned thereinto. From this plasmid, a chimeric DNA containing a part of RPS2 and the full-length ER was purified. By DNA sequencing of the junction between RPS2 and ER, it was confirmed that the reading frame of RPS2 coincided with the reading frame

of ER.

On the other hand, pSpeI linker was inserted into the RcoRV site of pRPS2-6. Subsequently, this plasmid was digested with XbaI and SacI, and then the above-described purified DNA fragment was cloned thereinto. This plasmid containing the chimeric gene composed of a 5' end region of RPS2 corresponding to about 1/3 of RPS2 and the full-length ER was digested with SpeI and SacI to cut out the chimeric gene, which was then cloned into pIG121-Hm plasmid. Thus, pBI(Hm)-CER, the vector of interest, was obtained.

As a result of DNA sequencing of the resultant vector, the nucleotide sequence shown in SEQ ID NO: 6 was obtained. The amino acid sequence encoded by the nucleotide sequence of SEQ ID NO: 6 is shown in SEQ ID NO: 5.

(5) Construction of pBI(Hm)-IER (see Fig. 6)

An ER cDNA fragment coding for the N-terminal 537 amino acid was purified from pER23-1 using BamHI and NspI. On the other hand, a RPS2 cDNA fragment was purified from pRPS2-16 plasmid using EcoRV and SphI. These two DNA fragments were cloned between the BamHI and SmaI sites of pBluescript SKII(+) plasmid. This plasmid was designated pERRPS-8. By DNA sequencing of the junction between RPS2 and ER, it was confirmed that the reading frame of RPS2 coincided with the reading frame of ER. pERRPS-8 was digested with SpeI and SacI to obtain a fragment containing a fusion gene composed of a domain encoding an elicitor binding region and a domain encoding a partial-length RPS2. This fragment was separated from agarose gel and ligated to a plasmid fragment obtained by digesting a plant

expression type binary vector pIG121-Hm with XbaI and SacI, to thereby obtain pBI(Hm)-IER, the vector of interest.

As a result of DNA sequencing of the resultant vector, the nucleotide sequence shown in SEQ ID NO: 8 was obtained. The amino acid sequence encoded by the nucleotide sequence of SEQ ID NO: 8 is shown in SEQ ID NO: 7.

(6) Construction of pBI(Hm)-ISER (see Fig. 7)

pERRPS-8 plasmid was digested with NsiI and NotI. Then, the NotI-NsiI adaptors shown below were inserted into the plasmid. It was confirmed by DNA sequencing that the reading frame of the ATG within the adaptor coincided with the reading frame of ER. The resultant plasmid was digested with EcoRV and SalI, and then the resultant RPS2/ER chimeric DNA was purified.

5'-GGCCGCGATATCATGAATGCA-3' (SEQ ID NO: 23)

5'-TTCATGATATCGC-3' (SEQ ID NO: 24)

The synthetic DNAs shown below derived from the untranslated 5' end of β -1,3-glucanase were prepared and inserted into the NotI and EcoRI sites of pBluescript SKII(+), thereby preparing plasmid pL.

5'-GGCCGCCTTCTTTCCTCAACCTTCTTTCTTCTTATATATTCGAAC-3' (SEQ ID NO: 25)

5'-AATTGTTCTGAATATATAAGAAGAAAGAAGGTTGAGGAAAGAAGGC-3' (SEQ ID NO: 26)

Plasmid pL was digested with EcoRV and SalI, and the RPS2/ER chimeric DNA purified above was inserted between the restriction sites. The resultant plasmid is designated pSERRPS-8. pSERRPS-8 was digested with SpeI and SacI to obtain a fragment containing a fusion gene composed of a domain encoding an

elicitor binding site and a domain encoding a partial-length RPS2. This fragment was separated from agarose gel and ligated to a plasmid fragment obtained by digesting a plant expression type binary vector pIG121-Hm with XbaI and SacI, to thereby obtain pBI(Hm)-ISER, the vector of interest.

As a result of DNA sequencing of the resultant vector, the nucleotide sequence shown in SEQ ID NO: 10 was obtained. The amino acid sequence encoded by the nucleotide sequence of SEQ ID NO: 10 is shown in SEQ ID NO: 9.

[EXAMPLE 2] Transfer of the Chimeric Receptor Gene into
Arabidopsis thaliana Plants

(1) Preparation of *Agrobacterium* Infection Liquid

Agrobacterium tumefaciens EHA101 into which a transformation vector [pBI(Hm)-CER or pBI(Hm)-ISER] containing the chimeric receptor gene or a transformation vector [pBI(Hm)-ER] containing ER gene as a control had been transferred by electroporation was inoculated into 3 ml of YEB medium and cultured at 28°C for 16 hr. Then, the cells were harvested by centrifugation and suspended in 10 ml of an infection medium to prepare an *Agrobacterium* infection liquid for transformation. The compositions of YEB medium and the infection medium are as described below.

YEB medium: 5 g/L beef extract; 1 g/L yeast extract; 5 g/L peptone; 5 g/L sucrose; 2 mM magnesium sulfate

Infection medium: the inorganic salts & vitamins used in MS [Murashige & Skoog (1962) *Physiol. Plant.* 15: 473] medium at 0.5-fold concentrations; 15 g/L sucrose; 10 g/L glucose; 10 mM MES (pH 5.4)

(2) Method for Creating *Arabidopsis thaliana* Transformants

Seeds of *Arabidopsis thaliana* L. ecotypes Landsberg and Columbia were sown in MSHF medium aseptically and cultured for 3 weeks. Roots grown up to 2-3 cm were cut to approximately 5 mm, transferred onto a callus-inducing medium and cultured for 2 days. Roots cut into pieces were dipped in the *Agrobacterium* infection liquid. After excessive liquid on the root pieces was removed on a paper filter, the root pieces were transferred onto *Arabidopsis* co-culture medium and cultured in a dark place for 2 days. Then, the root pieces were transferred onto *Arabidopsis* redifferentiation medium and cultured for 2 days. Then, the root pieces were transferred onto *Arabidopsis* redifferentiation medium for selection and cultured for 1 week. Thereafter, the root pieces were transferred onto fresh *Arabidopsis* redifferentiation medium for selection every 1 week and cultured until regenerated seedlings were obtained.

Seedlings whose shoot had grown up to approximately 1 cm were transplanted to *Arabidopsis* root-inducing medium and cultured further. The presence of a transgene in seedlings whose roots had elongated to 1-2 cm was confirmed by PCR to obtain transformants. Unless otherwise specified, the cultivation was performed at 22°C under a 16 hr illumination/8 hr non-illumination regime. The components of each medium used in the cultivation are as described below.

MSHF medium: the inorganic salts & vitamins used in MS medium [Murashige & Skoog (1962) *Physiol. Plant.* 15: 473]; 20 g/L sucrose; 2 g/L gellan gum; 5 mM MES (pH 6.2)

Callus-inducing medium: the inorganic salts & vitamins used

in MS medium; 20 g/L sucrose; 0.5 mg/L 2,4-dichlorophenoxyacetic acid; 0.1 mg/L kinetin; 2 g/L gellan gum; 5 mM MES (pH 6.2)

Arabidopsis co-culture medium: the inorganic salts & vitamins used in MS medium; 20 g/L sucrose; 0.5 mg/L 2,4-dichlorophenoxyacetic acid; 0.1 mg/L kinetin; 0.2 mM acetosyringone; 2 g/L gellan gum; 5 mM MES (pH 6.2)

Arabidopsis redifferentiation medium: the inorganic salts & vitamins used in MS medium; 20 g/L sucrose; 1 mg/L 2-isopentenylamine; 0.15 mg/L indoleacetic acid; 250 mg/L Cefotaxime; 2 g/L gellan gum; 5 mM MES (pH 6.2)

Arabidopsis redifferentiation medium for selection: the inorganic salts & vitamins used in MS medium; 20 g/L sucrose; 1 mg/L 2-isopentenylamine; 0.15 mg/L indoleacetic acid; 2 g/L gellan gum; 50 mg/L kanamycin; 250 mg/L Cefotaxime; 5 mM MES (pH 6.2)

Arabidopsis root-inducing medium: the inorganic salts & vitamins used in MS medium; 20 g/L sucrose; 2 mg/L naphthaleneacetic acid; 2 g/L gellan gum; 250 mg/L Cefotaxime; 5 mM MES (pH 6.2)

[EXAMPLE 3] Transfer of the Chimeric Receptor Gene into Tobacco Plants

(1) Preparation of *Agrobacterium* Infection Liquid

Agrobacterium tumefaciens EHA101 into which a transformation vector [pBI(Hm)-IER] containing the chimeric receptor gene or a transformation vector [pBI(Hm)-ER] containing ER gene as a control had been transferred by electroporation was inoculated into 3 ml of YEB medium and cultured at 28°C for 16 hr. Then, the cells were harvested by centrifugation and suspended

in 10 ml of an infection medium to prepare an *Agrobacterium* infection liquid for transformation. The compositions of YEB medium and the infection medium are as described below.

YEB medium: 5 g/L beef extract; 1 g/L yeast extract; 5 g/L peptone; 5 g/L sucrose; 2 mM magnesium sulfate

Infection medium: the inorganic salts & vitamins used in MS [Murashige & Skoog (1962) *Physiol. Plant.* 15: 473] medium at 0.5-fold concentrations; 15 g/L sucrose; 10 g/L glucose; 10 mM MES (pH 5.4)

(2) Method for Creating Tobacco Transformants

Leaves of tobacco (*Nicotiana tabacum* L.) cv. Samsun were sterilized and cut into pieces about 1 cm square. These leaf pieces were floated on the *Agrobacterium* infection liquid for 10 min and then excessive liquid on them was removed on a paper filter. Subsequently, the leaf pieces were transferred onto tobacco co-culture medium and cultured in the dark for 2 days. After the leaf pieces were transferred onto tobacco redifferentiation medium and cultured for 1 week, they were transferred onto tobacco redifferentiation medium for selection and cultured for 2 weeks. Thereafter, they were transferred onto fresh tobacco redifferentiation medium for selection every 2 weeks and cultured until regenerated seedlings were obtained.

Seedlings whose shoot had grown up to approximately 1 cm were transplanted to tobacco root-inducing medium and cultured further. The presence of a transgene in seedlings whose roots had elongated to 1-2 cm was confirmed by PCR to obtain transformants. Unless otherwise specified, the cultivation was performed at 25°C under a 16 hr illumination/8 hr non-

illumination regime. The components of each medium used in the cultivation are as described below.

Tobacco co-culture medium: the inorganic salts & vitamins used in MS medium [Murashige & Skoog (1962) *Physiol. Plant.* 15: 473]; 30 g/L sucrose; 0.1 mg/L naphthaleneacetic acid; 1 mg/L benzyladenine; 0.2 mM acetosyringone; 8 g/L agar; 5 mM MES (pH 5.8)

Tobacco redifferentiation medium: the inorganic salts & vitamins used in MS medium; 30 g/L sucrose; 0.1 mg/L naphthaleneacetic acid; 1 mg/L benzyladenine; 250 mg/L Cefotaxime; 8 g/L agar; 5 mM MES (pH 5.8)

Tobacco redifferentiation medium for selection: the inorganic salts & vitamins used in MS medium; 30 g/L sucrose; 0.1 mg/L naphthaleneacetic acid; 1 mg/L benzyladenine; 100 mg/L kanamycin; 250 mg/L Cefotaxime; 8 g/L agar; 5 mM MES (pH 5.8)

Tobacco root-inducing medium: the inorganic salts & vitamins used in MS medium; 30 g/L sucrose; 250 mg/L Cefotaxime; 8 g/L agar; 5 mM MES (pH 5.8)

[EXAMPLE 4] Hypersensitive Response Test on Transformed *Arabidopsis thaliana*

(1) Hypersensitive Response Test on Transformed *Arabidopsis thaliana* (Ecotype Landsberg)

Induction of the hypersensitive response by a soybean elicitor was examined using leaves of *Arabidopsis thaliana* transformant (CER) in which high expression of one of the created chimeric receptor genes, i.e. the gene comprising the nucleotide sequence shown in SEQ ID NO: 6 was confirmed, as well as *Arabidopsis thaliana* transformant (ER; SEQ ID NO: 2) in which

high expression of ER protein was confirmed and non-transformed *Arabidopsis thaliana* (control).

A fungal cell wall-derived glucan elicitor (100 μ g/ml) dissolved in 10 mM MgCl₂ solution and 10 mM MgCl₂ solution alone were infiltrated separately into leaves of *Arabidopsis thaliana* plants grown up in a greenhouse, from the back side of the leaves using a syringe not equipped with a needle. The plants were cultured at 22 °C for 7 days under a 16 hr light/8 hr dark regime to observe changes appearing on the leaf surface. The fungal cell wall-derived glucan elicitor used in this experiment was obtained according to the method of Umemoto et al. [N. Umemoto et al. (1997) Proc. Natl. Acad. Sci. USA 94: 1029].

When the glucan elicitor was infiltrated into non-transformed *Arabidopsis thaliana*, no changes were observed on the leaf surface. This indicates that *Arabidopsis thaliana* cannot recognize the glucan elicitor. On the other hand, in the *Arabidopsis thaliana* plants (ER) expressing a glucan elicitor receptor, 2 clones out of the 6 tested exhibited weak browning reactions (resistance reaction). In the *Arabidopsis thaliana* plants (CER) expressing a chimeric receptor, 8 clones out of the 12 tested exhibited effective browning reactions (resistance reaction) and 1 clone exhibited weak browning reactions (Table 1).

Table 1. Hypersensitive Response Test on Transformed *Arabidopsis thaliana* (Ecotype Landsberg)

	Fungal cell wall-derived glucan elicitor	10 mM MgCl ₂
Control 1	—	—
Control 2	—	—
Control 3	—	—
Control 4	—	—
Control 5	—	—
Control 6	—	—
Control 7	—	—
Control 8	—	—
Control 9	—	—
Control 10	—	—
Control 11	—	—
Control 12	—	—
ER 1	—	—
ER 2	—	—
ER 3	+	—
ER 4	—	—
ER 5	—	—
ER 6	+	—
CER 1	++	—
CER 2	++	—
CER 3	++	—
CER 4	++	—
CER 5	—	—
CER 6	+++	—
CER 7	+++	—
CER 8	++	—
CER 9	—	—
CER 10	—	—
CER 11	++	—
CER 12	—	—

—: no change; +: slight browning reactions;

++: medium browning reactions; +++: strong browning reactions

From these results, it has been proved that recognition of a glucan elicitor which a host plant cannot recognize takes place not only when a soybean-derived ER protein is expressed alone but also when a chimeric receptor composed of the ER protein and an N-terminal domain of *Arabidopsis thaliana*-derived

RPS2 protein which is expected to contain a signal transduction domain is expressed in *Arabidopsis thaliana*. Furthermore, it has been proved that expressing a chimeric receptor is more advantageous for recognition of a glucan elicitor. This demonstrates that when a chimeric receptor is expressed, the signal transduction after the binding of a glucan elicitor to the receptor is more efficiently performed, as compared to the case where an ER protein alone is expressed.

The present invention provide a DNA and plants which can induce the resistance reaction deeply involved in disease resistance more strongly than an ER protein expressed alone, by recognizing the contact and/or invasion of plant pathogenic microorganisms such as fungi having glucan structures in their cell walls, etc. By creating such plants which strongly induce the resistance reaction caused by a glucan elicitor, it will become possible to breed plants which exhibit strong resistance to plant pathogenic microorganisms such as fungi.

(2) Hypersensitive Response Test on Transformed *Arabidopsis thaliana* Ecotype Columbia

Induction of the hypersensitive response by a soybean elicitor was examined using leaves of *Arabidopsis thaliana* transformant (ISER) in which high expression of one of the created chimeric receptor genes, i.e. the gene comprising the nucleotide sequence shown in SEQ ID NO: 10 was confirmed, as well as non-transformed *Arabidopsis thaliana* (control).

A fungal cell wall-derived glucan elicitor (100 μ g/ml) dissolved in 10 mM MgCl₂ solution and 10 mM MgCl₂ solution alone were infiltrated separately into leaves of *Arabidopsis thaliana*

plants grown up in a greenhouse from the back side of the leaves using a syringe not equipped with a needle. The plants were cultured at 22°C for 7 days under a 16 hr light/8 hr dark regime to observe changes appearing on the leaf surface. The fungal cell wall-derived glucan elicitor used in this experiment was obtained according to the method of Umemoto et al. [N. Umemoto et al. (1997) Proc. Natl. Acad. Sci. USA 94: 1029].

When the glucan elicitor was infiltrated into non-transformed *Arabidopsis thaliana*, no changes were observed on the leaf surface. This indicates that *Arabidopsis thaliana* cannot recognize the glucan elicitor. On the other hand, in the *Arabidopsis thaliana* plants (ISER) expressing a chimeric receptor, 3 clones out of the 5 tested exhibited effective browning reactions (resistance reaction) (Table 2).

Table 2. Hypersensitive Response Test on Transformed *Arabidopsis thaliana* (Ecotype Columbia)

	Fungal cell wall-derived glucan elicitor	10 mM MgCl ₂
Control 1	—	—
Control 2	—	—
Control 3	—	—
ISER 1	+++	—
ISER 2	—	—
ISER 3	++	—
ISER 4	+++	—
ISER 5	—	—

—: no change; +: slight browning reactions;

++: medium browning reactions; +++: strong browning reactions

From these results, it has been proved that recognition of a glucan elicitor which a host plant cannot recognize takes place by expressing in *Arabidopsis thaliana* a chimeric receptor

composed of an ER from which an N-terminal region corresponding to about 1/3 of this protein has been deleted and an N-terminal domain of *Arabidopsis thaliana*-derived RPS2 protein, the domain being expected to contain a signal transduction domain. The present invention provides plants which can recognize the contact and/or invasion of plant pathogenic microorganisms such as fungi having glucan structures in their cell walls, etc. By creating such plants which strongly induce the resistance reaction caused by a glucan elicitor, it will become possible to breed plants which exhibit strong resistance to plant pathogenic microorganisms such as fungi.

[EXAMPLE 5] Hypersensitive Response Test on Transformed Tobacco

Induction of the hypersensitive response by a soybean elicitor was examined using leaves of tobacco transformant (IER) in which high expression of one of the created chimeric receptor genes, i.e. the gene comprising the nucleotide sequence shown in SEQ ID NO: 8 was confirmed, as well as tobacco transformant (LER; SEQ ID NO: 4) in which high expression of ER protein was confirmed and non-transformed tobacco (control).

Leaves of tobacco plants grown up in a greenhouse were cut off at the petiole and placed in light-transmittable plastic boxes. Silicone tube pieces 5 mm in diameter and 5 mm in height were put on the upper surface of leaves so that a liquid could be retained. A fungal cell wall-derived glucan elicitor (100 μ g/ml) dissolved in 10 mM MgCl₂ solution, a chemically synthesized glucan elicitor (100 μ g/ml) dissolved in the same solution, and 10 mM MgCl₂ solution alone as a control were retained separately

in the silicone tube pieces so that each solution was in contact with the leaf surface. The leaves were cultured under excessive moisture to prevent drying, under a 16 hr light/8 hr dark regime at 25 °C for 7 days. The chemically synthesized glucan elicitor (β -D-glucohexaoside) [N. Hong and T. Ogawa (1990) Tetrahedron Lett. 31: 3179] was released from Prof. Ogawa (Institute of Physical and Chemical Research, and the University of Tokyo). After the cultivation, the silicone tube pieces were removed. Then, the accumulation of tobacco phytoalexins was observed on a UV illuminator (Funakoshi).

When the glucan elicitor was added to non-transformed tobacco, no changes were observed. This indicates that the non-transformed tobacco cannot recognize the glucan elicitor. In the ER-expressing tobacco plants, 1 clone out of the 5 tested only exhibited accumulation of very small amounts of fluorescent substances (phytoalexins). In the IER-expressing tobacco, 3 clones out of the 5 tested exhibited accumulation of remarkable amounts of fluorescent substances (phytoalexins) and 1 clone exhibited accumulation of very small amounts of fluorescent substances (phytoalexins) (Table 3).

Table 3. Hypersensitive Response Test on Transformed Tobacco

	Fungal cell wall-derived glucan elicitor	Chemically synthesized glucan elicitor	10mM MgCl ₂
Control 1	—	—	—
Control 2	—	—	—
LER 1	—	—	—
LER 2	—	—	—
LER 3	+	+	—
LER 4	—	—	—
LER 5	—	—	—
IER 1	+++	++	—
IER 2	+++	+++	—
IER 3	++	+	—
IER 4	—	—	—
IER 5	+	+	—

— : no change; +: slight accumulation of fluorescent substances;

++: medium accumulation of fluorescent substances;

+++ : large accumulation of fluorescent substances

From these results, it has been proved that not only when a soybean-derived ER protein is expressed alone but also when a chimeric receptor composed of the ER protein and an N-terminal domain of *Arabidopsis thaliana*-derived RPS2 protein is expressed in a host plant other than soybean and *Arabidopsis thaliana* (tobacco), recognition of a glucan elicitor which the host plant cannot recognize takes place. Also, it has been shown that expressing a chimeric receptor is more advantageous for recognition of a glucan elicitor. Thus, it has been shown that even in a plant other than *Arabidopsis thaliana*, signal transduction is performed more efficiently when a gene coding for a chimeric receptor containing an N-terminal domain of RPS2 is expressed therein, as compared to the case where a gene coding for an ER protein alone is expressed. The present invention provides a DNA coding for a chimeric receptor which

can recognize the contact and/or invasion of plant pathogenic microorganisms such as fungi having glucan structures in their cell walls, and plants which have such a chimeric receptor. As a result, resistance reactions deeply involved in disease resistance, such as phytoalexin induction, are induced more strongly than those cases where an ER protein alone is expressed. By creating such plants which strongly induce the resistance reactions caused by glucan elicitors, it becomes possible to breed plants having strong resistance to plant pathogenic microorganisms such as fungi.

INDUSTRIAL APPLICABILITY

According to the present invention, a DNA coding for a chimeric protein (chimeric receptor) comprising a domain containing at least an elicitor binding site of an elicitor receptor and a domain containing a signal transduction motif of a plant-derived resistance gene; disease resistant plants into which the above DNA is transferred; and a method for creating such plants are provided.

The chimeric receptor of the invention can be used to elucidate the mechanism of resistance to fungi.

Further, the DNA of the invention comprising a nucleotide sequence coding for a chimeric receptor and fragments thereof are useful as materials for establishing techniques to breed fungus-resistant plants. That is, if the DNA of the invention or a fragment thereof is transferred into various plants and expressed therein, the plants' resistance to fungi can be enhanced.

SEQUENCE LISTING

SEQ ID NO: 1

SEQUENCE LENGTH: 667

SEQUENCE TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION:

Val Asn Ile Gln Thr Asn Thr Ser Tyr Ile Phe Pro Gln Thr Gln Ser
 1 5 10 15
 Thr Val Leu Pro Asp Pro Ser Lys Phe Phe Ser Ser Asn Leu Leu Ser
 20 25 30
 Ser Pro Leu Pro Thr Asn Ser Phe Phe Gln Asn Phe Val Leu Lys Asn
 35 40 45
 Gly Asp Gln Gln Glu Tyr Ile His Pro Tyr Leu Ile Lys Ser Ser Asn
 50 55 60
 Ser Ser Leu Ser Leu Ser Tyr Pro Ser Arg Gln Ala Ser Ser Ala Val
 65 70 75 80
 Ile Phe Gln Val Phe Asn Pro Asp Leu Thr Ile Ser Ala Pro Gln Gly
 85 90 95
 Pro Lys Gln Gly Pro Pro Gly Lys His Leu Ile Ser Ser Tyr Ser Asp
 100 105 110
 Leu Ser Val Thr Leu Asp Phe Pro Ser Ser Asn Leu Ser Phe Phe Leu
 115 120 125
 Val Arg Gly Ser Pro Tyr Leu Thr Val Ser Val Thr Gln Pro Thr Pro
 130 135 140
 Leu Ser Ile Thr Thr Ile His Ser Ile Leu Ser Phe Ser Ser Asn Asp
 145 150 155 160
 Ser Asn Thr Lys Tyr Thr Phe Gln Phe Asn Asn Gly Gln Thr Trp Leu
 165 170 175

Leu Tyr Ala Thr Ser Pro Ile Lys Leu Asn His Thr Leu Ser Glu Ile
 180 185 190
 Thr Ser Asn Ala Phe Ser Gly Ile Ile Arg Ile Ala Leu Leu Pro Asp
 195 200 205
 Ser Asp Ser Lys His Glu Ala Val Leu Asp Lys Tyr Ser Ser Cys Tyr
 210 215 220
 Pro Val Ser Gly Lys Ala Val Phe Arg Glu Pro Phe Cys Val Glu Tyr
 225 230 235 240
 Asn Trp Glu Lys Lys Asp Ser Gly Asp Leu Leu Leu Leu Ala His Pro
 245 250 255
 Leu His Val Gln Leu Leu Arg Asn Gly Asp Asn Asp Val Lys Ile Leu
 260 265 270
 Glu Asp Leu Lys Tyr Lys Ser Ile Asp Gly Asp Leu Val Gly Val Val
 275 280 285
 Gly Asp Ser Trp Val Leu Lys Thr Asp Pro Leu Phe Val Thr Trp His
 290 295 300
 Ser Ile Lys Gly Ile Lys Glu Glu Ser His Asp Glu Ile Val Ser Ala
 305 310 315 320
 Leu Ser Lys Asp Val Glu Ser Leu Asp Ser Ser Ser Ile Thr Thr Thr
 325 330 335
 Glu Ser Tyr Phe Tyr Gly Lys Leu Ile Ala Arg Ala Ala Arg Leu Val
 340 345 350
 Leu Ile Ala Glu Glu Leu Asn Tyr Pro Asp Val Ile Pro Lys Val Arg
 355 360 365
 Asn Phe Leu Lys Glu Thr Ile Glu Pro Trp Leu Glu Gly Thr Phe Ser
 370 375 380
 Gly Asn Gly Phe Leu His Asp Glu Lys Trp Gly Gly Ile Ile Thr Gln
 385 390 395 400
 Lys Gly Ser Thr Asp Ala Gly Gly Asp Phe Gly Phe Gly Ile Tyr Asn

	405		410		415										
Asp	His	His	Tyr	His	Leu	Gly	Tyr	Phe	Ile	Tyr	Gly	Ile	Ala	Val	Leu
	420		425		430										
Thr	Lys	Leu	Asp	Pro	Ala	Trp	Gly	Arg	Lys	Tyr	Lys	Pro	Gln	Ala	Tyr
	435		440		445										
Ser	Ile	Val	Gln	Asp	Phe	Leu	Asn	Leu	Asp	Thr	Lys	Leu	Asn	Ser	Asn
	450		455		460										
Tyr	Thr	Arg	Leu	Arg	Cys	Phe	Asp	Pro	Tyr	Val	Leu	His	Ser	Trp	Ala
465			470		475				480						
Gly	Gly	Leu	Thr	Glu	Phe	Thr	Asp	Gly	Arg	Asn	Gln	Glu	Ser	Thr	Ser
	485		490		495										
Glu	Ala	Val	Ser	Ala	Tyr	Tyr	Ser	Ala	Ala	Leu	Met	Gly	Leu	Ala	Tyr
	500		505		510										
Gly	Asp	Ala	Pro	Leu	Val	Ala	Leu	Gly	Ser	Thr	Leu	Thr	Ala	Leu	Glu
	515		520		525										
Ile	Glu	Gly	Thr	Lys	Met	Trp	Trp	His	Val	Lys	Glu	Gly	Gly	Thr	Leu
	530		535		540										
Tyr	Glu	Lys	Glu	Phe	Thr	Gln	Glu	Asn	Arg	Val	Met	Gly	Val	Leu	Trp
545			550		555				560						
Ser	Asn	Lys	Arg	Asp	Thr	Gly	Leu	Trp	Phe	Ala	Pro	Ala	Glu	Trp	Lys
	565		570		575										
Glu	Cys	Arg	Leu	Gly	Ile	Gln	Leu	Leu	Pro	Leu	Ala	Pro	Ile	Ser	Glu
	580		585		590										
Ala	Ile	Phe	Ser	Asn	Val	Asp	Phe	Val	Lys	Glu	Leu	Val	Glu	Trp	Thr
	595		600		605										
Leu	Pro	Ala	Leu	Asp	Arg	Glu	Gly	Gly	Val	Gly	Glu	Gly	Trp	Lys	Gly
	610		615		620										
Phe	Val	Tyr	Ala	Leu	Glu	Gly	Val	Tyr	Asp	Asn	Glu	Ser	Ala	Leu	Gln
625			630		635				640						

Lys Ile Arg Asn Leu Lys Gly Phe Asp Gly Gly Asn Ser Leu Thr Asn

645

650

655

Leu Leu Trp Trp Ile His Ser Arg Ser Asp Glu

660

665

SEQ ID NO: 2

SEQUENCE LENGTH: 2240

SEQUENCE TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

FEATURE

NAME/KEY: CDS

LOCATION: 29..2029

SEQUENCE DESCRIPTION:

TCGAATTAAA CCACCTTCAG CAACAATG GTT AAC ATC CAA ACC AAT ACA TCT 52

Val Asn Ile Gln Thr Asn Thr Ser

1

5

TAC ATC TTC CCT CAA ACA CAA TCC ACT GTT CTT CCT GAT CCC TCC AAA 100

Tyr Ile Phe Pro Gln Thr Gln Ser Thr Val Leu Pro Asp Pro Ser Lys

10

15

20

TTC TTC TCC TCA AAC CTT CTC TCA AGT CCA CTC CCC ACA AAC TCT TTC 148

Phe Phe Ser Ser Asn Leu Leu Ser Ser Pro Leu Pro Thr Asn Ser Phe

25

30

35

40

TTC CAA AAC TTT GTC CTA AAA AAT GGT GAC CAA CAA GAA TAC ATT CAT 196

Phe Gln Asn Phe Val Leu Lys Asn Gly Asp Gln Gln Glu Tyr Ile His

45

50

55

CCT TAC CTC ATC AAA TCC TCC AAC TCT TCC CTC TCT CTC TCA TAC CCT 244

Pro Tyr Leu Ile Lys Ser Ser Asn Ser Ser Leu Ser Leu Ser Tyr Pro

60	65	70	
TCT CGC CAA GCC AGT TCA GCT GTC ATA TTC CAA GTC TTC AAT CCT GAT	292		
Ser Arg Gln Ala Ser Ser Ala Val Ile Phe Gln Val Phe Asn Pro Asp			
75	80	85	
CTT ACC ATT TCA GCC CCA CAA GGT CCC AAA CAA GGT CCC CCT GGT AAA	340		
Leu Thr Ile Ser Ala Pro Gln Gly Pro Lys Gln Gly Pro Pro Gly Lys			
90	95	100	
CAC CTT ATC TCC TCC TAC AGT GAT CTC AGT GTC ACC TTG GAT TTC CCT	388		
His Leu Ile Ser Ser Tyr Ser Asp Leu Ser Val Thr Leu Asp Phe Pro			
105	110	115	120
TCT TCC AAT CTG AGC TTC TTC CTT GTT AGG GGA AGC CCC TAT TTG ACT	436		
Ser Ser Asn Leu Ser Phe Phe Leu Val Arg Gly Ser Pro Tyr Leu Thr			
125	130	135	
GTG TCT GTG ACT CAA CCA ACT CCT CTT TCA ATT ACC ACC ATC CAT TCC	484		
Val Ser Val Thr Gln Pro Thr Pro Leu Ser Ile Thr Thr Ile His Ser			
140	145	150	
ATT CTC TCA TTC TCT TCA AAT GAC TCC AAC ACC AAG TAC ACC TTT CAG	532		
Ile Leu Ser Phe Ser Ser Asn Asp Ser Asn Thr Lys Tyr Thr Phe Gln			
155	160	165	
TTC AAC AAT GGT CAA ACA TGG CTT CTT TAT GCT ACC TCC CCC ATC AAG	580		
Phe Asn Asn Gly Gln Thr Trp Leu Leu Tyr Ala Thr Ser Pro Ile Lys			
170	175	180	
TTG AAC CAC ACC CTT TCT GAG ATA ACT TCT AAT GCA TTT TCT GGC ATA	628		
Leu Asn His Thr Leu Ser Glu Ile Thr Ser Asn Ala Phe Ser Gly Ile			
185	190	195	200
ATC CGG ATA GCT TTG TTG CCG GAT TCG GAT TCG AAA CAC GAG GCT GTT	676		
Ile Arg Ile Ala Leu Leu Pro Asp Ser Asp Ser Lys His Glu Ala Val			
205	210	215	
CTT GAC AAG TAT AGT TCT TGT TAC CCC GTG TCA GGT AAA GCT GTG TTC	724		

CCA TGG TTG GAG GGA ACT TTT AGT GGG AAT GGA TTC CTA CAT GAT GAA	1204
Pro Trp Leu Glu Gly Thr Phe Ser Gly Asn Gly Phe Leu His Asp Glu	
380 385 390	
AAA TGG GGT GGC ATT ATT ACC CAA AAG GGG TCC ACT GAT GCT GGT GGT	1252
Lys Trp Gly Gly Ile Ile Thr Gln Lys Gly Ser Thr Asp Ala Gly Gly	
395 400 405	
GAT TTT GGA TTT GGA ATT TAC AAT GAT CAC CAC TAT CAT TTG GGG TAC	1300
Asp Phe Gly Phe Gly Ile Tyr Asn Asp His His Tyr His Leu Gly Tyr	
410 415 420	
TTC ATT TAT GGA ATT GCG GTG CTC ACT AAG CTT GAT CCA GCA TGG GGT	1348
Phe Ile Tyr Gly Ile Ala Val Leu Thr Lys Leu Asp Pro Ala Trp Gly	
425 430 435 440	
AGG AAG TAC AAG CCT CAA GCC TAT TCA ATA GTG CAA GAC TTC TTG AAC	1396
Arg Lys Tyr Lys Pro Gln Ala Tyr Ser Ile Val Gln Asp Phe Leu Asn	
445 450 455	
TTG GAC ACA AAA TTA AAC TCC AAT TAC ACA CGT TTG AGG TGT TTT GAC	1444
Leu Asp Thr Lys Leu Asn Ser Asn Tyr Thr Arg Leu Arg Cys Phe Asp	
460 465 470	
CCT TAT GTG CTT CAC TCT TGG GCT GGA GGG TTA ACT GAG TTC ACA GAT	1492
Pro Tyr Val Leu His Ser Trp Ala Gly Gly Leu Thr Glu Phe Thr Asp	
475 480 485	
GGA AGG AAT CAA GAG AGC ACA AGT GAG GCT GTG AGT GCA TAT TAT TCT	1540
Gly Arg Asn Gln Glu Ser Thr Ser Glu Ala Val Ser Ala Tyr Tyr Ser	
490 495 500	
GCT GCT TTG ATG GGA TTA GCA TAT GGT GAT GCA CCT CTT GTT GCA CTT	1588
Ala Ala Leu Met Gly Leu Ala Tyr Gly Asp Ala Pro Leu Val Ala Leu	
505 510 515 520	
GGA TCA ACA CTC ACA GCA TTG GAA ATT GAA GGG ACT AAA ATG TGG TGG	1636
Gly Ser Thr Leu Thr Ala Leu Glu Ile Glu Gly Thr Lys Met Trp Trp	

525	530	535	
CAT GTG AAA GAG GGA GGT ACT TTG TAT GAG AAA GAG TTT ACA CAA GAG			1684
His Val Lys Glu Gly Gly Thr Leu Tyr Glu Lys Glu Phe Thr Gln Glu			
540	545	550	
AAT AGG GTG ATG GGT GTT CTA TGG TCT AAC AAG AGG GAC ACT GGA CTT			1732
Asn Arg Val Met Gly Val Leu Trp Ser Asn Lys Arg Asp Thr Gly Leu			
555	560	565	
TGG TTT GCT CCT GCT GAG TGG AAA GAG TGT AGG CTT GGC ATT CAG CTC			1780
Trp Phe Ala Pro Ala Glu Trp Lys Glu Cys Arg Leu Gly Ile Gln Leu			
570	575	580	
TTA CCA TTG GCT CCT ATT TCT GAA GCC ATT TTC TCC AAT GTT GAC TTT			1828
Leu Pro Leu Ala Pro Ile Ser Glu Ala Ile Phe Ser Asn Val Asp Phe			
585	590	595	600
GTA AAG GAG CTT GTG GAG TGG ACT TTG CCT GCT TTG GAT AGG GAG GGT			1876
Val Lys Glu Leu Val Glu Trp Thr Leu Pro Ala Leu Asp Arg Glu Gly			
605	610	615	
GGT GTT GGT GAA GGA TGG AAG GGG TTT GTG TAT GCC CTT GAA GGG GTT			1924
Gly Val Gly Glu Gly Trp Lys Gly Phe Val Tyr Ala Leu Glu Gly Val			
620	625	630	
TAT GAC AAT GAA AGT GCA CTG CAG AAG ATA AGA AAC CTG AAA GGT TTT			1972
Tyr Asp Asn Glu Ser Ala Leu Gln Lys Ile Arg Asn Leu Lys Gly Phe			
635	640	645	
GAT GGT GGA AAC TCT TTG ACC AAT CTC TTG TGG TGG ATT CAT AGC AGA			2020
Asp Gly Gly Asn Ser Leu Thr Asn Leu Leu Trp Trp Ile His Ser Arg			
650	655	660	
AGT GAT GAA TAGAATGGAT TTGGTTATTG AACATGACAA CACTGCTGAT			2069
Ser Asp Glu			
665			
TTGGTTATTA TTATTATTGC TACTATTACT ATGCCAATAA AATCATCAAA TATGTCTTAC			2129

Val	Asn	Ile	Gln	Thr	Asn	Thr	Ser	Tyr	Ile	Phe	Pro	Gln	Thr	Gln	Ser
1				5					10					15	
Thr	Val	Leu	Pro	Asp	Pro	Ser	Lys	Phe	Phe	Ser	Ser	Asn	Leu	Leu	Ser
			20					25					30		
Ser	Pro	Leu	Pro	Thr	Asn	Ser	Phe	Phe	Gln	Asn	Phe	Val	Leu	Lys	Asn
		35					40					45			
Gly	Asp	Gln	Gln	Glu	Tyr	Ile	His	Pro	Tyr	Leu	Ile	Lys	Ser	Ser	Asn
	50					55					60				
Ser	Ser	Leu	Ser	Leu	Ser	Tyr	Pro	Ser	Arg	Gln	Ala	Ser	Ser	Ala	Val
65					70					75					80
Ile	Phe	Gln	Val	Phe	Asn	Pro	Asp	Leu	Thr	Ile	Ser	Ala	Pro	Gln	Gly
				85					90					95	
Pro	Lys	Gln	Gly	Pro	Pro	Gly	Lys	His	Leu	Ile	Ser	Ser	Tyr	Ser	Asp
			100					105					110		
Leu	Ser	Val	Thr	Leu	Asp	Phe	Pro	Ser	Ser	Asn	Leu	Ser	Phe	Phe	Leu
		115						120				125			
Val	Arg	Gly	Ser	Pro	Tyr	Leu	Thr	Val	Ser	Val	Thr	Gln	Pro	Thr	Pro
	130					135					140				
Leu	Ser	Ile	Thr	Thr	Ile	His	Ser	Ile	Leu	Ser	Phe	Ser	Ser	Asn	Asp
145					150					155					160

Ser Asn Thr Lys Tyr Thr Phe Gln Phe Asn Asn Gly Gln Thr Trp Leu
 165 170 175
 Leu Tyr Ala Thr Ser Pro Ile Lys Leu Asn His Thr Leu Ser Glu Ile
 180 185 190
 Thr Ser Asn Ala Phe Ser Gly Ile Ile Arg Ile Ala Leu Leu Pro Asp
 195 200 205
 Ser Asp Ser Lys His Glu Ala Val Leu Asp Lys Tyr Ser Ser Cys Tyr
 210 215 220
 Pro Val Ser Gly Lys Ala Val Phe Arg Glu Pro Phe Cys Val Glu Tyr
 225 230 235 240
 Asn Trp Glu Lys Lys Asp Ser Gly Asp Leu Leu Leu Leu Ala His Pro
 245 250 255
 Leu His Val Gln Leu Leu Arg Asn Gly Asp Asn Asp Val Lys Ile Leu
 260 265 270
 Glu Asp Leu Lys Tyr Lys Ser Ile Asp Gly Asp Leu Val Gly Val Val
 275 280 285
 Gly Asp Ser Trp Val Leu Lys Thr Asp Pro Leu Phe Val Thr Trp His
 290 295 300
 Ser Ile Lys Gly Ile Lys Glu Glu Ser His Asp Glu Ile Val Ser Ala
 305 310 315 320
 Leu Ser Lys Asp Val Glu Ser Leu Asp Ser Ser Ser Ile Thr Thr Thr
 325 330 335
 Glu Ser Tyr Phe Tyr Gly Lys Leu Ile Ala Arg Ala Ala Arg Leu Val
 340 345 350
 Leu Ile Ala Glu Glu Leu Asn Tyr Pro Asp Val Ile Pro Lys Val Arg
 355 360 365
 Asn Phe Leu Lys Glu Thr Ile Glu Pro Trp Leu Glu Gly Thr Phe Ser
 370 375 380
 Gly Asn Gly Phe Leu His Asp Glu Lys Trp Gly Gly Ile Ile Thr Gln

385		390		395		400									
Lys	Gly	Ser	Thr	Asp	Ala	Gly	Gly	Asp	Phe	Gly	Phe	Gly	Ile	Tyr	Asn
				405					410					415	
Asp	His	His	Tyr	His	Leu	Gly	Tyr	Phe	Ile	Tyr	Gly	Ile	Ala	Val	Leu
				420					425					430	
Thr	Lys	Leu	Asp	Pro	Ala	Trp	Gly	Arg	Lys	Tyr	Lys	Pro	Gln	Ala	Tyr
				435					440					445	
Ser	Ile	Val	Gln	Asp	Phe	Leu	Asn	Leu	Asp	Thr	Lys	Leu	Asn	Ser	Asn
				450				455						460	
Tyr	Thr	Arg	Leu	Arg	Cys	Phe	Asp	Pro	Tyr	Val	Leu	His	Ser	Trp	Ala
465						470					475				480
Gly	Gly	Leu	Thr	Glu	Phe	Thr	Asp	Gly	Arg	Asn	Gln	Glu	Ser	Thr	Ser
						485				490				495	
Glu	Ala	Val	Ser	Ala	Tyr	Tyr	Ser	Ala	Ala	Leu	Met	Gly	Leu	Ala	Tyr
				500					505					510	
Gly	Asp	Ala	Pro	Leu	Val	Ala	Leu	Gly	Ser	Thr	Leu	Thr	Ala	Leu	Glu
				515					520					525	
Ile	Glu	Gly	Thr	Lys	Met	Trp	Trp	His	Val	Lys	Glu	Gly	Gly	Thr	Leu
				530				535						540	
Tyr	Glu	Lys	Glu	Phe	Thr	Gln	Glu	Asn	Arg	Val	Met	Gly	Val	Leu	Trp
545						550					555				560
Ser	Asn	Lys	Arg	Asp	Thr	Gly	Leu	Trp	Phe	Ala	Pro	Ala	Glu	Trp	Lys
						565				570				575	
Glu	Cys	Arg	Leu	Gly	Ile	Gln	Leu	Leu	Pro	Leu	Ala	Pro	Ile	Ser	Glu
						580				585				590	
Ala	Ile	Phe	Ser	Asn	Val	Asp	Phe	Val	Lys	Glu	Leu	Val	Glu	Trp	Thr
				595				600						605	
Leu	Pro	Ala	Leu	Asp	Arg	Glu	Gly	Gly	Val	Gly	Glu	Gly	Trp	Lys	Gly
				610				615						620	

Phe Val Tyr Ala Leu Glu Gly Val Tyr Asp Asn Glu Ser Ala Leu Gln
 625 630 635 640
 Lys Ile Arg Asn Leu Lys Gly Phe Asp Gly Gly Asn Ser Leu Thr Asn
 645 650 655
 Leu Leu Trp Trp Ile His Ser Arg Ser Asp Glu
 660 665

SEQ ID NO: 4

SEQUENCE LENGTH: 2291

SEQUENCE TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

FEATURE

NAME/KEY: CDS

LOCATION: 80..2080

SEQUENCE DESCRIPTION:

CTTCTTTCCT CAACCTTCTT TCTTCTTATA TATTCGAACG ATCCGGGCCG CTCGAATTAA 60
 ACCACCTTCA GCAACAATG GTT AAC ATC CAA ACC AAT ACA TCT TAC ATC TTC 112
 Val Asn Ile Gln Thr Asn Thr Ser Tyr Ile Phe
 1 5 10
 CCT CAA ACA CAA TCC ACT GTT CTT CCT GAT CCC TCC AAA TTC TTC TCC 160
 Pro Gln Thr Gln Ser Thr Val Leu Pro Asp Pro Ser Lys Phe Phe Ser
 15 20 25
 TCA AAC CTT CTC TCA AGT CCA CTC CCC ACA AAC TCT TTC TTC CAA AAC 208
 Ser Asn Leu Leu Ser Ser Pro Leu Pro Thr Asn Ser Phe Phe Gln Asn
 30 35 40
 TTT GTC CTA AAA AAT GGT GAC CAA CAA GAA TAC ATT CAT CCT TAC CTC 256
 Phe Val Leu Lys Asn Gly Asp Gln Gln Glu Tyr Ile His Pro Tyr Leu

45	50	55	
ATC AAA TCC TCC AAC TCT TCC CTC TCT CTC TCA TAC CCT TCT CGC CAA			304
Ile Lys Ser Ser Asn Ser Ser Leu Ser Leu Ser Tyr Pro Ser Arg Gln			
60	65	70	75
GCC AGT TCA GCT GTC ATA TTC CAA GTC TTC AAT CCT GAT CTT ACC ATT			352
Ala Ser Ser Ala Val Ile Phe Gln Val Phe Asn Pro Asp Leu Thr Ile			
	80	85	90
TCA GCC CCA CAA GGT CCC AAA CAA GGT CCC CCT GGT AAA CAC CTT ATC			400
Ser Ala Pro Gln Gly Pro Lys Gln Gly Pro Pro Gly Lys His Leu Ile			
	95	100	105
TCC TCC TAC AGT GAT CTC AGT GTC ACC TTG GAT TTC CCT TCT TCC AAT			448
Ser Ser Tyr Ser Asp Leu Ser Val Thr Leu Asp Phe Pro Ser Ser Asn			
	110	115	120
CTG AGC TTC TTC CTT GTT AGG GGA AGC CCC TAT TTG ACT GTG TCT GTG			496
Leu Ser Phe Phe Leu Val Arg Gly Ser Pro Tyr Leu Thr Val Ser Val			
	125	130	135
ACT CAA CCA ACT CCT CTT TCA ATT ACC ACC ATC CAT TCC ATT CTC TCA			544
Thr Gln Pro Thr Pro Leu Ser Ile Thr Thr Ile His Ser Ile Leu Ser			
140	145	150	155
TTC TCT TCA AAT GAC TCC AAC ACC AAG TAC ACC TTT CAG TTC AAC AAT			592
Phe Ser Ser Asn Asp Ser Asn Thr Lys Tyr Thr Phe Gln Phe Asn Asn			
	160	165	170
GGT CAA ACA TGG CTT CTT TAT GCT ACC TCC CCC ATC AAG TTG AAC CAC			640
Gly Gln Thr Trp Leu Leu Tyr Ala Thr Ser Pro Ile Lys Leu Asn His			
	175	180	185
ACC CTT TCT GAG ATA ACT TCT AAT GCA TTT TCT GGC ATA ATC CGG ATA			688
Thr Leu Ser Glu Ile Thr Ser Asn Ala Phe Ser Gly Ile Ile Arg Ile			
	190	195	200
GCT TTG TTG CCG GAT TCG GAT TCG AAA CAC GAG GCT GTT CTT GAC AAG			736

Ala	Leu	Leu	Pro	Asp	Ser	Asp	Ser	Lys	His	Glu	Ala	Val	Leu	Asp	Lys		
205							210					215					
TAT	AGT	TCT	TGT	TAC	CCC	GTG	TCA	GGT	AAA	GCT	GTG	TTC	AGA	GAA	CCT	784	
Tyr	Ser	Ser	Cys	Tyr	Pro	Val	Ser	Gly	Lys	Ala	Val	Phe	Arg	Glu	Pro		
220					225					230				235			
TTC	TGT	GTG	GAA	TAT	AAC	TGG	GAG	AAG	AAA	GAT	TCA	GGG	GAT	TTG	CTA	832	
Phe	Cys	Val	Glu	Tyr	Asn	Trp	Glu	Lys	Lys	Asp	Ser	Gly	Asp	Leu	Leu		
				240					245					250			
CTC	TTG	GCT	CAC	CCT	CTC	CAT	GTT	CAG	CTT	CTT	CGT	AAT	GGA	GAC	AAT	880	
Leu	Leu	Ala	His	Pro	Leu	His	Val	Gln	Leu	Leu	Arg	Asn	Gly	Asp	Asn		
			255					260					265				
GAT	GTC	AAA	ATT	CTT	GAA	GAT	TTA	AAG	TAT	AAA	AGC	ATT	GAT	GGG	GAT	928	
Asp	Val	Lys	Ile	Leu	Glu	Asp	Leu	Lys	Tyr	Lys	Ser	Ile	Asp	Gly	Asp		
		270						275				280					
CTT	GTT	GGT	GTT	GTC	GGG	GAT	TCA	TGG	GTT	TTG	AAA	ACA	GAT	CCT	TTG	976	
Leu	Val	Gly	Val	Val	Gly	Asp	Ser	Trp	Val	Leu	Lys	Thr	Asp	Pro	Leu		
		285						290				295					
TTT	GTA	ACA	TGG	CAT	TCA	ATC	AAG	GGA	ATC	AAA	GAA	GAA	TCC	CAT	GAT	1024	
Phe	Val	Thr	Trp	His	Ser	Ile	Lys	Gly	Ile	Lys	Glu	Glu	Ser	His	Asp		
300					305					310				315			
GAG	ATT	GTC	TCA	GCC	CTT	TCT	AAA	GAT	GTT	GAG	AGC	CTA	GAT	TCA	TCA	1072	
Glu	Ile	Val	Ser	Ala	Leu	Ser	Lys	Asp	Val	Glu	Ser	Leu	Asp	Ser	Ser		
				320						325				330			
TCA	ATA	ACT	ACA	ACA	GAG	TCA	TAT	TTT	TAT	GGG	AAG	TTG	ATT	GCA	AGG	1120	
Ser	Ile	Thr	Thr	Thr	Glu	Ser	Tyr	Phe	Tyr	Gly	Lys	Leu	Ile	Ala	Arg		
		335						340				345					
GCT	GCA	AGG	TTG	GTA	TTG	ATT	GCT	GAG	GAG	TTG	AAC	TAC	CCT	GAT	GTG	1168	
Ala	Ala	Arg	Leu	Val	Leu	Ile	Ala	Glu	Glu	Leu	Asn	Tyr	Pro	Asp	Val		
		350						355				360					

ATT CCA AAG GTT AGG AAT TTT TTG AAA GAA ACC ATT GAG CCA TGG TTG	1216
Ile Pro Lys Val Arg Asn Phe Leu Lys Glu Thr Ile Glu Pro Trp Leu	
365 370 375	
GAG GGA ACT TTT AGT GGG AAT GGA TTC CTA CAT GAT GAA AAA TGG GGT	1264
Glu Gly Thr Phe Ser Gly Asn Gly Phe Leu His Asp Glu Lys Trp Gly	
380 385 390 395	
GGC ATT ATT ACC CAA AAG GGG TCC ACT GAT GCT GGT GGT GAT TTT GGA	1312
Gly Ile Ile Thr Gln Lys Gly Ser Thr Asp Ala Gly Gly Asp Phe Gly	
400 405 410	
TTT GGA ATT TAC AAT GAT CAC CAC TAT CAT TTG GGG TAC TTC ATT TAT	1360
Phe Gly Ile Tyr Asn Asp His His Tyr His Leu Gly Tyr Phe Ile Tyr	
415 420 425	
GGA ATT GCG GTG CTC ACT AAG CTT GAT CCA GCA TGG GGT AGG AAG TAC	1408
Gly Ile Ala Val Leu Thr Lys Leu Asp Pro Ala Trp Gly Arg Lys Tyr	
430 435 440	
AAG CCT CAA GCC TAT TCA ATA GTG CAA GAC TTC TTG AAC TTG GAC ACA	1456
Lys Pro Gln Ala Tyr Ser Ile Val Gln Asp Phe Leu Asn Leu Asp Thr	
445 450 455	
AAA TTA AAC TCC AAT TAC ACA CGT TTG AGG TGT TTT GAC CCT TAT GTG	1504
Lys Leu Asn Ser Asn Tyr Thr Arg Leu Arg Cys Phe Asp Pro Tyr Val	
460 465 470 475	
CTT CAC TCT TGG GCT GGA GGG TTA ACT GAG TTC ACA GAT GGA AGG AAT	1552
Leu His Ser Trp Ala Gly Gly Leu Thr Glu Phe Thr Asp Gly Arg Asn	
480 485 490	
CAA GAG AGC ACA AGT GAG GCT GTG AGT GCA TAT TAT TCT GCT GCT TTG	1600
Gln Glu Ser Thr Ser Glu Ala Val Ser Ala Tyr Tyr Ser Ala Ala Leu	
495 500 505	
ATG GGA TTA GCA TAT GGT GAT GCA CCT CTT GTT GCA CTT GGA TCA ACA	1648
Met Gly Leu Ala Tyr Gly Asp Ala Pro Leu Val Ala Leu Gly Ser Thr	

510	515	520	
CTC ACA GCA TTG GAA ATT GAA GGG ACT AAA ATG TGG TGG CAT GTG AAA			1696
Leu Thr Ala Leu Glu Ile Glu Gly Thr Lys Met Trp Trp His Val Lys			
525	530	535	
GAG GGA GGT ACT TTG TAT GAG AAA GAG TTT ACA CAA GAG AAT AGG GTG			1744
Glu Gly Gly Thr Leu Tyr Glu Lys Glu Phe Thr Gln Glu Asn Arg Val			
540	545	550	555
ATG GGT GTT CTA TGG TCT AAC AAG AGG GAC ACT GGA CTT TGG TTT GCT			1792
Met Gly Val Leu Trp Ser Asn Lys Arg Asp Thr Gly Leu Trp Phe Ala			
560	565	570	
CCT GCT GAG TGG AAA GAG TGT AGG CTT GGC ATT CAG CTC TTA CCA TTG			1840
Pro Ala Glu Trp Lys Glu Cys Arg Leu Gly Ile Gln Leu Leu Pro Leu			
575	580	585	
GCT CCT ATT TCT GAA GCC ATT TTC TCC AAT GTT GAC TTT GTA AAG GAG			1888
Ala Pro Ile Ser Glu Ala Ile Phe Ser Asn Val Asp Phe Val Lys Glu			
590	595	600	
CTT GTG GAG TGG ACT TTG CCT GCT TTG GAT AGG GAG GGT GGT GTT GGT			1936
Leu Val Glu Trp Thr Leu Pro Ala Leu Asp Arg Glu Gly Gly Val Gly			
605	610	615	
GAA GGA TGG AAG GGG TTT GTG TAT GCC CTT GAA GGG GTT TAT GAC AAT			1984
Glu Gly Trp Lys Gly Phe Val Tyr Ala Leu Glu Gly Val Tyr Asp Asn			
620	625	630	635
GAA AGT GCA CTG CAG AAG ATA AGA AAC CTG AAA GGT TTT GAT GGT GGA			2032
Glu Ser Ala Leu Gln Lys Ile Arg Asn Leu Lys Gly Phe Asp Gly Gly			
640	645	650	
AAC TCT TTG ACC AAT CTC TTG TGG TGG ATT CAT AGC AGA AGT GAT GAA			2080
Asn Ser Leu Thr Asn Leu Leu Trp Trp Ile His Ser Arg Ser Asp Glu			
655	660	665	
TAGAATGGAT TTGGTTATTG AACATGACAA CACTGCTGAT TTGGTTATTA TTATTATTGC			2140

TACTATTACT ATGCCAATAA AATCATCAAA TATGTCTTAC ATCATAATAA AAAAATTATT 2200
ATTATGCCAA TAAGTCAATG ATTTGATAT GTTATAGAAG AAGAAAATA TAATAAGTT 2260
ACAACAATAT CATGATCTAT AACATGTCTG A 2291

SEQ ID NO: 5

SEQUENCE LENGTH: 1054

SEQUENCE TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION:

Met Asp Phe Ile Ser Ser Leu Ile Val Gly Cys Ala Gln Val Leu Cys
1 5 10 15
Glu Ser Met Asn Met Ala Glu Arg Arg Gly His Lys Thr Asp Leu Arg
20 25 30
Gln Ala Ile Thr Asp Leu Glu Thr Ala Ile Gly Asp Leu Lys Ala Ile
35 40 45
Arg Asp Asp Leu Thr Leu Arg Ile Gln Gln Asp Gly Leu Glu Gly Arg
50 55 60
Ser Cys Ser Asn Arg Ala Arg Glu Trp Leu Ser Ala Val Gln Val Thr
65 70 75 80
Glu Thr Lys Thr Ala Leu Leu Leu Val Arg Phe Arg Arg Arg Glu Gln
85 90 95
Arg Thr Arg Met Arg Arg Arg Tyr Leu Ser Cys Phe Gly Cys Ala Asp
100 105 110
Tyr Lys Leu Cys Lys Lys Val Ser Ala Ile Leu Lys Ser Ile Gly Glu
115 120 125
Leu Arg Glu Arg Ser Glu Ala Ile Lys Thr Asp Gly Gly Ser Ile Gln
130 135 140
Val Thr Cys Arg Glu Ile Pro Ile Lys Ser Val Val Gly Asn Thr Thr

145		150		155		160									
Met	Met	Glu	Gln	Val	Leu	Glu	Phe	Leu	Ser	Glu	Glu	Glu	Glu	Arg	Gly
		165				170				175					
Ile	Ile	Gly	Val	Tyr	Gly	Pro	Gly	Gly	Val	Gly	Lys	Thr	Thr	Leu	Met
		180				185				190					
Gln	Ser	Ile	Asn	Asn	Glu	Leu	Ile	Thr	Lys	Gly	His	Gln	Tyr	Asp	Val
	195					200				205					
Leu	Ile	Trp	Val	Gln	Met	Ser	Arg	Glu	Phe	Gly	Glu	Cys	Thr	Ile	Gln
	210					215				220					
Gln	Ala	Val	Gly	Ala	Arg	Leu	Gly	Leu	Ser	Trp	Asp	Glu	Lys	Glu	Thr
225					230					235				240	
Gly	Glu	Asn	Arg	Ala	Leu	Lys	Ile	Tyr	Arg	Ala	Leu	Arg	Gln	Lys	Arg
		245				250				255					
Phe	Leu	Leu	Leu	Leu	Asp	Asp	Val	Trp	Glu	Glu	Ile	Asp	Leu	Glu	Lys
		260				265				270					
Thr	Gly	Val	Pro	Arg	Pro	Asp	Arg	Glu	Asn	Lys	Cys	Lys	Val	Met	Phe
	275					280				285					
Thr	Thr	Arg	Ser	Ile	Ala	Leu	Cys	Asn	Asn	Met	Gly	Ala	Glu	Tyr	Lys
	290					295				300					
Leu	Arg	Val	Glu	Phe	Leu	Glu	Lys	Lys	His	Ala	Trp	Glu	Leu	Phe	Cys
305					310					315				320	
Ser	Lys	Val	Trp	Arg	Lys	Asp	Leu	Leu	Glu	Ser	Ser	Ser	Ile	Arg	Arg
		325				330				335					
Leu	Ala	Glu	Ile	Ile	Val	Ser	Lys	Cys	Gly	Gly	Leu	Pro	Leu	Ala	Leu
		340				345				350					
Ile	Thr	Leu	Gly	Gly	Ala	Met	Ala	Gly	Leu	Gln	Glu	Phe	Asp	Pro	Gly
	355					360				365					
Ser	Pro	Gly	Leu	Gln	Glu	Phe	Ala	Ala	Ala	Arg	Ile	Lys	Pro	Pro	Ser
	370					375				380					

Ala Thr Met Val Asn Ile Gln Thr Asn Thr Ser Tyr Ile Phe Pro Gln			
385	390	395	400
Thr Gln Ser Thr Val Leu Pro Asp Pro Ser Lys Phe Phe Ser Ser Asn			
	405	410	415
Leu Leu Ser Ser Pro Leu Pro Thr Asn Ser Phe Phe Gln Asn Phe Val			
	420	425	430
Leu Lys Asn Gly Asp Gln Gln Glu Tyr Ile His Pro Tyr Leu Ile Lys			
	435	440	445
Ser Ser Asn Ser Ser Leu Ser Leu Ser Tyr Pro Ser Arg Gln Ala Ser			
	450	455	460
Ser Ala Val Ile Phe Gln Val Phe Asn Pro Asp Leu Thr Ile Ser Ala			
465	470	475	480
Pro Gln Gly Pro Lys Gln Gly Pro Pro Gly Lys His Leu Ile Ser Ser			
	485	490	495
Tyr Ser Asp Leu Ser Val Thr Leu Asp Phe Pro Ser Ser Asn Leu Ser			
	500	505	510
Phe Phe Leu Val Arg Gly Ser Pro Tyr Leu Thr Val Ser Val Thr Gln			
	515	520	525
Pro Thr Pro Leu Ser Ile Thr Thr Ile His Ser Ile Leu Ser Phe Ser			
	530	535	540
Ser Asn Asp Ser Asn Thr Lys Tyr Thr Phe Gln Phe Asn Asn Gly Gln			
545	550	555	560
Thr Trp Leu Leu Tyr Ala Thr Ser Pro Ile Lys Leu Asn His Thr Leu			
	565	570	575
Ser Glu Ile Thr Ser Asn Ala Phe Ser Gly Ile Ile Arg Ile Ala Leu			
	580	585	590
Leu Pro Asp Ser Asp Ser Lys His Glu Ala Val Leu Asp Lys Tyr Ser			
	595	600	605
Ser Cys Tyr Pro Val Ser Gly Lys Ala Val Phe Arg Glu Pro Phe Cys			

610	615	620
Val Glu Tyr Asn Trp Glu Lys Lys Asp Ser Gly Asp Leu Leu Leu Leu		
625	630	635
Ala His Pro Leu His Val Gln Leu Leu Arg Asn Gly Asp Asn Asp Val		640
	645	650
		655
Lys Ile Leu Glu Asp Leu Lys Tyr Lys Ser Ile Asp Gly Asp Leu Val		
	660	665
		670
Gly Val Val Gly Asp Ser Trp Val Leu Lys Thr Asp Pro Leu Phe Val		
	675	680
		685
Thr Trp His Ser Ile Lys Gly Ile Lys Glu Glu Ser His Asp Glu Ile		
	690	695
		700
Val Ser Ala Leu Ser Lys Asp Val Glu Ser Leu Asp Ser Ser Ser Ile		
705	710	715
		720
Thr Thr Thr Glu Ser Tyr Phe Tyr Gly Lys Leu Ile Ala Arg Ala Ala		
	725	730
		735
Arg Leu Val Leu Ile Ala Glu Glu Leu Asn Tyr Pro Asp Val Ile Pro		
	740	745
		750
Lys Val Arg Asn Phe Leu Lys Glu Thr Ile Glu Pro Trp Leu Glu Gly		
	755	760
		765
Thr Phe Ser Gly Asn Gly Phe Leu His Asp Glu Lys Trp Gly Gly Ile		
	770	775
		780
Ile Thr Gln Lys Gly Ser Thr Asp Ala Gly Gly Asp Phe Gly Phe Gly		
785	790	795
		800
Ile Tyr Asn Asp His His Tyr His Leu Gly Tyr Phe Ile Tyr Gly Ile		
	805	810
		815
Ala Val Leu Thr Lys Leu Asp Pro Ala Trp Gly Arg Lys Tyr Lys Pro		
	820	825
		830
Gln Ala Tyr Ser Ile Val Gln Asp Phe Leu Asn Leu Asp Thr Lys Leu		
	835	840
		845

```

Asn Ser Asn Tyr Thr Arg Leu Arg Cys Phe Asp Pro Tyr Val Leu His
      850                      855                      860
Ser Trp Ala Gly Gly Leu Thr Glu Phe Thr Asp Gly Arg Asn Gln Glu
865                      870                      875                      880
Ser Thr Ser Glu Ala Val Ser Ala Tyr Tyr Ser Ala Ala Leu Met Gly
                      885                      890                      895
Leu Ala Tyr Gly Asp Ala Pro Leu Val Ala Leu Gly Ser Thr Leu Thr
                      900                      905                      910
Ala Leu Glu Ile Glu Gly Thr Lys Met Trp Trp His Val Lys Glu Gly
                      915                      920                      925
Gly Thr Leu Tyr Glu Lys Glu Phe Thr Gln Glu Asn Arg Val Met Gly
                      930                      935                      940
Val Leu Trp Ser Asn Lys Arg Asp Thr Gly Leu Trp Phe Ala Pro Ala
945                      950                      955                      960
Glu Trp Lys Glu Cys Arg Leu Gly Ile Gln Leu Leu Pro Leu Ala Pro
                      965                      970                      975
Ile Ser Glu Ala Ile Phe Ser Asn Val Asp Phe Val Lys Glu Leu Val
                      980                      985                      990
Glu Trp Thr Leu Pro Ala Leu Asp Arg Glu Gly Gly Val Gly Glu Gly
                      995                      1000                      1005
Trp Lys Gly Phe Val Tyr Ala Leu Glu Gly Val Tyr Asp Asn Glu Ser
                      1010                      1015                      1020
Ala Leu Gln Lys Ile Arg Asn Leu Lys Gly Phe Asp Gly Gly Asn Ser
1025                      1030                      1035                      1040
Leu Thr Asn Leu Leu Trp Trp Ile His Ser Arg Ser Asp Glu
                      1045                      1050

```

SEQ ID NO: 6

SEQUENCE LENGTH: 3405

SEQUENCE TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

FEATURE

NAME/KEY: CDS

LOCATION: 33..3194

SEQUENCE DESCRIPTION:

ACAAGTAAAA GAAAGAGCGA GAAATCATCG AA ATG GAT TTC ATC TCA TCT CTT	53
Met Asp Phe Ile Ser Ser Leu	
1 5	
ATC GTT GGC TGT GCT CAG GTG TTG TGT GAA TCT ATG AAT ATG GCG GAG	101
Ile Val Gly Cys Ala Gln Val Leu Cys Glu Ser Met Asn Met Ala Glu	
10 15 20	
AGA AGA GGA CAT AAG ACT GAT CTT AGA CAA GCC ATC ACT GAT CTT GAA	149
Arg Arg Gly His Lys Thr Asp Leu Arg Gln Ala Ile Thr Asp Leu Glu	
25 30 35	
ACA GCC ATC GGT GAC TTG AAG GCC ATA CGT GAT GAC CTG ACT TTA CGG	197
Thr Ala Ile Gly Asp Leu Lys Ala Ile Arg Asp Asp Leu Thr Leu Arg	
40 45 50 55	
ATC CAA CAA GAC GGT CTA GAG GGA CGA AGC TGC TCA AAT CGT GCC AGA	245
Ile Gln Gln Asp Gly Leu Glu Gly Arg Ser Cys Ser Asn Arg Ala Arg	
60 65 70	
GAG TGG CTT AGT GCG GTG CAA GTA ACG GAG ACT AAA ACA GCC CTA CTT	293
Glu Trp Leu Ser Ala Val Gln Val Thr Glu Thr Lys Thr Ala Leu Leu	
75 80 85	
TTA GTG AGG TTT AGG CGT CGG GAA CAG AGG ACG CGA ATG AGG AGG AGA	341
Leu Val Arg Phe Arg Arg Arg Glu Gln Arg Thr Arg Met Arg Arg Arg	
90 95 100	

TAC CTC AGT TGT TTC GGT TGT GCC GAC TAC AAA CTG TGC AAG AAG GTT	389
Tyr Leu Ser Cys Phe Gly Cys Ala Asp Tyr Lys Leu Cys Lys Lys Val	
105 110 115	
TCT GCC ATA TTG AAG AGC ATT GGT GAG CTG AGA GAA CGC TCT GAA GCT	437
Ser Ala Ile Leu Lys Ser Ile Gly Glu Leu Arg Glu Arg Ser Glu Ala	
120 125 130 135	
ATC AAA ACA GAT GGC GGG TCA ATT CAA GTA ACT TGT AGA GAG ATA CCC	485
Ile Lys Thr Asp Gly Gly Ser Ile Gln Val Thr Cys Arg Glu Ile Pro	
140 145 150	
ATC AAG TCC GTT GTC GGA AAT ACC ACG ATG ATG GAA CAG GTT TTG GAA	533
Ile Lys Ser Val Val Gly Asn Thr Thr Met Met Glu Gln Val Leu Glu	
155 160 165	
TTT CTC AGT GAA GAA GAA GAA AGA GGA ATC ATT GGT GTT TAT GGA CCT	581
Phe Leu Ser Glu Glu Glu Glu Arg Gly Ile Ile Gly Val Tyr Gly Pro	
170 175 180	
GGT GGG GTT GGG AAG ACA ACG TTA ATG CAG AGC ATT AAC AAC GAG CTG	629
Gly Gly Val Gly Lys Thr Thr Leu Met Gln Ser Ile Asn Asn Glu Leu	
185 190 195	
ATC ACA AAA GGA CAT CAG TAT GAT GTA CTG ATT TGG GTT CAA ATG TCC	677
Ile Thr Lys Gly His Gln Tyr Asp Val Leu Ile Trp Val Gln Met Ser	
200 205 210 215	
AGA GAA TTC GGC GAG TGT ACA ATT CAG CAA GCC GTT GGA GCA CGG TTG	725
Arg Glu Phe Gly Glu Cys Thr Ile Gln Gln Ala Val Gly Ala Arg Leu	
220 225 230	
GGT TTA TCT TGG GAC GAG AAG GAG ACC GGC GAA AAC AGA GCT TTG AAG	773
Gly Leu Ser Trp Asp Glu Lys Glu Thr Gly Glu Asn Arg Ala Leu Lys	
235 240 245	
ATA TAC AGA GCT TTG AGA CAG AAA CGT TTC TTG TTG TTG CTA GAT GAT	821
Ile Tyr Arg Ala Leu Arg Gln Lys Arg Phe Leu Leu Leu Leu Asp Asp	

250	255	260	
GTC TGG GAA GAG ATA GAC TTG GAG AAA ACT GGA GTT CCT CGA CCT GAC			869
Val Trp Glu Glu Ile Asp Leu Glu Lys Thr Gly Val Pro Arg Pro Asp			
265	270	275	
AGG GAA AAC AAA TGC AAG GTG ATG TTC ACG ACA CGG TCT ATA GCA TTA			917
Arg Glu Asn Lys Cys Lys Val Met Phe Thr Thr Arg Ser Ile Ala Leu			
280	285	290	295
TGC AAC AAT ATG GGT GCG GAA TAC AAG TTG AGA GTG GAG TTT CTG GAG			965
Cys Asn Asn Met Gly Ala Glu Tyr Lys Leu Arg Val Glu Phe Leu Glu			
300	305	310	
AAG AAA CAC GCG TGG GAG CTG TTC TGT AGT AAG GTA TGG AGA AAA GAT			1013
Lys Lys His Ala Trp Glu Leu Phe Cys Ser Lys Val Trp Arg Lys Asp			
315	320	325	
CTT TTA GAG TCA TCA TCA ATT CGC CGG CTC GCG GAG ATT ATA GTG AGT			1061
Leu Leu Glu Ser Ser Ser Ile Arg Arg Leu Ala Glu Ile Ile Val Ser			
330	335	340	
AAA TGT GGA GGA TTG CCA CTA GCG TTG ATC ACT TTA GGA GGA GCC ATG			1109
Lys Cys Gly Gly Leu Pro Leu Ala Leu Ile Thr Leu Gly Gly Ala Met			
345	350	355	
GCT GGG CTG CAG GAA TTC GAT CCC GGA TCC CCC GGG CTG CAG GAA TTC			1157
Ala Gly Leu Gln Glu Phe Asp Pro Gly Ser Pro Gly Leu Gln Glu Phe			
360	365	370	375
GCG GCC GCT CGA ATT AAA CCA CCT TCA GCA ACA ATG GTT AAC ATC CAA			1205
Ala Ala Ala Arg Ile Lys Pro Pro Ser Ala Thr Met Val Asn Ile Gln			
380	385	390	
ACC AAT ACA TCT TAC ATC TTC CCT CAA ACA CAA TCC ACT GTT CTT CCT			1253
Thr Asn Thr Ser Tyr Ile Phe Pro Gln Thr Gln Ser Thr Val Leu Pro			
395	400	405	
GAT CCC TCC AAA TTC TTC TCC TCA AAC CTT CTC TCA AGT CCA CTC CCC			1301

Asp	Pro	Ser	Lys	Phe	Phe	Ser	Ser	Asn	Leu	Leu	Ser	Ser	Pro	Leu	Pro		
410				415				420									
ACA	AAC	TCT	TTC	TTC	CAA	AAC	TTT	GTC	CTA	AAA	AAT	GGT	GAC	CAA	CAA	1349	
Thr	Asn	Ser	Phe	Phe	Gln	Asn	Phe	Val	Leu	Lys	Asn	Gly	Asp	Gln	Gln		
425				430				435									
GAA	TAC	ATT	CAT	CCT	TAC	CTC	ATC	AAA	TCC	TCC	AAC	TCT	TCC	CTC	TCT	1397	
Glu	Tyr	Ile	His	Pro	Tyr	Leu	Ile	Lys	Ser	Ser	Asn	Ser	Ser	Leu	Ser		
440				445				450				455					
CTC	TCA	TAC	CCT	TCT	CGC	CAA	GCC	AGT	TCA	GCT	GTC	ATA	TTC	CAA	GTC	1445	
Leu	Ser	Tyr	Pro	Ser	Arg	Gln	Ala	Ser	Ser	Ala	Val	Ile	Phe	Gln	Val		
				460				465				470					
TTC	AAT	CCT	GAT	CTT	ACC	ATT	TCA	GCC	CCA	CAA	GGT	CCC	AAA	CAA	GGT	1493	
Phe	Asn	Pro	Asp	Leu	Thr	Ile	Ser	Ala	Pro	Gln	Gly	Pro	Lys	Gln	Gly		
				475				480				485					
CCC	CCT	GGT	AAA	CAC	CTT	ATC	TCC	TCC	TAC	AGT	GAT	CTC	AGT	GTC	ACC	1541	
Pro	Pro	Gly	Lys	His	Leu	Ile	Ser	Ser	Tyr	Ser	Asp	Leu	Ser	Val	Thr		
490				495				500									
TTG	GAT	TTC	CCT	TCT	TCC	AAT	CTG	AGC	TTC	TTC	CTT	GTT	AGG	GGA	AGC	1589	
Leu	Asp	Phe	Pro	Ser	Ser	Asn	Leu	Ser	Phe	Phe	Leu	Val	Arg	Gly	Ser		
505				510				515									
CCC	TAT	TTG	ACT	GTG	TCT	GTG	ACT	CAA	CCA	ACT	CCT	CTT	TCA	ATT	ACC	1637	
Pro	Tyr	Leu	Thr	Val	Ser	Val	Thr	Gln	Pro	Thr	Pro	Leu	Ser	Ile	Thr		
520				525				530				535					
ACC	ATC	CAT	TCC	ATT	CTC	TCA	TTC	TCT	TCA	AAT	GAC	TCC	AAC	ACC	AAG	1685	
Thr	Ile	His	Ser	Ile	Leu	Ser	Phe	Ser	Ser	Asn	Asp	Ser	Asn	Thr	Lys		
				540				545				550					
TAC	ACC	TTT	CAG	TTC	AAC	AAT	GGT	CAA	ACA	TGG	CTT	CTT	TAT	GCT	ACC	1733	
Tyr	Thr	Phe	Gln	Phe	Asn	Asn	Gly	Gln	Thr	Trp	Leu	Leu	Tyr	Ala	Thr		
555				560				565									

TCC CCC ATC AAG TTG AAC CAC ACC CTT TCT GAG ATA ACT TCT AAT GCA	1781
Ser Pro Ile Lys Leu Asn His Thr Leu Ser Glu Ile Thr Ser Asn Ala	
570 575 580	
TTT TCT GGC ATA ATC CGG ATA GCT TTG TTG CCG GAT TCG GAT TCG AAA	1829
Phe Ser Gly Ile Ile Arg Ile Ala Leu Leu Pro Asp Ser Asp Ser Lys	
585 590 595	
CAC GAG GCT GTT CTT GAC AAG TAT AGT TCT TGT TAC CCC GTG TCA GGT	1877
His Glu Ala Val Leu Asp Lys Tyr Ser Ser Cys Tyr Pro Val Ser Gly	
600 605 610 615	
AAA GCT GTG TTC AGA GAA CCT TTC TGT GTG GAA TAT AAC TGG GAG AAG	1925
Lys Ala Val Phe Arg Glu Pro Phe Cys Val Glu Tyr Asn Trp Glu Lys	
620 625 630	
AAA GAT TCA GGG GAT TTG CTA CTC TTG GCT CAC CCT CTC CAT GTT CAG	1973
Lys Asp Ser Gly Asp Leu Leu Leu Leu Ala His Pro Leu His Val Gln	
635 640 645	
CTT CTT CGT AAT GGA GAC AAT GAT GTC AAA ATT CTT GAA GAT TTA AAG	2021
Leu Leu Arg Asn Gly Asp Asn Asp Val Lys Ile Leu Glu Asp Leu Lys	
650 655 660	
TAT AAA AGC ATT GAT GGG GAT CTT GTT GGT GTT GTC GGG GAT TCA TGG	2069
Tyr Lys Ser Ile Asp Gly Asp Leu Val Gly Val Val Gly Asp Ser Trp	
665 670 675	
GTT TTG AAA ACA GAT CCT TTG TTT GTA ACA TGG CAT TCA ATC AAG GGA	2117
Val Leu Lys Thr Asp Pro Leu Phe Val Thr Trp His Ser Ile Lys Gly	
680 685 690 695	
ATC AAA GAA GAA TCC CAT GAT GAG ATT GTC TCA GCC CTT TCT AAA GAT	2165
Ile Lys Glu Glu Ser His Asp Glu Ile Val Ser Ala Leu Ser Lys Asp	
700 705 710	
GTT GAG AGC CTA GAT TCA TCA TCA ATA ACT ACA ACA GAG TCA TAT TTT	2213
Val Glu Ser Leu Asp Ser Ser Ser Ile Thr Thr Thr Glu Ser Tyr Phe	

715	720	725	
TAT GGG AAG TTG ATT GCA AGG GCT GCA AGG TTG GTA TTG ATT GCT GAG			2261
Tyr Gly Lys Leu Ile Ala Arg Ala Ala Arg Leu Val Leu Ile Ala Glu			
730	735	740	
GAG TTG AAC TAC CCT GAT GTG ATT CCA AAG GTT AGG AAT TTT TTG AAA			2309
Glu Leu Asn Tyr Pro Asp Val Ile Pro Lys Val Arg Asn Phe Leu Lys			
745	750	755	
GAA ACC ATT GAG CCA TGG TTG GAG GGA ACT TTT AGT GGG AAT GGA TTC			2357
Glu Thr Ile Glu Pro Trp Leu Glu Gly Thr Phe Ser Gly Asn Gly Phe			
760	765	770	775
CTA CAT GAT GAA AAA TGG GGT GGC ATT ATT ACC CAA AAG GGG TCC ACT			2405
Leu His Asp Glu Lys Trp Gly Gly Ile Ile Thr Gln Lys Gly Ser Thr			
780	785	790	
GAT GCT GGT GGT GAT TTT GGA TTT GGA ATT TAC AAT GAT CAC CAC TAT			2453
Asp Ala Gly Gly Asp Phe Gly Phe Gly Ile Tyr Asn Asp His His Tyr			
795	800	805	
CAT TTG GGG TAC TTC ATT TAT GGA ATT GCG GTG CTC ACT AAG CTT GAT			2501
His Leu Gly Tyr Phe Ile Tyr Gly Ile Ala Val Leu Thr Lys Leu Asp			
810	815	820	
CCA GCA TGG GGT AGG AAG TAC AAG CCT CAA GCC TAT TCA ATA GTG CAA			2549
Pro Ala Trp Gly Arg Lys Tyr Lys Pro Gln Ala Tyr Ser Ile Val Gln			
825	830	835	
GAC TTC TTG AAC TTG GAC ACA AAA TTA AAC TCC AAT TAC ACA CGT TTG			2597
Asp Phe Leu Asn Leu Asp Thr Lys Leu Asn Ser Asn Tyr Thr Arg Leu			
840	845	850	855
AGG TGT TTT GAC CCT TAT GTG CTT CAC TCT TGG GCT GGA GGG TTA ACT			2645
Arg Cys Phe Asp Pro Tyr Val Leu His Ser Trp Ala Gly Gly Leu Thr			
860	865	870	
GAG TTC ACA GAT GGA AGG AAT CAA GAG AGC ACA AGT GAG GCT GTG AGT			2693

Glu Phe Thr Asp Gly Arg Asn Gln Glu Ser Thr Ser Glu Ala Val Ser	
875	880
885	
GCA TAT TAT TCT GCT GCT TTG ATG GGA TTA GCA TAT GGT GAT GCA CCT	2741
Ala Tyr Tyr Ser Ala Ala Leu Met Gly Leu Ala Tyr Gly Asp Ala Pro	
890	895
900	
CTT GTT GCA CTT GGA TCA ACA CTC ACA GCA TTG GAA ATT GAA GGG ACT	2789
Leu Val Ala Leu Gly Ser Thr Leu Thr Ala Leu Glu Ile Glu Gly Thr	
905	910
915	
AAA ATG TGG TGG CAT GTG AAA GAG GGA GGT ACT TTG TAT GAG AAA GAG	2837
Lys Met Trp Trp His Val Lys Glu Gly Gly Thr Leu Tyr Glu Lys Glu	
920	925
930	935
TTT ACA CAA GAG AAT AGG GTG ATG GGT GTT CTA TGG TCT AAC AAG AGG	2885
Phe Thr Gln Glu Asn Arg Val Met Gly Val Leu Trp Ser Asn Lys Arg	
940	945
950	
GAC ACT GGA CTT TGG TTT GCT CCT GCT GAG TGG AAA GAG TGT AGG CTT	2933
Asp Thr Gly Leu Trp Phe Ala Pro Ala Glu Trp Lys Glu Cys Arg Leu	
955	960
965	
GGC ATT CAG CTC TTA CCA TTG GCT CCT ATT TCT GAA GCC ATT TTC TCC	2981
Gly Ile Gln Leu Leu Pro Leu Ala Pro Ile Ser Glu Ala Ile Phe Ser	
970	975
980	
AAT GTT GAC TTT GTA AAG GAG CTT GTG GAG TGG ACT TTG CCT GCT TTG	3029
Asn Val Asp Phe Val Lys Glu Leu Val Glu Trp Thr Leu Pro Ala Leu	
985	990
995	
GAT AGG GAG GGT GGT GTT GGT GAA GGA TGG AAG GGG TTT GTG TAT GCC	3077
Asp Arg Glu Gly Gly Val Gly Glu Gly Trp Lys Gly Phe Val Tyr Ala	
1000	1005
1010	1015
CTT GAA GGG GTT TAT GAC AAT GAA AGT GCA CTG CAG AAG ATA AGA AAC	3125
Leu Glu Gly Val Tyr Asp Asn Glu Ser Ala Leu Gln Lys Ile Arg Asn	
1020	1025
1030	

CTG AAA GGT TTT GAT GGT GGA AAC TCT TTG ACC AAT CTC TTG TGG TGG 3173
 Leu Lys Gly Phe Asp Gly Gly Asn Ser Leu Thr Asn Leu Leu Trp Trp

1035

1040

1045

ATT CAT AGC AGA AGT GAT GAA TAGAATGGAT TTGGTTATTG AACATGACAA 3224
 Ile His Ser Arg Ser Asp Glu

1050

CACTGCTGAT TTGGTTATTA TTATTATTGC TACTATTACT ATGCCAATAA AATCATCAAA 3284
 TATGTCTTAC ATCATAATAA AAAAATTATT ATTATGCCAA TAAGTCAATG ATTCGATAT 3344
 GTTATAGAAG AAGAAACTA TAATAAAGTT ACAACAATAT CATGATCTAT AACATGTCTG 3404
 A 3405

SEQ ID NO: 7

SEQUENCE LENGTH: 877

SEQUENCE TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION:

Val Asn Ile Gln Thr Asn Thr Ser Tyr Ile Phe Pro Gln Thr Gln Ser
 1 5 10 15
 Thr Val Leu Pro Asp Pro Ser Lys Phe Phe Ser Ser Asn Leu Leu Ser
 20 25 30
 Ser Pro Leu Pro Thr Asn Ser Phe Phe Gln Asn Phe Val Leu Lys Asn
 35 40 45
 Gly Asp Gln Gln Glu Tyr Ile His Pro Tyr Leu Ile Lys Ser Ser Asn
 50 55 60
 Ser Ser Leu Ser Leu Ser Tyr Pro Ser Arg Gln Ala Ser Ser Ala Val
 65 70 75 80
 Ile Phe Gln Val Phe Asn Pro Asp Leu Thr Ile Ser Ala Pro Gln Gly
 85 90 95

Pro Lys Gln Gly Pro Pro Gly Lys His Leu Ile Ser Ser Tyr Ser Asp
 100 105 110
 Leu Ser Val Thr Leu Asp Phe Pro Ser Ser Asn Leu Ser Phe Phe Leu
 115 120 125
 Val Arg Gly Ser Pro Tyr Leu Thr Val Ser Val Thr Gln Pro Thr Pro
 130 135 140
 Leu Ser Ile Thr Thr Ile His Ser Ile Leu Ser Phe Ser Ser Asn Asp
 145 150 155 160
 Ser Asn Thr Lys Tyr Thr Phe Gln Phe Asn Asn Gly Gln Thr Trp Leu
 165 170 175
 Leu Tyr Ala Thr Ser Pro Ile Lys Leu Asn His Thr Leu Ser Glu Ile
 180 185 190
 Thr Ser Asn Ala Phe Ser Gly Ile Ile Arg Ile Ala Leu Leu Pro Asp
 195 200 205
 Ser Asp Ser Lys His Glu Ala Val Leu Asp Lys Tyr Ser Ser Cys Tyr
 210 215 220
 Pro Val Ser Gly Lys Ala Val Phe Arg Glu Pro Phe Cys Val Glu Tyr
 225 230 235 240
 Asn Trp Glu Lys Lys Asp Ser Gly Asp Leu Leu Leu Leu Ala His Pro
 245 250 255
 Leu His Val Gln Leu Leu Arg Asn Gly Asp Asn Asp Val Lys Ile Leu
 260 265 270
 Glu Asp Leu Lys Tyr Lys Ser Ile Asp Gly Asp Leu Val Gly Val Val
 275 280 285
 Gly Asp Ser Trp Val Leu Lys Thr Asp Pro Leu Phe Val Thr Trp His
 290 295 300
 Ser Ile Lys Gly Ile Lys Glu Glu Ser His Asp Glu Ile Val Ser Ala
 305 310 315 320
 Leu Ser Lys Asp Val Glu Ser Leu Asp Ser Ser Ser Ile Thr Thr Thr

	325		330		335
Glu Ser Tyr Phe Tyr Gly Lys Leu Ile Ala Arg Ala Ala Arg Leu Val					
	340		345		350
Leu Ile Ala Glu Glu Leu Asn Tyr Pro Asp Val Ile Pro Lys Val Arg					
	355		360		365
Asn Phe Leu Lys Glu Thr Ile Glu Pro Trp Leu Glu Gly Thr Phe Ser					
	370		375		380
Gly Asn Gly Phe Leu His Asp Glu Lys Trp Gly Gly Ile Ile Thr Gln					
385		390		395	400
Lys Gly Ser Thr Asp Ala Gly Gly Asp Phe Gly Phe Gly Ile Tyr Asn					
	405		410		415
Asp His His Tyr His Leu Gly Tyr Phe Ile Tyr Gly Ile Ala Val Leu					
	420		425		430
Thr Lys Leu Asp Pro Ala Trp Gly Arg Lys Tyr Lys Pro Gln Ala Tyr					
	435		440		445
Ser Ile Val Gln Asp Phe Leu Asn Leu Asp Thr Lys Leu Asn Ser Asn					
	450		455		460
Tyr Thr Arg Leu Arg Cys Phe Asp Pro Tyr Val Leu His Ser Trp Ala					
465		470		475	480
Gly Gly Leu Thr Glu Phe Thr Asp Gly Arg Asn Gln Glu Ser Thr Ser					
	485		490		495
Glu Ala Val Ser Ala Tyr Tyr Ser Ala Ala Leu Met Gly Leu Ala Tyr					
	500		505		510
Gly Asp Ala Pro Leu Val Ala Leu Gly Ser Thr Leu Thr Ala Leu Glu					
	515		520		525
Ile Glu Gly Thr Lys Met Trp Trp His Ala Met Asp Phe Ile Ser Ser					
	530		535		540
Leu Ile Val Gly Cys Ala Gln Val Leu Cys Glu Ser Met Asn Met Ala					
545		550		555	560

Glu Arg Arg Gly His Lys Thr Asp Leu Arg Gln Ala Ile Thr Asp Leu
 565 570 575
 Glu Thr Ala Ile Gly Asp Leu Lys Ala Ile Arg Asp Asp Leu Thr Leu
 580 585 590
 Arg Ile Gln Gln Asp Gly Leu Glu Gly Arg Ser Cys Ser Asn Arg Ala
 595 600 605
 Arg Glu Trp Leu Ser Ala Val Gln Val Thr Glu Thr Lys Thr Ala Leu
 610 615 620
 Leu Leu Val Arg Phe Arg Arg Arg Glu Gln Arg Thr Arg Met Arg Arg
 625 630 635 640
 Arg Tyr Leu Ser Cys Phe Gly Cys Ala Asp Tyr Lys Leu Cys Lys Lys
 645 650 655
 Val Ser Ala Ile Leu Lys Ser Ile Gly Glu Leu Arg Glu Arg Ser Glu
 660 665 670
 Ala Ile Lys Thr Asp Gly Gly Ser Ile Gln Val Thr Cys Arg Glu Ile
 675 680 685
 Pro Ile Lys Ser Val Val Gly Asn Thr Thr Met Met Glu Gln Val Leu
 690 695 700
 Glu Phe Leu Ser Glu Glu Glu Glu Arg Gly Ile Ile Gly Val Tyr Gly
 705 710 715 720
 Pro Gly Gly Val Gly Lys Thr Thr Leu Met Gln Ser Ile Asn Asn Glu
 725 730 735
 Leu Ile Thr Lys Gly His Gln Tyr Asp Val Leu Ile Trp Val Gln Met
 740 745 750
 Ser Arg Glu Phe Gly Glu Cys Thr Ile Gln Gln Ala Val Gly Ala Arg
 755 760 765
 Leu Gly Leu Ser Trp Asp Glu Lys Glu Thr Gly Glu Asn Arg Ala Leu
 770 775 780
 Lys Ile Tyr Arg Ala Leu Arg Gln Lys Arg Phe Leu Leu Leu Leu Asp

785	790	795	800
Asp Val Trp Glu Glu Ile Asp Leu Glu Lys Thr Gly Val Pro Arg Pro			
	805	810	815
Asp Arg Glu Asn Lys Cys Lys Val Met Phe Thr Thr Arg Ser Ile Ala			
	820	825	830
Leu Cys Asn Asn Met Gly Ala Glu Tyr Lys Leu Arg Val Glu Phe Leu			
	835	840	845
Glu Lys Lys His Ala Trp Glu Leu Phe Cys Ser Lys Val Trp Arg Lys			
	850	855	860
Asp Leu Leu Glu Ser Ser Ser Ile Arg Arg Leu Ala Glu			
865	870	875	

SEQ ID NO: 8

SEQUENCE LENGTH: 2666

SEQUENCE TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

FEATURE

NAME/KEY: CDS

LOCATION: 29..2659

SEQUENCE DESCRIPTION:

TCGAATTAAA CCACCTTCAG CAACAATG GTT AAC ATC CAA ACC AAT ACA TCT	52
Val Asn Ile Gln Thr Asn Thr Ser	
1 5	
TAC ATC TTC CCT CAA ACA CAA TCC ACT GTT CTT CCT GAT CCC TCC AAA	100
Tyr Ile Phe Pro Gln Thr Gln Ser Thr Val Leu Pro Asp Pro Ser Lys	
10 15 20	
TTC TTC TCC TCA AAC CTT CTC TCA AGT CCA CTC CCC ACA AAC TCT TTC	148

Phe	Phe	Ser	Ser	Asn	Leu	Leu	Ser	Ser	Pro	Leu	Pro	Thr	Asn	Ser	Phe	
25					30					35					40	
TTC	CAA	AAC	TTT	GTC	CTA	AAA	AAT	GGT	GAC	CAA	CAA	GAA	TAC	ATT	CAT	196
Phe	Gln	Asn	Phe	Val	Leu	Lys	Asn	Gly	Asp	Gln	Gln	Glu	Tyr	Ile	His	
				45					50					55		
CCT	TAC	CTC	ATC	AAA	TCC	TCC	AAC	TCT	TCC	CTC	TCT	CTC	TCA	TAC	CCT	244
Pro	Tyr	Leu	Ile	Lys	Ser	Ser	Asn	Ser	Ser	Leu	Ser	Leu	Ser	Tyr	Pro	
				60					65					70		
TCT	CGC	CAA	GCC	AGT	TCA	GCT	GTC	ATA	TTC	CAA	GTC	TTC	AAT	CCT	GAT	292
Ser	Arg	Gln	Ala	Ser	Ser	Ala	Val	Ile	Phe	Gln	Val	Phe	Asn	Pro	Asp	
				75				80						85		
CTT	ACC	ATT	TCA	GCC	CCA	CAA	GGT	CCC	AAA	CAA	GGT	CCC	CCT	GGT	AAA	340
Leu	Thr	Ile	Ser	Ala	Pro	Gln	Gly	Pro	Lys	Gln	Gly	Pro	Pro	Gly	Lys	
				90				95						100		
CAC	CTT	ATC	TCC	TCC	TAC	AGT	GAT	CTC	AGT	GTC	ACC	TTG	GAT	TTC	CCT	388
His	Leu	Ile	Ser	Ser	Tyr	Ser	Asp	Leu	Ser	Val	Thr	Leu	Asp	Phe	Pro	
105					110					115				120		
TCT	TCC	AAT	CTG	AGC	TTC	TTC	CTT	GTT	AGG	GGA	AGC	CCC	TAT	TTG	ACT	436
Ser	Ser	Asn	Leu	Ser	Phe	Phe	Leu	Val	Arg	Gly	Ser	Pro	Tyr	Leu	Thr	
					125				130					135		
GTG	TCT	GTG	ACT	CAA	CCA	ACT	CCT	CTT	TCA	ATT	ACC	ACC	ATC	CAT	TCC	484
Val	Ser	Val	Thr	Gln	Pro	Thr	Pro	Leu	Ser	Ile	Thr	Thr	Ile	His	Ser	
				140					145					150		
ATT	CTC	TCA	TTC	TCT	TCA	AAT	GAC	TCC	AAC	ACC	AAG	TAC	ACC	TTT	CAG	532
Ile	Leu	Ser	Phe	Ser	Ser	Asn	Asp	Ser	Asn	Thr	Lys	Tyr	Thr	Phe	Gln	
				155					160					165		
TTC	AAC	AAT	GGT	CAA	ACA	TGG	CTT	CTT	TAT	GCT	ACC	TCC	CCC	ATC	AAG	580
Phe	Asn	Asn	Gly	Gln	Thr	Trp	Leu	Leu	Tyr	Ala	Thr	Ser	Pro	Ile	Lys	
				170					175					180		

TTG AAC CAC ACC CTT TCT GAG ATA ACT TCT AAT GCA TTT TCT GGC ATA	628
Leu Asn His Thr Leu Ser Glu Ile Thr Ser Asn Ala Phe Ser Gly Ile	
185 190 195 200	
ATC CGG ATA GCT TTG TTG CCG GAT TCG GAT TCG AAA CAC GAG GCT GTT	676
Ile Arg Ile Ala Leu Leu Pro Asp Ser Asp Ser Lys His Glu Ala Val	
205 210 215	
CTT GAC AAG TAT AGT TCT TGT TAC CCC GTG TCA GGT AAA GCT GTG TTC	724
Leu Asp Lys Tyr Ser Ser Cys Tyr Pro Val Ser Gly Lys Ala Val Phe	
220 225 230	
AGA GAA CCT TTC TGT GTG GAA TAT AAC TGG GAG AAG AAA GAT TCA GGG	772
Arg Glu Pro Phe Cys Val Glu Tyr Asn Trp Glu Lys Lys Asp Ser Gly	
235 240 245	
GAT TTG CTA CTC TTG GCT CAC CCT CTC CAT GTT CAG CTT CTT CGT AAT	820
Asp Leu Leu Leu Leu Ala His Pro Leu His Val Gln Leu Leu Arg Asn	
250 255 260	
GGA GAC AAT GAT GTC AAA ATT CTT GAA GAT TTA AAG TAT AAA AGC ATT	868
Gly Asp Asn Asp Val Lys Ile Leu Glu Asp Leu Lys Tyr Lys Ser Ile	
265 270 275 280	
GAT GGG GAT CTT GTT GGT GTT GTC GGG GAT TCA TGG GTT TTG AAA ACA	916
Asp Gly Asp Leu Val Gly Val Val Gly Asp Ser Trp Val Leu Lys Thr	
285 290 295	
GAT CCT TTG TTT GTA ACA TGG CAT TCA ATC AAG GGA ATC AAA GAA GAA	964
Asp Pro Leu Phe Val Thr Trp His Ser Ile Lys Gly Ile Lys Glu Glu	
300 305 310	
TCC CAT GAT GAG ATT GTC TCA GCC CTT TCT AAA GAT GTT GAG AGC CTA	1012
Ser His Asp Glu Ile Val Ser Ala Leu Ser Lys Asp Val Glu Ser Leu	
315 320 325	
GAT TCA TCA TCA ATA ACT ACA ACA GAG TCA TAT TTT TAT GGG AAG TTG	1060
Asp Ser Ser Ser Ile Thr Thr Thr Glu Ser Tyr Phe Tyr Gly Lys Leu	

330	335	340	
ATT GCA AGG GCT GCA AGG TTG GTA TTG ATT GCT GAG GAG TTG AAC TAC			1108
Ile Ala Arg Ala Ala Arg Leu Val Leu Ile Ala Glu Glu Leu Asn Tyr			
345	350	355	360
CCT GAT GTG ATT CCA AAG GTT AGG AAT TTT TTG AAA GAA ACC ATT GAG			1156
Pro Asp Val Ile Pro Lys Val Arg Asn Phe Leu Lys Glu Thr Ile Glu			
	365	370	375
CCA TGG TTG GAG GGA ACT TTT AGT GGG AAT GGA TTC CTA CAT GAT GAA			1204
Pro Trp Leu Glu Gly Thr Phe Ser Gly Asn Gly Phe Leu His Asp Glu			
	380	385	390
AAA TGG GGT GGC ATT ATT ACC CAA AAG GGG TCC ACT GAT GCT GGT GGT			1252
Lys Trp Gly Gly Ile Ile Thr Gln Lys Gly Ser Thr Asp Ala Gly Gly			
	395	400	405
GAT TTT GGA TTT GGA ATT TAC AAT GAT CAC CAC TAT CAT TTG GGG TAC			1300
Asp Phe Gly Phe Gly Ile Tyr Asn Asp His His Tyr His Leu Gly Tyr			
	410	415	420
TTC ATT TAT GGA ATT GCG GTG CTC ACT AAG CTT GAT CCA GCA TGG GGT			1348
Phe Ile Tyr Gly Ile Ala Val Leu Thr Lys Leu Asp Pro Ala Trp Gly			
425	430	435	440
AGG AAG TAC AAG CCT CAA GCC TAT TCA ATA GTG CAA GAC TTC TTG AAC			1396
Arg Lys Tyr Lys Pro Gln Ala Tyr Ser Ile Val Gln Asp Phe Leu Asn			
	445	450	455
TTG GAC ACA AAA TTA AAC TCC AAT TAC ACA CGT TTG AGG TGT TTT GAC			1444
Leu Asp Thr Lys Leu Asn Ser Asn Tyr Thr Arg Leu Arg Cys Phe Asp			
	460	465	470
CCT TAT GTG CTT CAC TCT TGG GCT GGA GGG TTA ACT GAG TTC ACA GAT			1492
Pro Tyr Val Leu His Ser Trp Ala Gly Gly Leu Thr Glu Phe Thr Asp			
	475	480	485
GGA AGG AAT CAA GAG AGC ACA AGT GAG GCT GTG AGT GCA TAT TAT TCT			1540

Gly	Arg	Asn	Gln	Glu	Ser	Thr	Ser	Glu	Ala	Val	Ser	Ala	Tyr	Tyr	Ser	
490						495					500					
GCT	GCT	TTG	ATG	GGA	TTA	GCA	TAT	GGT	GAT	GCA	CCT	CTT	GTT	GCA	CTT	1588
Ala	Ala	Leu	Met	Gly	Leu	Ala	Tyr	Gly	Asp	Ala	Pro	Leu	Val	Ala	Leu	
505					510					515					520	
GGA	TCA	ACA	CTC	ACA	GCA	TTG	GAA	ATT	GAA	GGG	ACT	AAA	ATG	TGG	TGG	1636
Gly	Ser	Thr	Leu	Thr	Ala	Leu	Glu	Ile	Glu	Gly	Thr	Lys	Met	Trp	Trp	
				525					530					535		
CAT	GCG	ATG	GAT	TTC	ATC	TCA	TCT	CTT	ATC	GTT	GGC	TGT	GCT	CAG	GTG	1684
His	Ala	Met	Asp	Phe	Ile	Ser	Ser	Leu	Ile	Val	Gly	Cys	Ala	Gln	Val	
			540						545					550		
TTG	TGT	GAA	TCT	ATG	AAT	ATG	GCG	GAG	AGA	AGA	GGA	CAT	AAG	ACT	GAT	1732
Leu	Cys	Glu	Ser	Met	Asn	Met	Ala	Glu	Arg	Arg	Gly	His	Lys	Thr	Asp	
			555						560					565		
CTT	AGA	CAA	GCC	ATC	ACT	GAT	CTT	GAA	ACA	GCC	ATC	GGT	GAC	TTG	AAG	1780
Leu	Arg	Gln	Ala	Ile	Thr	Asp	Leu	Glu	Thr	Ala	Ile	Gly	Asp	Leu	Lys	
			570						575					580		
GCC	ATA	CGT	GAT	GAC	CTG	ACT	TTA	CGG	ATC	CAA	CAA	GAC	GGT	CTA	GAG	1828
Ala	Ile	Arg	Asp	Asp	Leu	Thr	Leu	Arg	Ile	Gln	Gln	Asp	Gly	Leu	Glu	
585					590					595					600	
GGA	CGA	AGC	TGC	TCA	AAT	CGT	GCC	AGA	GAG	TGG	CTT	AGT	GCG	GTG	CAA	1876
Gly	Arg	Ser	Cys	Ser	Asn	Arg	Ala	Arg	Glu	Trp	Leu	Ser	Ala	Val	Gln	
				605					610					615		
GTA	ACG	GAG	ACT	AAA	ACA	GCC	CTA	CTT	TTA	GTG	AGG	TTT	AGG	CGT	CGG	1924
Val	Thr	Glu	Thr	Lys	Thr	Ala	Leu	Leu	Leu	Val	Arg	Phe	Arg	Arg	Arg	
				620					625					630		
GAA	CAG	AGG	ACG	CGA	ATG	AGG	AGG	AGA	TAC	CTC	AGT	TGT	TTC	GGT	TGT	1972
Glu	Gln	Arg	Thr	Arg	Met	Arg	Arg	Arg	Tyr	Leu	Ser	Cys	Phe	Gly	Cys	
			635						640					645		

GCC GAC TAC AAA CTG TGC AAG AAG GTT TCT GCC ATA TTG AAG AGC ATT	2020
Ala Asp Tyr Lys Leu Cys Lys Lys Val Ser Ala Ile Leu Lys Ser Ile	
650 655 660	
GGT GAG CTG AGA GAA CGC TCT GAA GCT ATC AAA ACA GAT GGC GGG TCA	2068
Gly Glu Leu Arg Glu Arg Ser Glu Ala Ile Lys Thr Asp Gly Gly Ser	
665 670 675 680	
ATT CAA GTA ACT TGT AGA GAG ATA CCC ATC AAG TCC GTT GTC GGA AAT	2116
Ile Gln Val Thr Cys Arg Glu Ile Pro Ile Lys Ser Val Val Gly Asn	
685 690 695	
ACC ACG ATG ATG GAA CAG GTT TTG GAA TTT CTC AGT GAA GAA GAA GAA	2164
Thr Thr Met Met Glu Gln Val Leu Glu Phe Leu Ser Glu Glu Glu Glu	
700 705 710	
AGA GGA ATC ATT GGT GTT TAT GGA CCT GGT GGG GTT GGG AAG ACA ACG	2212
Arg Gly Ile Ile Gly Val Tyr Gly Pro Gly Gly Val Gly Lys Thr Thr	
715 720 725	
TTA ATG CAG AGC ATT AAC AAC GAG CTG ATC ACA AAA GGA CAT CAG TAT	2260
Leu Met Gln Ser Ile Asn Asn Glu Leu Ile Thr Lys Gly His Gln Tyr	
730 735 740	
GAT GTA CTG ATT TGG GTT CAA ATG TCC AGA GAA TTC GGC GAG TGT ACA	2308
Asp Val Leu Ile Trp Val Gln Met Ser Arg Glu Phe Gly Glu Cys Thr	
745 750 755 760	
ATT CAG CAA GCC GTT GGA GCA CGG TTG GGT TTA TCT TGG GAC GAG AAG	2356
Ile Gln Gln Ala Val Gly Ala Arg Leu Gly Leu Ser Trp Asp Glu Lys	
765 770 775	
GAG ACC GGC GAA AAC AGA GCT TTG AAG ATA TAC AGA GCT TTG AGA CAG	2404
Glu Thr Gly Glu Asn Arg Ala Leu Lys Ile Tyr Arg Ala Leu Arg Gln	
780 785 790	
AAA CGT TTC TTG TTG TTG CTA GAT GAT GTC TGG GAA GAG ATA GAC TTG	2452
Lys Arg Phe Leu Leu Leu Leu Asp Asp Val Trp Glu Glu Ile Asp Leu	

795	800	805	
GAG AAA ACT GGA GTT CCT CGA CCT GAC AGG GAA AAC AAA TGC AAG GTG			2500
Glu Lys Thr Gly Val Pro Arg Pro Asp Arg Glu Asn Lys Cys Lys Val			
810	815	820	
ATG TTC ACG ACA CGG TCT ATA GCA TTA TGC AAC AAT ATG GGT GCG GAA			2548
Met Phe Thr Thr Arg Ser Ile Ala Leu Cys Asn Asn Met Gly Ala Glu			
825	830	835	840
TAC AAG TTG AGA GTG GAG TTT CTG GAG AAG AAA CAC GCG TGG GAG CTG			2596
Tyr Lys Leu Arg Val Glu Phe Leu Glu Lys Lys His Ala Trp Glu Leu			
	845	850	855
TTC TGT AGT AAG GTA TGG AGA AAA GAT CTT TTA GAG TCA TCA TCA ATT			2644
Phe Cys Ser Lys Val Trp Arg Lys Asp Leu Leu Glu Ser Ser Ser Ile			
	860	865	870
CGC CGG CTC GCG GAG TGAATTC			2666
Arg Arg Leu Ala Glu			
875			

SEQ ID NO: 9

SEQUENCE LENGTH: 683

SEQUENCE TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION:

Asn Ala Phe Ser Gly Ile Ile Arg Ile Ala Leu Leu Pro Asp Ser Asp

1

5

10

15

Ser Lys His Glu Ala Val Leu Asp Lys Tyr Ser Ser Cys Tyr Pro Val

20

25

30

Ser Gly Lys Ala Val Phe Arg Glu Pro Phe Cys Val Glu Tyr Asn Trp

35

40

45

Glu Lys Lys Asp Ser Gly Asp Leu Leu Leu Leu Ala His Pro Leu His
 50 55 60
 Val Gln Leu Leu Arg Asn Gly Asp Asn Asp Val Lys Ile Leu Glu Asp
 65 70 75 80
 Leu Lys Tyr Lys Ser Ile Asp Gly Asp Leu Val Gly Val Val Gly Asp
 85 90 95
 Ser Trp Val Leu Lys Thr Asp Pro Leu Phe Val Thr Trp His Ser Ile
 100 105 110
 Lys Gly Ile Lys Glu Glu Ser His Asp Glu Ile Val Ser Ala Leu Ser
 115 120 125
 Lys Asp Val Glu Ser Leu Asp Ser Ser Ser Ile Thr Thr Thr Glu Ser
 130 135 140
 Tyr Phe Tyr Gly Lys Leu Ile Ala Arg Ala Ala Arg Leu Val Leu Ile
 145 150 155 160
 Ala Glu Glu Leu Asn Tyr Pro Asp Val Ile Pro Lys Val Arg Asn Phe
 165 170 175
 Leu Lys Glu Thr Ile Glu Pro Trp Leu Glu Gly Thr Phe Ser Gly Asn
 180 185 190
 Gly Phe Leu His Asp Glu Lys Trp Gly Gly Ile Ile Thr Gln Lys Gly
 195 200 205
 Ser Thr Asp Ala Gly Gly Asp Phe Gly Phe Gly Ile Tyr Asn Asp His
 210 215 220
 His Tyr His Leu Gly Tyr Phe Ile Tyr Gly Ile Ala Val Leu Thr Lys
 225 230 235 240
 Leu Asp Pro Ala Trp Gly Arg Lys Tyr Lys Pro Gln Ala Tyr Ser Ile
 245 250 255
 Val Gln Asp Phe Leu Asn Leu Asp Thr Lys Leu Asn Ser Asn Tyr Thr
 260 265 270
 Arg Leu Arg Cys Phe Asp Pro Tyr Val Leu His Ser Trp Ala Gly Gly

275	280	285
Leu Thr Glu Phe Thr Asp Gly Arg Asn Gln Glu Ser Thr Ser Glu Ala		
290	295	300
Val Ser Ala Tyr Tyr Ser Ala Ala Leu Met Gly Leu Ala Tyr Gly Asp		
305	310	315
Ala Pro Leu Val Ala Leu Gly Ser Thr Leu Thr Ala Leu Glu Ile Glu		
	325	330
		335
Gly Thr Lys Met Trp Trp His Ala Met Asp Phe Ile Ser Ser Leu Ile		
	340	345
		350
Val Gly Cys Ala Gln Val Leu Cys Glu Ser Met Asn Met Ala Glu Arg		
	355	360
		365
Arg Gly His Lys Thr Asp Leu Arg Gln Ala Ile Thr Asp Leu Glu Thr		
	370	375
		380
Ala Ile Gly Asp Leu Lys Ala Ile Arg Asp Asp Leu Thr Leu Arg Ile		
385	390	395
		400
Gln Gln Asp Gly Leu Glu Gly Arg Ser Cys Ser Asn Arg Ala Arg Glu		
	405	410
		415
Trp Leu Ser Ala Val Gln Val Thr Glu Thr Lys Thr Ala Leu Leu Leu		
	420	425
		430
Val Arg Phe Arg Arg Arg Glu Gln Arg Thr Arg Met Arg Arg Arg Tyr		
	435	440
		445
Leu Ser Cys Phe Gly Cys Ala Asp Tyr Lys Leu Cys Lys Lys Val Ser		
	450	455
		460
Ala Ile Leu Lys Ser Ile Gly Glu Leu Arg Glu Arg Ser Glu Ala Ile		
465	470	475
		480
Lys Thr Asp Gly Gly Ser Ile Gln Val Thr Cys Arg Glu Ile Pro Ile		
	485	490
		495
Lys Ser Val Val Gly Asn Thr Thr Met Met Glu Gln Val Leu Glu Phe		
	500	505
		510

Leu Ser Glu Glu Glu Glu Arg Gly Ile Ile Gly Val Tyr Gly Pro Gly
 515 520 525
 Gly Val Gly Lys Thr Thr Leu Met Gln Ser Ile Asn Asn Glu Leu Ile
 530 535 540
 Thr Lys Gly His Gln Tyr Asp Val Leu Ile Trp Val Gln Met Ser Arg
 545 550 555 560
 Glu Phe Gly Glu Cys Thr Ile Gln Gln Ala Val Gly Ala Arg Leu Gly
 565 570 575
 Leu Ser Trp Asp Glu Lys Glu Thr Gly Glu Asn Arg Ala Leu Lys Ile
 580 585 590
 Tyr Arg Ala Leu Arg Gln Lys Arg Phe Leu Leu Leu Leu Asp Asp Val
 595 600 605
 Trp Glu Glu Ile Asp Leu Glu Lys Thr Gly Val Pro Arg Pro Asp Arg
 610 615 620
 Glu Asn Lys Cys Lys Val Met Phe Thr Thr Arg Ser Ile Ala Leu Cys
 625 630 635 640
 Asn Asn Met Gly Ala Glu Tyr Lys Leu Arg Val Glu Phe Leu Glu Lys
 645 650 655
 Lys His Ala Trp Glu Leu Phe Cys Ser Lys Val Trp Arg Lys Asp Leu
 660 665 670
 Leu Glu Ser Ser Ser Ile Arg Arg Leu Ala Glu
 675 680

SEQ ID NO: 10

SEQUENCE LENGTH: 2117

SEQUENCE TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

FEATURE

NAME/KEY: CDS

LOCATION: 62..2110

SEQUENCE DESCRIPTION:

```

GCGGCCGCCT TCTTTCCTCA ACCTTCTTTC TTCTTATATA TTCGAACAAT TCGATATCAT    60
G AAT GCA TTT TCT GGC ATA ATC CGG ATA GCT TTG TTG CCG GAT TCG          106
  Asn Ala Phe Ser Gly Ile Ile Arg Ile Ala Leu Leu Pro Asp Ser
      1             5             10             15
GAT TCG AAA CAC GAG GCT GTT CTT GAC AAG TAT AGT TCT TGT TAC CCC          154
Asp Ser Lys His Glu Ala Val Leu Asp Lys Tyr Ser Ser Cys Tyr Pro
      20             25             30
GTG TCA GGT AAA GCT GTG TTC AGA GAA CCT TTC TGT GTG GAA TAT AAC          202
Val Ser Gly Lys Ala Val Phe Arg Glu Pro Phe Cys Val Glu Tyr Asn
      35             40             45
TGG GAG AAG AAA GAT TCA GGG GAT TTG CTA CTC TTG GCT CAC CCT CTC          250
Trp Glu Lys Lys Asp Ser Gly Asp Leu Leu Leu Leu Ala His Pro Leu
      50             55             60
CAT GTT CAG CTT CTT CGT AAT GGA GAC AAT GAT GTC AAA ATT CTT GAA          298
His Val Gln Leu Leu Arg Asn Gly Asp Asn Asp Val Lys Ile Leu Glu
      65             70             75
GAT TTA AAG TAT AAA AGC ATT GAT GGG GAT CTT GTT GGT GTT GTC GGG          346
Asp Leu Lys Tyr Lys Ser Ile Asp Gly Asp Leu Val Gly Val Val Gly
      80             85             90             95
GAT TCA TGG GTT TTG AAA ACA GAT CCT TTG TTT GTA ACA TGG CAT TCA          394
Asp Ser Trp Val Leu Lys Thr Asp Pro Leu Phe Val Thr Trp His Ser
      100            105            110
ATC AAG GGA ATC AAA GAA GAA TCC CAT GAT GAG ATT GTC TCA GCC CTT          442
Ile Lys Gly Ile Lys Glu Glu Ser His Asp Glu Ile Val Ser Ala Leu
      115            120            125

```

TCT AAA GAT GTT GAG AGC CTA GAT TCA TCA TCA ATA ACT ACA ACA GAG	490
Ser Lys Asp Val Glu Ser Leu Asp Ser Ser Ser Ile Thr Thr Thr Glu	
130 135 140	
TCA TAT TTT TAT GGG AAG TTG ATT GCA AGG GCT GCA AGG TTG GTA TTG	538
Ser Tyr Phe Tyr Gly Lys Leu Ile Ala Arg Ala Ala Arg Leu Val Leu	
145 150 155	
ATT GCT GAG GAG TTG AAC TAC CCT GAT GTG ATT CCA AAG GTT AGG AAT	586
Ile Ala Glu Glu Leu Asn Tyr Pro Asp Val Ile Pro Lys Val Arg Asn	
160 165 170 175	
TTT TTG AAA GAA ACC ATT GAG CCA TGG TTG GAG GGA ACT TTT AGT GGG	634
Phe Leu Lys Glu Thr Ile Glu Pro Trp Leu Glu Gly Thr Phe Ser Gly	
180 185 190	
AAT GGA TTC CTA CAT GAT GAA AAA TGG GGT GGC ATT ATT ACC CAA AAG	682
Asn Gly Phe Leu His Asp Glu Lys Trp Gly Gly Ile Ile Thr Gln Lys	
195 200 205	
GGG TCC ACT GAT GCT GGT GGT GAT TTT GGA TTT GGA ATT TAC AAT GAT	730
Gly Ser Thr Asp Ala Gly Gly Asp Phe Gly Phe Gly Ile Tyr Asn Asp	
210 215 220	
CAC CAC TAT CAT TTG GGG TAC TTC ATT TAT GGA ATT GCG GTG CTC ACT	778
His His Tyr His Leu Gly Tyr Phe Ile Tyr Gly Ile Ala Val Leu Thr	
225 230 235	
AAG CTT GAT CCA GCA TGG GGT AGG AAG TAC AAG CCT CAA GCC TAT TCA	826
Lys Leu Asp Pro Ala Trp Gly Arg Lys Tyr Lys Pro Gln Ala Tyr Ser	
240 245 250 255	
ATA GTG CAA GAC TTC TTG AAC TTG GAC ACA AAA TTA AAC TCC AAT TAC	874
Ile Val Gln Asp Phe Leu Asn Leu Asp Thr Lys Leu Asn Ser Asn Tyr	
260 265 270	
ACA CGT TTG AGG TGT TTT GAC CCT TAT GTG CTT CAC TCT TGG GCT GGA	922
Thr Arg Leu Arg Cys Phe Asp Pro Tyr Val Leu His Ser Trp Ala Gly	

275	280	285	
GGG TTA ACT GAG TTC ACA GAT GGA AGG AAT CAA GAG AGC ACA AGT GAG			970
Gly Leu Thr Glu Phe Thr Asp Gly Arg Asn Gln Glu Ser Thr Ser Glu			
290	295	300	
GCT GTG AGT GCA TAT TAT TCT GCT GCT TTG ATG GGA TTA GCA TAT GGT			1018
Ala Val Ser Ala Tyr Tyr Ser Ala Ala Leu Met Gly Leu Ala Tyr Gly			
305	310	315	
GAT GCA CCT CTT GTT GCA CTT GGA TCA ACA CTC ACA GCA TTG GAA ATT			1066
Asp Ala Pro Leu Val Ala Leu Gly Ser Thr Leu Thr Ala Leu Glu Ile			
320	325	330	335
GAA GGG ACT AAA ATG TGG TGG CAT GCG ATG GAT TTC ATC TCA TCT CTT			1114
Glu Gly Thr Lys Met Trp Trp His Ala Met Asp Phe Ile Ser Ser Leu			
340	345	350	
ATC GTT GGC TGT GCT CAG GTG TTG TGT GAA TCT ATG AAT ATG GCG GAG			1162
Ile Val Gly Cys Ala Gln Val Leu Cys Glu Ser Met Asn Met Ala Glu			
355	360	365	
AGA AGA GGA CAT AAG ACT GAT CTT AGA CAA GCC ATC ACT GAT CTT GAA			1210
Arg Arg Gly His Lys Thr Asp Leu Arg Gln Ala Ile Thr Asp Leu Glu			
370	375	380	
ACA GCC ATC GGT GAC TTG AAG GCC ATA CGT GAT GAC CTG ACT TTA CGG			1258
Thr Ala Ile Gly Asp Leu Lys Ala Ile Arg Asp Asp Leu Thr Leu Arg			
385	390	395	
ATC CAA CAA GAC GGT CTA GAG GGA CGA AGC TGC TCA AAT CGT GCC AGA			1306
Ile Gln Gln Asp Gly Leu Glu Gly Arg Ser Cys Ser Asn Arg Ala Arg			
400	405	410	415
GAG TGG CTT AGT GCG GTG CAA GTA ACG GAG ACT AAA ACA GCC CTA CTT			1354
Glu Trp Leu Ser Ala Val Gln Val Thr Glu Thr Lys Thr Ala Leu Leu			
420	425	430	
TTA GTG AGG TTT AGG CGT CGG GAA CAG AGG ACG CGA ATG AGG AGG AGA			1402

Leu	Val	Arg	Phe	Arg	Arg	Arg	Glu	Gln	Arg	Thr	Arg	Met	Arg	Arg	Arg		
			435					440					445				
TAC	CTC	AGT	TGT	TTC	GGT	TGT	GCC	GAC	TAC	AAA	CTG	TGC	AAG	AAG	GTT	1450	
Tyr	Leu	Ser	Cys	Phe	Gly	Cys	Ala	Asp	Tyr	Lys	Leu	Cys	Lys	Lys	Val		
			450					455					460				
TCT	GCC	ATA	TTG	AAG	AGC	ATT	GGT	GAG	CTG	AGA	GAA	CGC	TCT	GAA	GCT	1498	
Ser	Ala	Ile	Leu	Lys	Ser	Ile	Gly	Glu	Leu	Arg	Glu	Arg	Ser	Glu	Ala		
			465					470					475				
ATC	AAA	ACA	GAT	GGC	GGG	TCA	ATT	CAA	GTA	ACT	TGT	AGA	GAG	ATA	CCC	1546	
Ile	Lys	Thr	Asp	Gly	Gly	Ser	Ile	Gln	Val	Thr	Cys	Arg	Glu	Ile	Pro		
480						485				490					495		
ATC	AAG	TCC	GTT	GTC	GGA	AAT	ACC	ACG	ATG	ATG	GAA	CAG	GTT	TTG	GAA	1594	
Ile	Lys	Ser	Val	Val	Gly	Asn	Thr	Thr	Met	Met	Glu	Gln	Val	Leu	Glu		
				500					505					510			
TTT	CTC	AGT	GAA	GAA	GAA	GAA	AGA	GGA	ATC	ATT	GGT	GTT	TAT	GGA	CCT	1642	
Phe	Leu	Ser	Glu	Glu	Glu	Glu	Arg	Gly	Ile	Ile	Gly	Val	Tyr	Gly	Pro		
			515					520					525				
GGT	GGG	GTT	GGG	AAG	ACA	ACG	TTA	ATG	CAG	AGC	ATT	AAC	AAC	GAG	CTG	1690	
Gly	Gly	Val	Gly	Lys	Thr	Thr	Leu	Met	Gln	Ser	Ile	Asn	Asn	Glu	Leu		
			530					535					540				
ATC	ACA	AAA	GGA	CAT	CAG	TAT	GAT	GTA	CTG	ATT	TGG	GTT	CAA	ATG	TCC	1738	
Ile	Thr	Lys	Gly	His	Gln	Tyr	Asp	Val	Leu	Ile	Trp	Val	Gln	Met	Ser		
			545					550					555				
AGA	GAA	TTC	GGC	GAG	TGT	ACA	ATT	CAG	CAA	GCC	GTT	GGA	GCA	CGG	TTG	1786	
Arg	Glu	Phe	Gly	Glu	Cys	Thr	Ile	Gln	Gln	Ala	Val	Gly	Ala	Arg	Leu		
560						565				570					575		
GGT	TTA	TCT	TGG	GAC	GAG	AAG	GAG	ACC	GGC	GAA	AAC	AGA	GCT	TTG	AAG	1834	
Gly	Leu	Ser	Trp	Asp	Glu	Lys	Glu	Thr	Gly	Glu	Asn	Arg	Ala	Leu	Lys		
				580					585					590			

ATA TAC AGA GCT TTG AGA CAG AAA CGT TTC TTG TTG TTG CTA GAT GAT	1882
Ile Tyr Arg Ala Leu Arg Gln Lys Arg Phe Leu Leu Leu Leu Asp Asp	
595 600 605	
GTC TGG GAA GAG ATA GAC TTG GAG AAA ACT GGA GTT CCT CGA CCT GAC	1930
Val Trp Glu Glu Ile Asp Leu Glu Lys Thr Gly Val Pro Arg Pro Asp	
610 615 620	
AGG GAA AAC AAA TGC AAG GTG ATG TTC ACG ACA CGG TCT ATA GCA TTA	1978
Arg Glu Asn Lys Cys Lys Val Met Phe Thr Thr Arg Ser Ile Ala Leu	
625 630 635	
TGC AAC AAT ATG GGT GCG GAA TAC AAG TTG AGA GTG GAG TTT CTG GAG	2026
Cys Asn Asn Met Gly Ala Glu Tyr Lys Leu Arg Val Glu Phe Leu Glu	
640 645 650 655	
AAG AAA CAC GCG TGG GAG CTG TTC TGT AGT AAG GTA TGG AGA AAA GAT	2074
Lys Lys His Ala Trp Glu Leu Phe Cys Ser Lys Val Trp Arg Lys Asp	
660 665 670	
CTT TTA GAG TCA TCA TCA ATT CGC CGG CTC GCG GAG TGAATTC	2117
Leu Leu Glu Ser Ser Ser Ile Arg Arg Leu Ala Glu	
675 680	

SEQ ID NO: 11

SEQUENCE LENGTH: 347

SEQUENCE TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION:

Met Ala Lys Tyr His Ser Ser Gly Lys Ser Ser Ser Met Thr Ala Ile

1

5

10

15

Ala Phe Leu Phe Ile Leu Leu Ile Thr Tyr Thr Gly Thr Thr Asp Ala

20

25

30

Gln Ser Gly Val Cys Tyr Gly Arg Leu Gly Asn Asn Leu Pro Thr Pro
 35 40 45
 Gln Glu Val Val Ala Leu Tyr Asn Gln Ala Asn Ile Arg Arg Met Arg
 50 55 60
 Ile Tyr Gly Pro Ser Pro Glu Val Leu Glu Ala Leu Arg Gly Ser Asn
 65 70 75 80
 Ile Glu Leu Leu Leu Asp Ile Pro Asn Asp Asn Leu Arg Asn Leu Ala
 85 90 95
 Ser Ser Gln Asp Asn Ala Asn Lys Trp Val Gln Asp Asn Ile Lys Asn
 100 105 110
 Tyr Ala Asn Asn Val Arg Phe Arg Tyr Val Ser Val Gly Asn Glu Val
 115 120 125
 Lys Pro Glu His Ser Phe Ala Gln Phe Leu Val Pro Ala Leu Glu Asn
 130 135 140
 Ile Gln Arg Ala Ile Ser Asn Ala Gly Leu Gly Asn Gln Val Lys Val
 145 150 155 160
 Ser Thr Ala Ile Asp Thr Gly Ala Leu Ala Glu Ser Phe Pro Pro Ser
 165 170 175
 Lys Gly Ser Phe Lys Ser Asp Tyr Arg Gly Ala Tyr Leu Asp Gly Val
 180 185 190
 Ile Arg Phe Leu Val Asn Asn Asn Ala Pro Leu Met Val Asn Val Tyr
 195 200 205
 Ser Tyr Phe Ala Tyr Thr Ala Asn Pro Lys Asp Ile Ser Leu Asp Tyr
 210 215 220
 Ala Leu Phe Arg Ser Pro Ser Val Val Val Gln Asp Gly Ser Leu Gly
 225 230 235 240
 Tyr Arg Asn Leu Phe Asp Ala Ser Val Asp Ala Val Tyr Ala Ala Leu
 245 250 255
 Glu Lys Ala Gly Gly Gly Ser Leu Asn Ile Val Val Ser Glu Ser Gly

260	265	270
Trp Pro Ser Ser Gly Gly Thr Ala Thr Ser Leu Asp Asn Ala Arg Thr		
275	280	285
Tyr Asn Thr Asn Leu Val Arg Asn Val Lys Gln Gly Thr Pro Lys Arg		
290	295	300
Pro Gly Ala Pro Leu Glu Thr Tyr Val Phe Ala Met Phe Asp Glu Asn		
305	310	315
Gln Lys Gln Pro Glu Phe Glu Lys Phe Trp Gly Leu Phe Ser Pro Ile		
325	330	335
Thr Lys Gln Pro Lys Tyr Ser Ile Asn Phe Asn		
340	345	

SEQ ID NO: 12

SEQUENCE LENGTH: 1044

SEQUENCE TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

SEQUENCE DESCRIPTION:

ATG GCT AAG TAT CAT TCA AGT GGG AAA AGC TCT TCC ATG ACT GCT ATA	48
Met Ala Lys Tyr His Ser Ser Gly Lys Ser Ser Ser Met Thr Ala Ile	
1 5 10 15	
GCC TTC CTG TTT ATC CTT CTA ATC ACT TAT ACA GGC ACA ACA GAT GCA	96
Ala Phe Leu Phe Ile Leu Leu Ile Thr Tyr Thr Gly Thr Thr Asp Ala	
20 25 30	
CAA TCC GGG GTA TGT TAT GGA AGA CTT GGC AAC AAC TTA CCA ACC CCT	144
Gln Ser Gly Val Cys Tyr Gly Arg Leu Gly Asn Asn Leu Pro Thr Pro	
35 40 45	
CAA GAA GTT GTG GCC CTC TAC AAT CAA GCC AAC ATT CGC AGG ATG CGA	192

Gln	Glu	Val	Val	Ala	Leu	Tyr	Asn	Gln	Ala	Asn	Ile	Arg	Arg	Met	Arg	
50				55				60								
ATC	TAC	GGT	CCA	AGC	CCA	GAA	GTC	CTC	GAA	GCA	CTA	AGA	GGT	TCC	AAC	240
Ile	Tyr	Gly	Pro	Ser	Pro	Glu	Val	Leu	Glu	Ala	Leu	Arg	Gly	Ser	Asn	
65				70				75				80				
ATT	GAG	CTT	TTG	CTA	GAC	ATT	CCA	AAT	GAC	AAC	CTC	AGA	AAC	CTA	GCA	288
Ile	Glu	Leu	Leu	Leu	Asp	Ile	Pro	Asn	Asp	Asn	Leu	Arg	Asn	Leu	Ala	
				85				90				95				
TCT	AGC	CAA	GAC	AAT	GCA	AAC	AAA	TGG	GTG	CAA	GAC	AAC	ATC	AAA	AAC	336
Ser	Ser	Gln	Asp	Asn	Ala	Asn	Lys	Trp	Val	Gln	Asp	Asn	Ile	Lys	Asn	
				100				105				110				
TAT	GCC	AAC	AAT	GTC	AGA	TTC	AGA	TAC	GTT	TCA	GTG	GGA	AAT	GAA	GTG	384
Tyr	Ala	Asn	Asn	Val	Arg	Phe	Arg	Tyr	Val	Ser	Val	Gly	Asn	Glu	Val	
				115				120				125				
AAA	CCC	GAA	CAC	TCA	TTT	GCA	CAA	TTT	CTA	GTG	CCT	GCA	TTG	GAA	AAC	432
Lys	Pro	Glu	His	Ser	Phe	Ala	Gln	Phe	Leu	Val	Pro	Ala	Leu	Glu	Asn	
				130				135				140				
ATT	CAG	AGG	GCC	ATT	TCT	AAT	GCT	GGC	CTT	GGA	AAC	CAA	GTA	AAA	GTT	480
Ile	Gln	Arg	Ala	Ile	Ser	Asn	Ala	Gly	Leu	Gly	Asn	Gln	Val	Lys	Val	
145				150				155				160				
TCC	ACT	GCC	ATT	GAT	ACT	GGT	GCC	TTG	GCA	GAA	TCA	TTC	CCA	CCA	TCA	528
Ser	Thr	Ala	Ile	Asp	Thr	Gly	Ala	Leu	Ala	Glu	Ser	Phe	Pro	Pro	Ser	
				165				170				175				
AAG	GGT	TCC	TTC	AAA	TCT	GAT	TAT	AGA	GGA	GCA	TAT	CTT	GAT	GGT	GTC	576
Lys	Gly	Ser	Phe	Lys	Ser	Asp	Tyr	Arg	Gly	Ala	Tyr	Leu	Asp	Gly	Val	
				180				185				190				
ATC	AGA	TTT	CTA	GTG	AAC	AAT	AAT	GCC	CCA	TTA	ATG	GTT	AAT	GTG	TAC	624
Ile	Arg	Phe	Leu	Val	Asn	Asn	Asn	Ala	Pro	Leu	Met	Val	Asn	Val	Tyr	
				195				200				205				

TCT TAC TTC GCT TAC ACT GCA AAC CCT AAG GAC ATT AGT CTT GAC TAT	672
Ser Tyr Phe Ala Tyr Thr Ala Asn Pro Lys Asp Ile Ser Leu Asp Tyr	
210 215 220	
GCA CTT TTT AGG TCT CCT TCG GTG GTA GTG CAA GAT GGT TCA CTT GGT	720
Ala Leu Phe Arg Ser Pro Ser Val Val Val Gln Asp Gly Ser Leu Gly	
225 230 235 240	
TAC CGT AAC CTC TTT GAT GCT TCG GTT GAT GCT GTT TAT GCT GCA TTG	768
Tyr Arg Asn Leu Phe Asp Ala Ser Val Asp Ala Val Tyr Ala Ala Leu	
245 250 255	
GAG AAA GCA GGA GGA GGG TCA TTG AAC ATA GTT GTG TCT GAG AGT GGA	816
Glu Lys Ala Gly Gly Gly Ser Leu Asn Ile Val Val Ser Glu Ser Gly	
260 265 270	
TGG CCT TCT TCT GGT GGA ACT GCA ACT TCA CTT GAT AAT GCA AGA ACT	864
Trp Pro Ser Ser Gly Gly Thr Ala Thr Ser Leu Asp Asn Ala Arg Thr	
275 280 285	
TAC AAC ACA AAC TTG GTT CGG AAT GTG AAG CAA GGA ACC CCT AAA AGG	912
Tyr Asn Thr Asn Leu Val Arg Asn Val Lys Gln Gly Thr Pro Lys Arg	
290 295 300	
CCT GGT GCA CCC CTT GAA ACT TAT GTG TTT GCC ATG TTT GAT GAA AAT	960
Pro Gly Ala Pro Leu Glu Thr Tyr Val Phe Ala Met Phe Asp Glu Asn	
305 310 315 320	
CAG AAG CAG CCA GAG TTT GAA AAA TTT TGG GGG CTC TTT TCT CCT ATA	1008
Gln Lys Gln Pro Glu Phe Glu Lys Phe Trp Gly Leu Phe Ser Pro Ile	
325 330 335	
ACT AAG CAG CCC AAA TAC TCG ATT AAT TTC AAT TAA	1044
Thr Lys Gln Pro Lys Tyr Ser Ile Asn Phe Asn	
340 345	

SEQ ID NO: 13

SEQUENCE LENGTH: 331

SEQUENCE TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION:

```

Met Ser Ala Leu Leu Leu Leu Leu Gly Val Leu Ser Ser Thr Gly Val
 1           5           10           15
Leu Leu Thr Gly Val Glu Ser Val Gly Val Cys Tyr Gly Gly Asn Gly
          20           25           30
Asn Asn Leu Pro Thr Lys Gln Ala Val Val Asn Leu Tyr Lys Ser Asn
        35           40           45
Gly Ile Gly Lys Ile Arg Leu Tyr Tyr Pro Asp Glu Gly Ala Leu Gln
        50           55           60
Ala Leu Arg Gly Ser Asn Ile Glu Val Ile Leu Ala Val Pro Asn Asp
        65           70           75           80
Gln Leu Gln Ser Val Ser Asn Asn Gly Gly Ala Thr Asn Trp Val Asn
          85           90           95
Lys Tyr Val Lys Pro Tyr Ala Gly Asn Val Lys Leu Lys Tyr Ile Ala
          100          105          110
Val Gly Asn Glu Val His Pro Gly Asp Ala Leu Ala Gly Ser Val Leu
          115          120          125
Pro Ala Leu Gln Ser Ile Gln Asn Ala Ile Ser Ala Ala Asn Leu Gln
          130          135          140
Arg Gln Ile Lys Val Ser Thr Ala Ile Asp Thr Thr Leu Leu Gly Asn
        145          150          155          160
Ser Tyr Pro Pro Lys Asp Gly Val Phe Ser Asn Ser Ala Ser Ser Tyr
          165          170          175
Ile Thr Pro Ile Ile Asn Phe Leu Ala Lys Asn Gly Ala Pro Leu Leu
          180          185          190

```

Ala Asn Val Tyr Pro Tyr Phe Ala Tyr Val Asn Asn Gln Gln Asn Ile
 195 200 205
 Gly Leu Asp Tyr Ala Leu Phe Thr Lys Gln Gly Asn Asn Glu Val Gly
 210 215 220
 Tyr Gln Asn Leu Phe Asp Ala Leu Val Asp Ser Leu Tyr Ala Ala Leu
 225 230 235 240
 Glu Lys Val Gly Ala Ser Asn Val Lys Val Val Val Ser Glu Ser Gly
 245 250 255
 Trp Pro Ser Gln Gly Gly Val Gly Ala Thr Val Gln Asn Ala Gly Thr
 260 265 270
 Tyr Tyr Arg Asn Leu Ile Lys His Val Lys Gly Gly Thr Pro Lys Arg
 275 280 285
 Pro Asn Gly Pro Ile Glu Thr Tyr Leu Phe Ala Met Phe Asp Glu Asn
 290 295 300
 Gln Lys Gly Gly Ala Glu Thr Glu Lys His Phe Gly Leu Phe Arg Pro
 305 310 315 320
 Asp Lys Ser Pro Lys Tyr Gln Leu Ser Phe Asn
 325 330

SEQ ID NO: 14

SEQUENCE LENGTH: 993

SEQUENCE TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

SEQUENCE DESCRIPTION:

ATG TCT GCC TTA TTG CTG CTT CTT GGA GTA TTA TCT TCC ACT GGA GTA 48

Met Ser Ala Leu Leu Leu Leu Leu Gly Val Leu Ser Ser Thr Gly Val

1

5

10

15

CTG CTT ACT GGG GTA GAA TCT GTG GGT GTG TGT TAT GGA GGA AAT GGA	96
Leu Leu Thr Gly Val Glu Ser Val Gly Val Cys Tyr Gly Gly Asn Gly	
20 25 30	
AAC AAT CTA CCA ACA AAG CAA GCA GTG GTG AAT CTC TAC AAA TCA AAC	144
Asn Asn Leu Pro Thr Lys Gln Ala Val Val Asn Leu Tyr Lys Ser Asn	
35 40 45	
GGA ATT GGC AAA ATC CGT TTA TAC TAT CCA GAT GAA GGT GCC CTT CAA	192
Gly Ile Gly Lys Ile Arg Leu Tyr Tyr Pro Asp Glu Gly Ala Leu Gln	
50 55 60	
GCC CTC AGA GGT TCA AAC ATA GAA GTG ATA CTT GCT GTT CCT AAT GAT	240
Ala Leu Arg Gly Ser Asn Ile Glu Val Ile Leu Ala Val Pro Asn Asp	
65 70 75 80	
CAA CTT CAA TCT GTC TCC AAC AAT GGA GGT GCA ACA AAT TGG GTC AAC	288
Gln Leu Gln Ser Val Ser Asn Asn Gly Gly Ala Thr Asn Trp Val Asn	
85 90 95	
AAG TAC GTG AAA CCC TAT GCA GGA AAC GTG AAA TTG AAG TAC ATT GCA	336
Lys Tyr Val Lys Pro Tyr Ala Gly Asn Val Lys Leu Lys Tyr Ile Ala	
100 105 110	
GTT GGC AAC GAA GTT CAC CCT GGT GAT GCT CTA GCA GGC TCA GTT CTT	384
Val Gly Asn Glu Val His Pro Gly Asp Ala Leu Ala Gly Ser Val Leu	
115 120 125	
CCA GCA CTT CAA AGC ATT CAG AAC GCA ATT TCT GCA GCA AAT TTG CAA	432
Pro Ala Leu Gln Ser Ile Gln Asn Ala Ile Ser Ala Ala Asn Leu Gln	
130 135 140	
CGC CAA ATC AAA GTC TCC ACA GCA ATA GAC ACC ACT CTA CTG GGC AAC	480
Arg Gln Ile Lys Val Ser Thr Ala Ile Asp Thr Thr Leu Leu Gly Asn	
145 150 155 160	
TCT TAC CCA CCA AAA GAT GGC GTT TTC AGC AAC AGT GCA AGT TCA TAC	528
Ser Tyr Pro Pro Lys Asp Gly Val Phe Ser Asn Ser Ala Ser Ser Tyr	

	165	170	175	
ATA ACT CCA ATC ATA AAC TTT TTA GCC AAA AAC GGT GCC CCA CTT CTT				576
Ile Thr Pro Ile Ile Asn Phe Leu Ala Lys Asn Gly Ala Pro Leu Leu				
	180	185	190	
GCA AAC GTG TAC CCT TAC TTC GCC TAC GTT AAC AAT CAA CAA AAC ATT				624
Ala Asn Val Tyr Pro Tyr Phe Ala Tyr Val Asn Asn Gln Gln Asn Ile				
	195	200	205	
GGT CTT GAT TAT GCC TTG TTT ACC AAA CAA GGC AAC AAC GAA GTT GGG				672
Gly Leu Asp Tyr Ala Leu Phe Thr Lys Gln Gly Asn Asn Glu Val Gly				
	210	215	220	
TAC CAA AAC CTG TTT GAT GCA TTG GTG GAT TCT CTG TAC GCA GCT CTT				720
Tyr Gln Asn Leu Phe Asp Ala Leu Val Asp Ser Leu Tyr Ala Ala Leu				
225	230	235	240	
GAG AAA GTG GGA GCA TCA AAT GTG AAG GTT GTT GTG TCT GAG AGT GGG				768
Glu Lys Val Gly Ala Ser Asn Val Lys Val Val Val Ser Glu Ser Gly				
	245	250	255	
TGG CCA TCA CAA GGT GGA GTT GGA GCC ACT GTT CAA AAC GCA GGA ACG				816
Trp Pro Ser Gln Gly Gly Val Gly Ala Thr Val Gln Asn Ala Gly Thr				
	260	265	270	
TAT TAC AGG AAT TTG ATC AAA CAT GTT AAG GGT GGC ACC CCA AAG AGG				864
Tyr Tyr Arg Asn Leu Ile Lys His Val Lys Gly Gly Thr Pro Lys Arg				
	275	280	285	
CCT AAT GGA CCC ATA GAG ACT TAC CTC TTT GCC ATG TTT GAT GAA AAC				912
Pro Asn Gly Pro Ile Glu Thr Tyr Leu Phe Ala Met Phe Asp Glu Asn				
	290	295	300	
CAG AAG GGT GGT GCA GAA ACT GAG AAA CAC TTT GGT CTC TTC AGG CCT				960
Gln Lys Gly Gly Ala Glu Thr Glu Lys His Phe Gly Leu Phe Arg Pro				
305	310	315	320	
GAT AAA TCA CCA AAA TAC CAA CTC AGT TTC AAT				993

Asp Lys Ser Pro Lys Tyr Gln Leu Ser Phe Asn

325

330

SEQ ID NO: 15

SEQUENCE LENGTH: 34

SEQUENCE TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid (synthetic DNA)

SEQUENCE DESCRIPTION:

CTAGAGGATC CGGTACCCCC GGGGTCGACG AGCT

34

SEQ ID NO: 16

SEQUENCE LENGTH: 26

SEQUENCE TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid (synthetic DNA)

SEQUENCE DESCRIPTION:

CGTCGACCCC GGGGGTACCG GATCCT

26

SEQ ID NO: 17

SEQUENCE LENGTH: 44

SEQUENCE TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid (synthetic DNA)

SEQUENCE DESCRIPTION:

CTAGACTTCT TTCCTCAACC TTCTTTCTTC TTATATATTC GAAC

44

SEQ ID NO: 18
SEQUENCE LENGTH: 44
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
GATCGTTCGA ATATATAAGA AGAAAGAAGG TTGAGGAAAG AAGT 44

SEQ ID NO: 19
SEQUENCE LENGTH: 30
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
TGGCATGCCA TGGATTTTCAT CTCATCTCTT 30

SEQ ID NO: 20
SEQUENCE LENGTH: 27
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid (synthetic DNA)
SEQUENCE DESCRIPTION:
GGGAATTCAC TCCGCGAGCC GGCGAAT 27

SEQ ID NO: 21

SEQUENCE LENGTH: 35

SEQUENCE TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid (synthetic DNA)

SEQUENCE DESCRIPTION:

ACAAGTAAAA GAAAGAGCGA GAAATCATCG AAATG

35

SEQ ID NO: 22

SEQUENCE LENGTH: 24

SEQUENCE TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid (synthetic DNA)

SEQUENCE DESCRIPTION:

AGCCATGGCT CCTCCTAAAG TGAT

24

SEQ ID NO: 23

SEQUENCE LENGTH: 21

SEQUENCE TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid (synthetic DNA)

SEQUENCE DESCRIPTION:

GGCCGCGATA TCATGAATGC A

21

SEQ ID NO: 24

SEQUENCE LENGTH: 13

SEQUENCE TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid (synthetic DNA)

SEQUENCE DESCRIPTION:

TTCATGATAT CGC

13

SEQ ID NO: 25

SEQUENCE LENGTH: 45

SEQUENCE TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid (synthetic DNA)

SEQUENCE DESCRIPTION:

GGCCGCCTTC TTCCTCAAC CTTCTTTCTT CTTATATATT CGAAC

45

SEQ ID NO: 26

SEQUENCE LENGTH: 45

SEQUENCE TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid (synthetic DNA)

SEQUENCE DESCRIPTION:

AATTGTTCGA ATATATAAGA AGAAAGAAGG TTGAGGAAAG AAGGC

45

SEQ ID NO: 27

SEQUENCE LENGTH: 204

SEQUENCE TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION:

Glu	Tyr	Asn	Trp	Glu	Lys	Lys	Asp	Ser	Gly	Asp	Leu	Leu	Leu	Leu	Ala
1				5					10					15	
His	Pro	Leu	His	Val	Gln	Leu	Leu	Arg	Asn	Gly	Asp	Asn	Asp	Val	Lys
				20				25						30	
Ile	Leu	Glu	Asp	Leu	Lys	Tyr	Lys	Ser	Ile	Asp	Gly	Asp	Leu	Val	Gly
				35				40						45	
Val	Val	Gly	Asp	Ser	Trp	Val	Leu	Lys	Thr	Asp	Pro	Leu	Phe	Val	Thr
				50				55						60	
Trp	His	Ser	Ile	Lys	Gly	Ile	Lys	Glu	Glu	Ser	His	Asp	Glu	Ile	Val
65					70					75					80
Ser	Ala	Leu	Ser	Lys	Asp	Val	Glu	Ser	Leu	Asp	Ser	Ser	Ser	Ile	Thr
					85					90					95
Thr	Thr	Glu	Ser	Tyr	Phe	Tyr	Gly	Lys	Leu	Ile	Ala	Arg	Ala	Ala	Arg
					100					105					110
Leu	Val	Leu	Ile	Ala	Glu	Glu	Leu	Asn	Tyr	Pro	Asp	Val	Ile	Pro	Lys
					115					120					125
Val	Arg	Asn	Phe	Leu	Lys	Glu	Thr	Ile	Glu	Pro	Trp	Leu	Glu	Gly	Thr
					130										140
Phe	Ser	Gly	Asn	Gly	Phe	Leu	His	Asp	Glu	Lys	Trp	Gly	Gly	Ile	Ile
145						150				155					160
Thr	Gln	Lys	Gly	Ser	Thr	Asp	Ala	Gly	Gly	Asp	Phe	Gly	Phe	Gly	Ile
						165				170					175
Tyr	Asn	Asp	His	His	Tyr	His	Leu	Gly	Tyr	Phe	Ile	Tyr	Gly	Ile	Ala
						180				185					190
Val	Leu	Thr	Lys	Leu	Asp	Pro	Ala	Trp	Gly	Arg	Lys				
						195									200

SEQ ID NO: 28

SEQUENCE LENGTH: 612

SEQUENCE TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

SEQUENCE DESCRIPTION:

GAA TAT AAC TGG GAG AAG AAA GAT TCA GGG GAT TTG CTA CTC TTG GCT	48
Glu Tyr Asn Trp Glu Lys Lys Asp Ser Gly Asp Leu Leu Leu Leu Ala	
1 5 10 15	
CAC CCT CTC CAT GTT CAG CTT CTT CGT AAT GGA GAC AAT GAT GTC AAA	96
His Pro Leu His Val Gln Leu Leu Arg Asn Gly Asp Asn Asp Val Lys	
20 25 30	
ATT CTT GAA GAT TTA AAG TAT AAA AGC ATT GAT GGG GAT CTT GTT GGT	144
Ile Leu Glu Asp Leu Lys Tyr Lys Ser Ile Asp Gly Asp Leu Val Gly	
35 40 45	
GTT GTC GGG GAT TCA TGG GTT TTG AAA ACA GAT CCT TTG TTT GTA ACA	192
Val Val Gly Asp Ser Trp Val Leu Lys Thr Asp Pro Leu Phe Val Thr	
50 55 60	
TGG CAT TCA ATC AAG GGA ATC AAA GAA GAA TCC CAT GAT GAG ATT GTC	240
Trp His Ser Ile Lys Gly Ile Lys Glu Glu Ser His Asp Glu Ile Val	
65 70 75 80	
TCA GCC CTT TCT AAA GAT GTT GAG AGC CTA GAT TCA TCA TCA ATA ACT	288
Ser Ala Leu Ser Lys Asp Val Glu Ser Leu Asp Ser Ser Ser Ile Thr	
85 90 95	
ACA ACA GAG TCA TAT TTT TAT GGG AAG TTG ATT GCA AGG GCT GCA AGG	336
Thr Thr Glu Ser Tyr Phe Tyr Gly Lys Leu Ile Ala Arg Ala Ala Arg	
100 105 110	
TTG GTA TTG ATT GCT GAG GAG TTG AAC TAC CCT GAT GTG ATT CCA AAG	384
Leu Val Leu Ile Ala Glu Glu Leu Asn Tyr Pro Asp Val Ile Pro Lys	

115	120	125	
GTT AGG AAT TTT TTG AAA GAA ACC ATT GAG CCA TGG TTG GAG GGA ACT			432
Val Arg Asn Phe Leu Lys Glu Thr Ile Glu Pro Trp Leu Glu Gly Thr			
130	135	140	
TTT AGT GGG AAT GGA TTC CTA CAT GAT GAA AAA TGG GGT GGC ATT ATT			480
Phe Ser Gly Asn Gly Phe Leu His Asp Glu Lys Trp Gly Gly Ile Ile			
145	150	155	160
ACC CAA AAG GGG TCC ACT GAT GCT GGT GGT GAT TTT GGA TTT GGA ATT			528
Thr Gln Lys Gly Ser Thr Asp Ala Gly Gly Asp Phe Gly Phe Gly Ile			
	165	170	175
TAC AAT GAT CAC CAC TAT CAT TTG GGG TAC TTC ATT TAT GGA ATT GCG			576
Tyr Asn Asp His His Tyr His Leu Gly Tyr Phe Ile Tyr Gly Ile Ala			
	180	185	190
GTG CTC ACT AAG CTT GAT CCA GCA TGG GGT AGG AAG			612
Val Leu Thr Lys Leu Asp Pro Ala Trp Gly Arg Lys			
195	200		

SEQ ID NO: 29

SEQUENCE LENGTH: 1080

SEQUENCE TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

SEQUENCE DESCRIPTION:

ATG GAT TTC ATC TCA TCT CTT ATC GTT GGC TGT GCT CAG GTG TTG TGT	48
Met Asp Phe Ile Ser Ser Leu Ile Val Gly Cys Ala Gln Val Leu Cys	
1 5 10 15	
GAA TCT ATG AAT ATG GCG GAG AGA AGA GGA CAT AAG ACT GAT CTT AGA	96
Glu Ser Met Asn Met Ala Glu Arg Arg Gly His Lys Thr Asp Leu Arg	

20	25	30	
CAA GCC ATC ACT GAT CTT GAA ACA GCC ATC GGT GAC TTG AAG GCC ATA			144
Gln Ala Ile Thr Asp Leu Glu Thr Ala Ile Gly Asp Leu Lys Ala Ile			
35	40	45	
CGT GAT GAC CTG ACT TTA CGG ATC CAA CAA GAC GGT CTA GAG GGA CGA			192
Arg Asp Asp Leu Thr Leu Arg Ile Gln Gln Asp Gly Leu Glu Gly Arg			
50	55	60	
AGC TGC TCA AAT CGT GCC AGA GAG TGG CTT AGT GCG GTG CAA GTA ACG			240
Ser Cys Ser Asn Arg Ala Arg Glu Trp Leu Ser Ala Val Gln Val Thr			
65	70	75	80
GAG ACT AAA ACA GCC CTA CTT TTA GTG AGG TTT AGG CGT CGG GAA CAG			288
Glu Thr Lys Thr Ala Leu Leu Leu Val Arg Phe Arg Arg Arg Glu Gln			
85	90	95	
AGG ACG CGA ATG AGG AGG AGA TAC CTC AGT TGT TTC GGT TGT GCC GAC			336
Arg Thr Arg Met Arg Arg Arg Tyr Leu Ser Cys Phe Gly Cys Ala Asp			
100	105	110	
TAC AAA CTG TGC AAG AAG GTT TCT GCC ATA TTG AAG AGC ATT GGT GAG			384
Tyr Lys Leu Cys Lys Lys Val Ser Ala Ile Leu Lys Ser Ile Gly Glu			
115	120	125	
CTG AGA GAA CGC TCT GAA GCT ATC AAA ACA GAT GGC GGG TCA ATT CAA			432
Leu Arg Glu Arg Ser Glu Ala Ile Lys Thr Asp Gly Gly Ser Ile Gln			
130	135	140	
GTA ACT TGT AGA GAG ATA CCC ATC AAG TCC GTT GTC GGA AAT ACC ACG			480
Val Thr Cys Arg Glu Ile Pro Ile Lys Ser Val Val Gly Asn Thr Thr			
145	150	155	160
ATG ATG GAA CAG GTT TTG GAA TTT CTC AGT GAA GAA GAA GAA AGA GGA			528
Met Met Glu Gln Val Leu Glu Phe Leu Ser Glu Glu Glu Glu Arg Gly			
165	170	175	
ATC ATT GGT GTT TAT GGA CCT GGT GGG GTT GGG AAG ACA ACG TTA ATG			576

Ile	Ile	Gly	Val	Tyr	Gly	Pro	Gly	Gly	Val	Gly	Lys	Thr	Thr	Leu	Met	
			180					185						190		
CAG	AGC	ATT	AAC	AAC	GAG	CTG	ATC	ACA	AAA	GGA	CAT	CAG	TAT	GAT	GTA	624
Gln	Ser	Ile	Asn	Asn	Glu	Leu	Ile	Thr	Lys	Gly	His	Gln	Tyr	Asp	Val	
			195					200						205		
CTG	ATT	TGG	GTT	CAA	ATG	TCC	AGA	GAA	TTC	GGC	GAG	TGT	ACA	ATT	CAG	672
Leu	Ile	Trp	Val	Gln	Met	Ser	Arg	Glu	Phe	Gly	Glu	Cys	Thr	Ile	Gln	
			210					215						220		
CAA	GCC	GTT	GGA	GCA	CGG	TTG	GGT	TTA	TCT	TGG	GAC	GAG	AAG	GAG	ACC	720
Gln	Ala	Val	Gly	Ala	Arg	Leu	Gly	Leu	Ser	Trp	Asp	Glu	Lys	Glu	Thr	
225						230					235				240	
GGC	GAA	AAC	AGA	GCT	TTG	AAG	ATA	TAC	AGA	GCT	TTG	AGA	CAG	AAA	CGT	768
Gly	Glu	Asn	Arg	Ala	Leu	Lys	Ile	Tyr	Arg	Ala	Leu	Arg	Gln	Lys	Arg	
				245					250					255		
TTC	TTG	TTG	TTG	CTA	GAT	GAT	GTC	TGG	GAA	GAG	ATA	GAC	TTG	GAG	AAA	816
Phe	Leu	Leu	Leu	Leu	Asp	Asp	Val	Trp	Glu	Glu	Ile	Asp	Leu	Glu	Lys	
				260					265					270		
ACT	GGA	GTT	CCT	CGA	CCT	GAC	AGG	GAA	AAC	AAA	TGC	AAG	GTG	ATG	TTC	864
Thr	Gly	Val	Pro	Arg	Pro	Asp	Arg	Glu	Asn	Lys	Cys	Lys	Val	Met	Phe	
				275					280					285		
ACG	ACA	CGG	TCT	ATA	GCA	TTA	TGC	AAC	AAT	ATG	GGT	GCG	GAA	TAC	AAG	912
Thr	Thr	Arg	Ser	Ile	Ala	Leu	Cys	Asn	Asn	Met	Gly	Ala	Glu	Tyr	Lys	
				290					295					300		
TTG	AGA	GTG	GAG	TTT	CTG	GAG	AAG	AAA	CAC	GCG	TGG	GAG	CTG	TTC	TGT	960
Leu	Arg	Val	Glu	Phe	Leu	Glu	Lys	Lys	His	Ala	Trp	Glu	Leu	Phe	Cys	
305						310					315				320	
AGT	AAG	GTA	TGG	AGA	AAA	GAT	CTT	TTA	GAG	TCA	TCA	TCA	ATT	CGC	CGG	1008
Ser	Lys	Val	Trp	Arg	Lys	Asp	Leu	Leu	Glu	Ser	Ser	Ser	Ile	Arg	Arg	
				325							330				335	

CTC GCG GAG ATT ATA GTG AGT AAA TGT GGA GGA TTG CCA CTA GCG TTG 1056

Leu Ala Glu Ile Ile Val Ser Lys Cys Gly Gly Leu Pro Leu Ala Leu

340

345

350

ATC ACT TTA GGA GGA GCC ATG GCT

1080

Ile Thr Leu Gly Gly Ala Met Ala

355

360

CLAIMS

1. A DNA coding for a chimeric protein comprising a domain containing at least an elicitor binding site of an elicitor receptor and a domain containing at least a signal transduction motif of a protein which confers plant disease resistance.
2. The DNA of claim 1, wherein the elicitor is a glucan, polygalacturonate, N-acetylchitosaccharide, elicitin, *Cladosporium fulvum*-derived Avr gene product, or *Rhynchosporium secalis*-derived *nip1* gene product.
3. The DNA of claim 1, wherein the elicitor binding site comprises the following recombinant protein (a) or (b):
 - (a) a protein consisting of the amino acid sequence shown in SEQ ID NO: 27
 - (b) a protein which consists of the amino acid sequence shown in SEQ ID NO: 27 having deletion, substitution, addition or insertion of at least one amino acid, and which has ability to bind to a glucan elicitor.
4. The DNA of any one of claims 1 to 3, wherein the domain containing a signal transduction motif contains at least one motif selected from the group consisting of leucine-rich repeats, leucine zippers, nucleotide binding sites and serine/threonine kinase domains.
5. The DNA of any one of claims 1 to 4, wherein the domain containing a signal transduction motif is a signal transduction

domain of an expression product of tomato-derived *Pto*, *Prf*, *Cf-2* or *Cf-9* gene; a signal transduction domain of an expression product of rice-derived *Xa21* gene; a signal transduction domain of an expression product of *Arabidopsis thaliana*-derived *RPS2* or *RPM1* gene; a signal transduction domain of an expression product of flax-derived *L6* gene; or a signal transduction domain of an expression product of tobacco-derived *N* gene.

6. The DNA of claim 5, wherein *RPS2* gene comprises the nucleotide sequence shown in SEQ ID NO: 29.

7. A DNA coding for the following protein (c) or (d):

(c) a chimeric protein comprising the amino acid sequence shown in SEQ ID NO: 5, 7 or 9;

(d) a chimeric protein which comprises the amino acid sequence shown in SEQ ID NO: 5, 7 or 9 having deletion, substitution, addition or insertion of at least one amino acid, and which confers plant disease resistance.

8. A DNA coding for a chimeric protein which confers plant disease resistance, said DNA comprising the nucleotide sequence shown in SEQ ID NO: 6, 8 or 10.

9. A recombinant vector comprising the DNA of any one of claims 1 to 8.

10. A transformed plant transformed with the recombinant vector of claim 9, or the progeny of said plant.

11. A method for creating a plant disease resistant plant, comprising transferring the recombinant vector of claim 9 into a plant.

12. A plant having plant disease resistance, said plant having the DNA of any one of claims 1 to 8 transferred thereinto and expressing said DNA, or the progeny of said plant.

13. A method for creating a plant having plant disease resistance or the progeny of said plant, comprising incorporating the DNA of any one of claims 1 to 8 into a chromosome of a plant and expressing said DNA.

**Smart & Biggar
Ottawa, Canada
Patent Agents**

FIG. 1

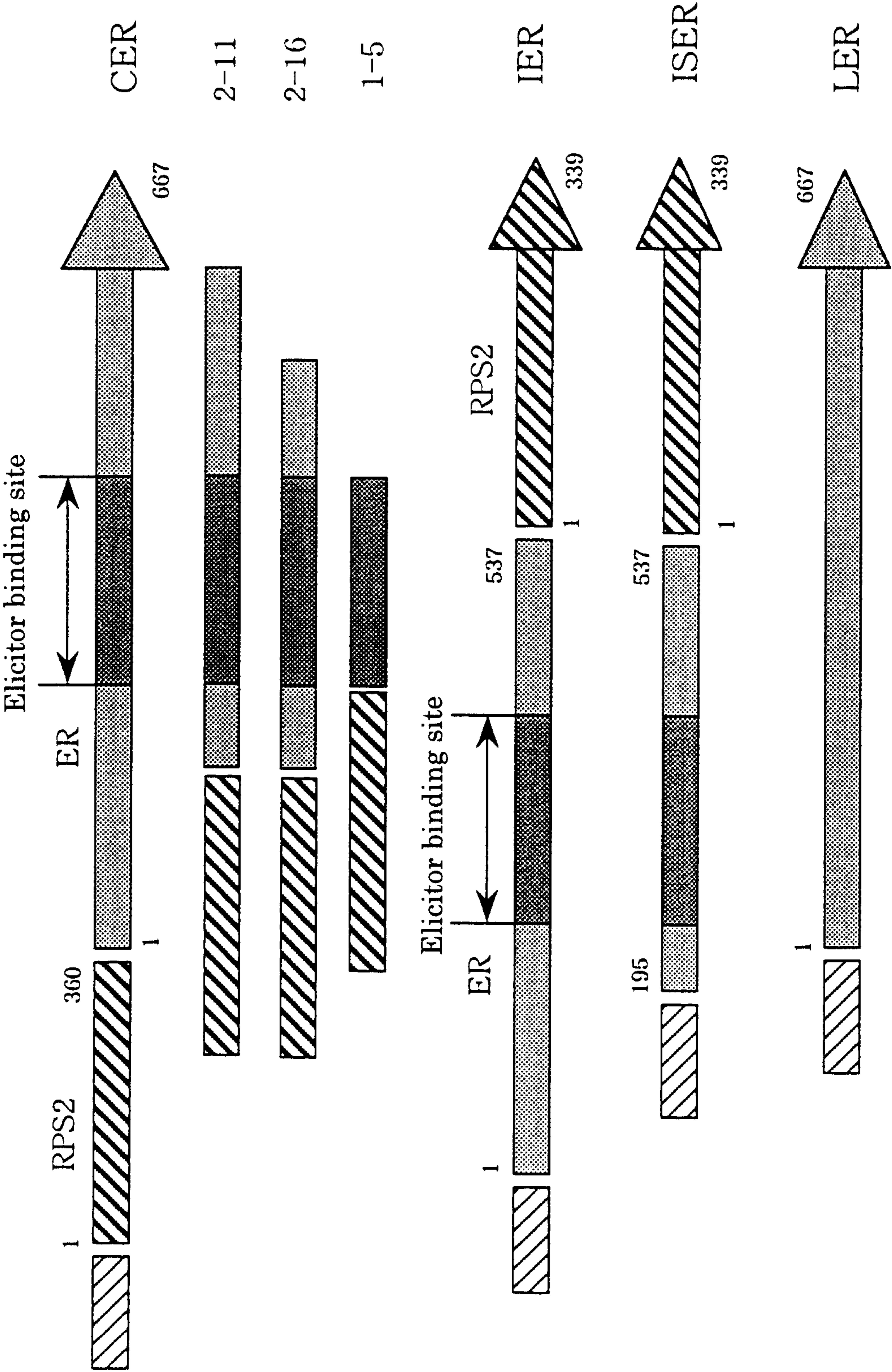


FIG. 2

Construction of pBI(Hm)-ER

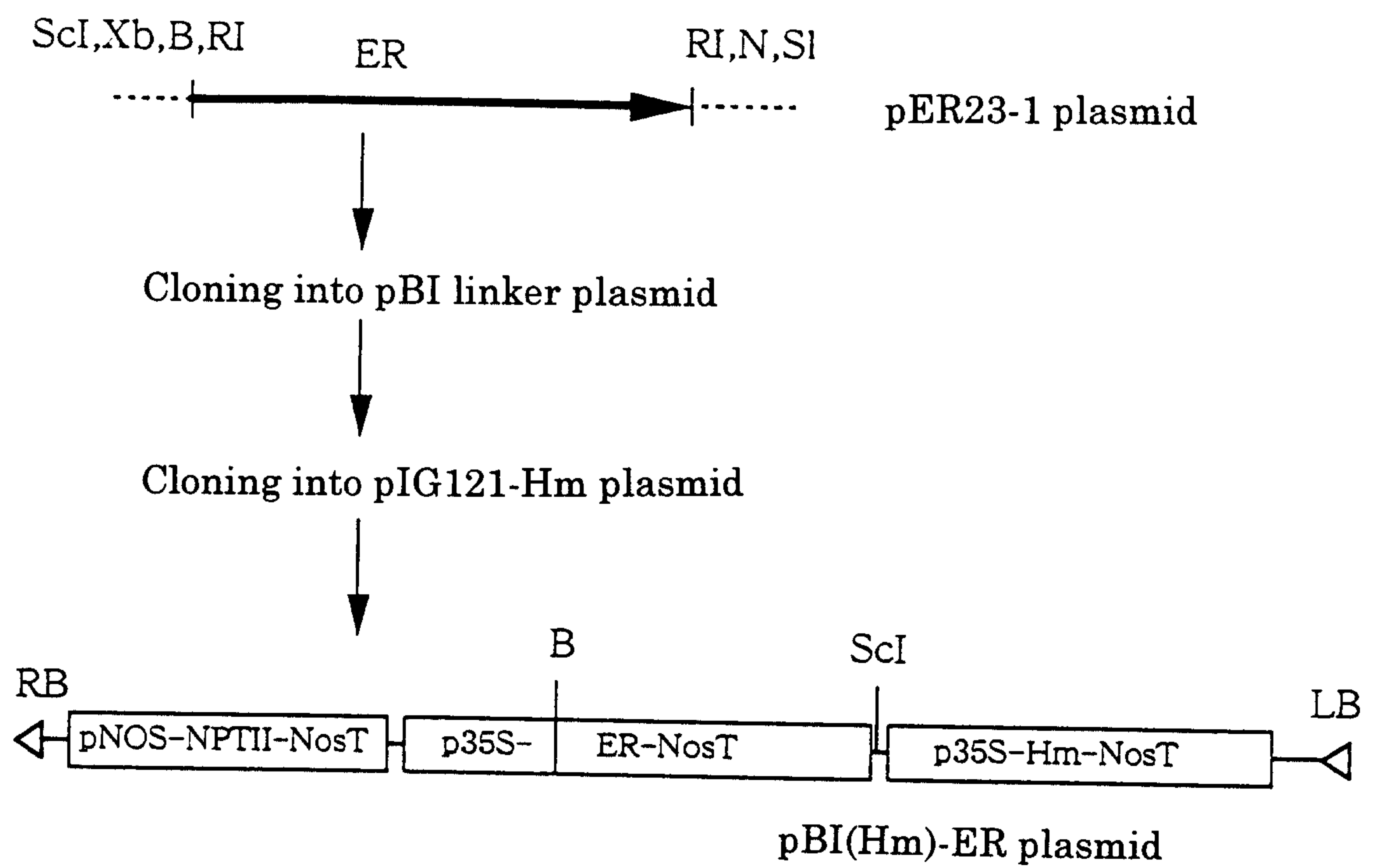


FIG. 3

Construction of pBI(Hm)-LER

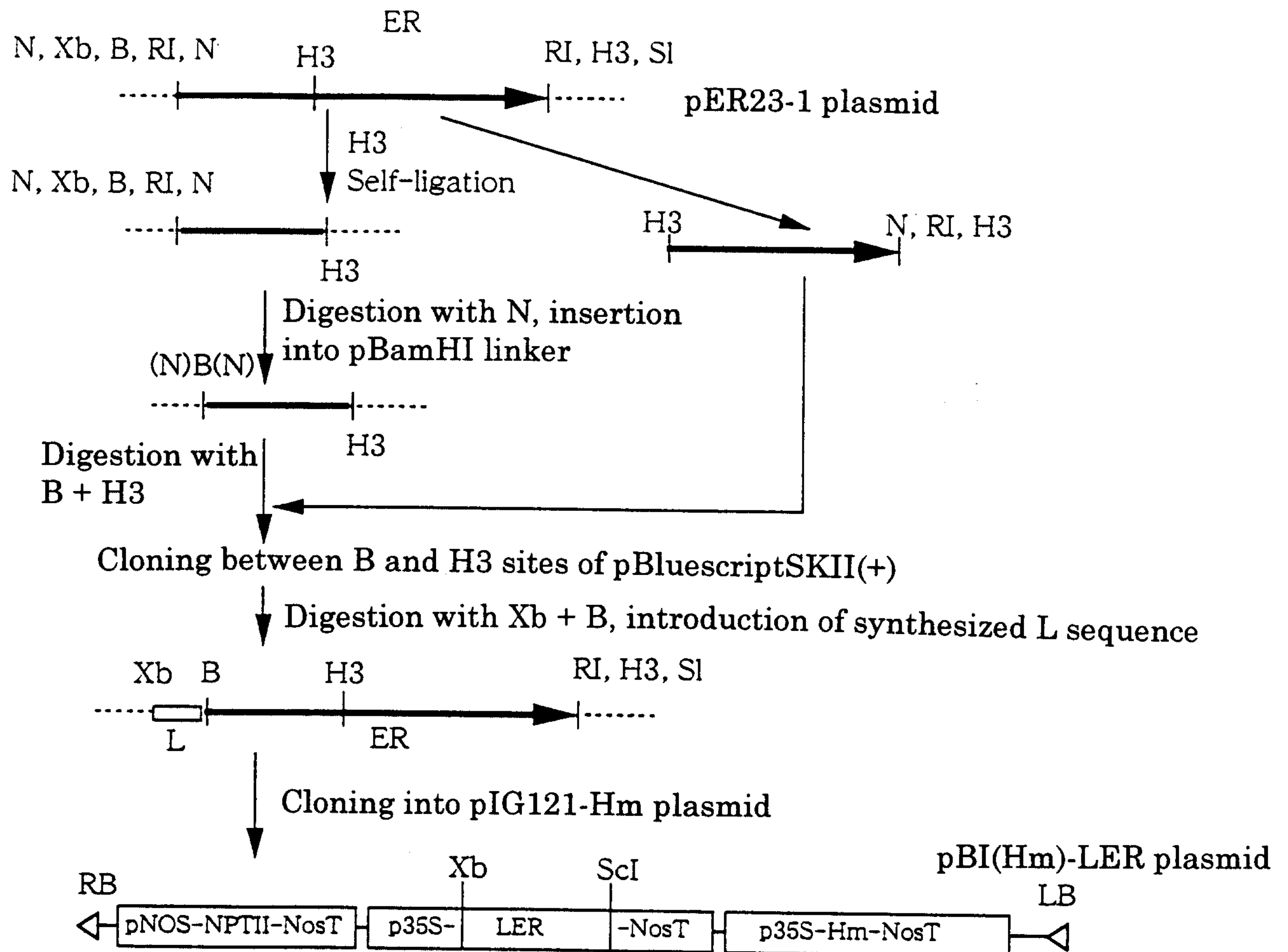


FIG. 4

Plasmids Comprising a DNA Encoding an N-Terminal Region
of RPS2 Corresponding to about 1/3 of This Protein

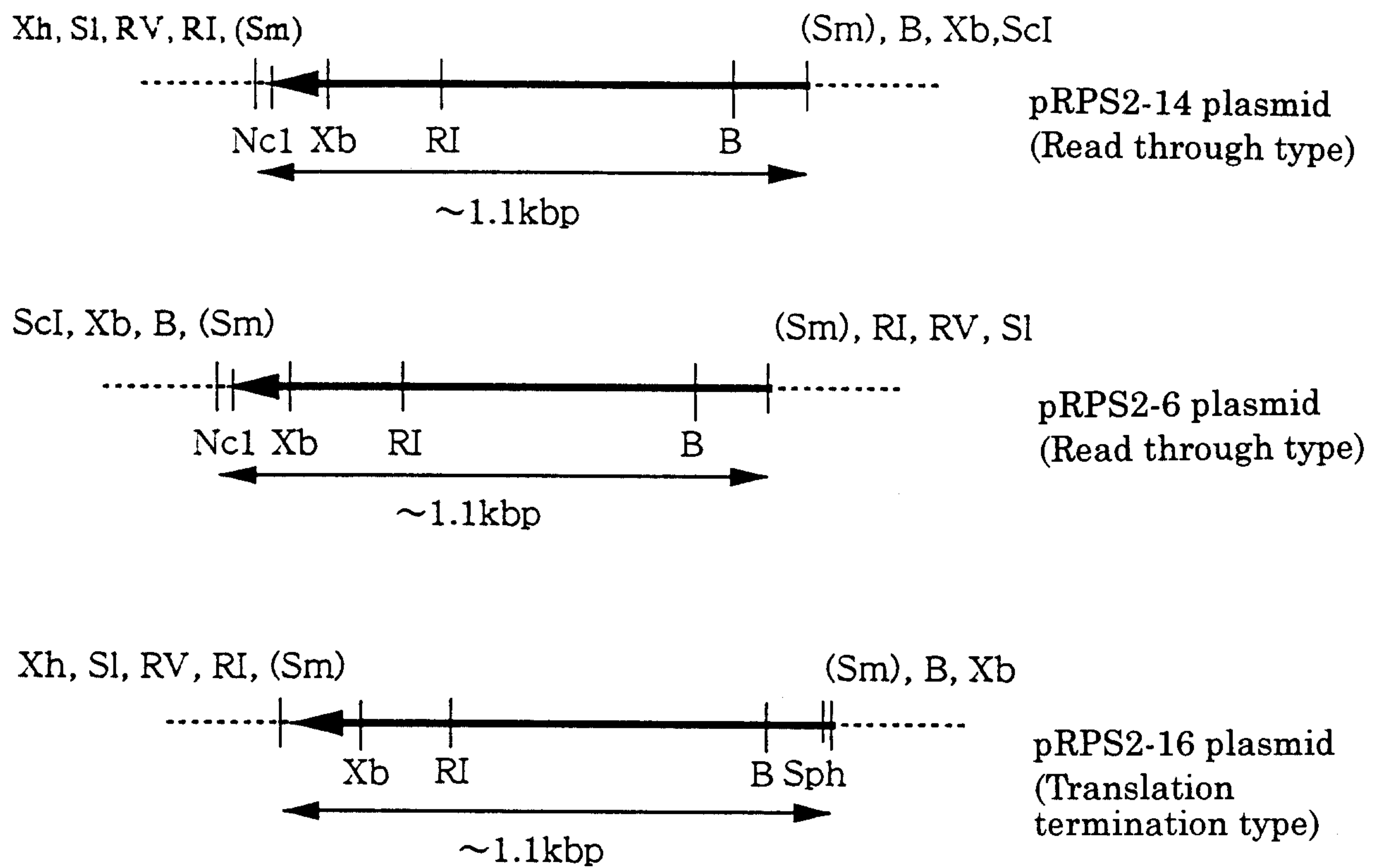


FIG. 5

Construction of pBI(Hm)-CER

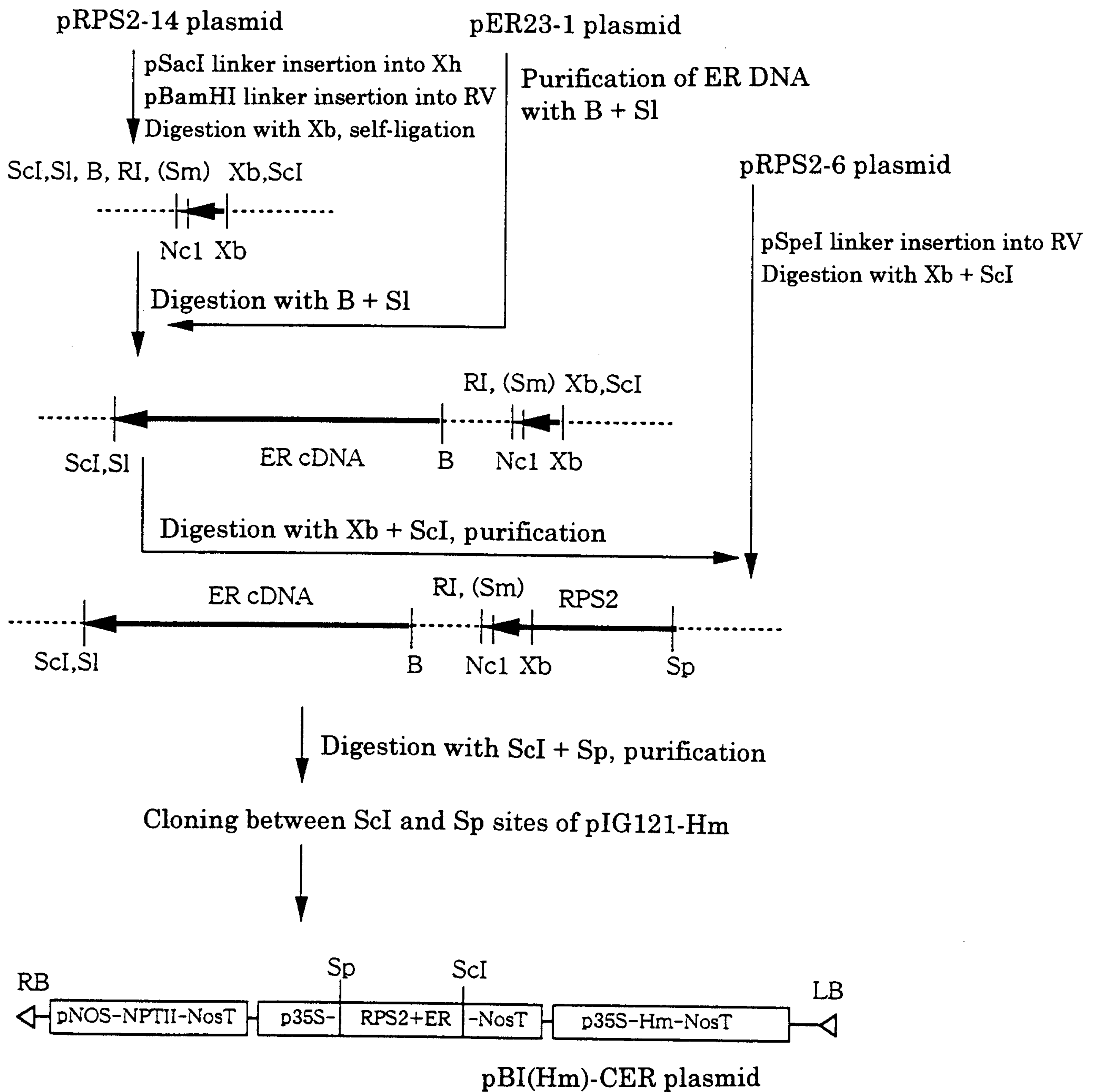


FIG. 6

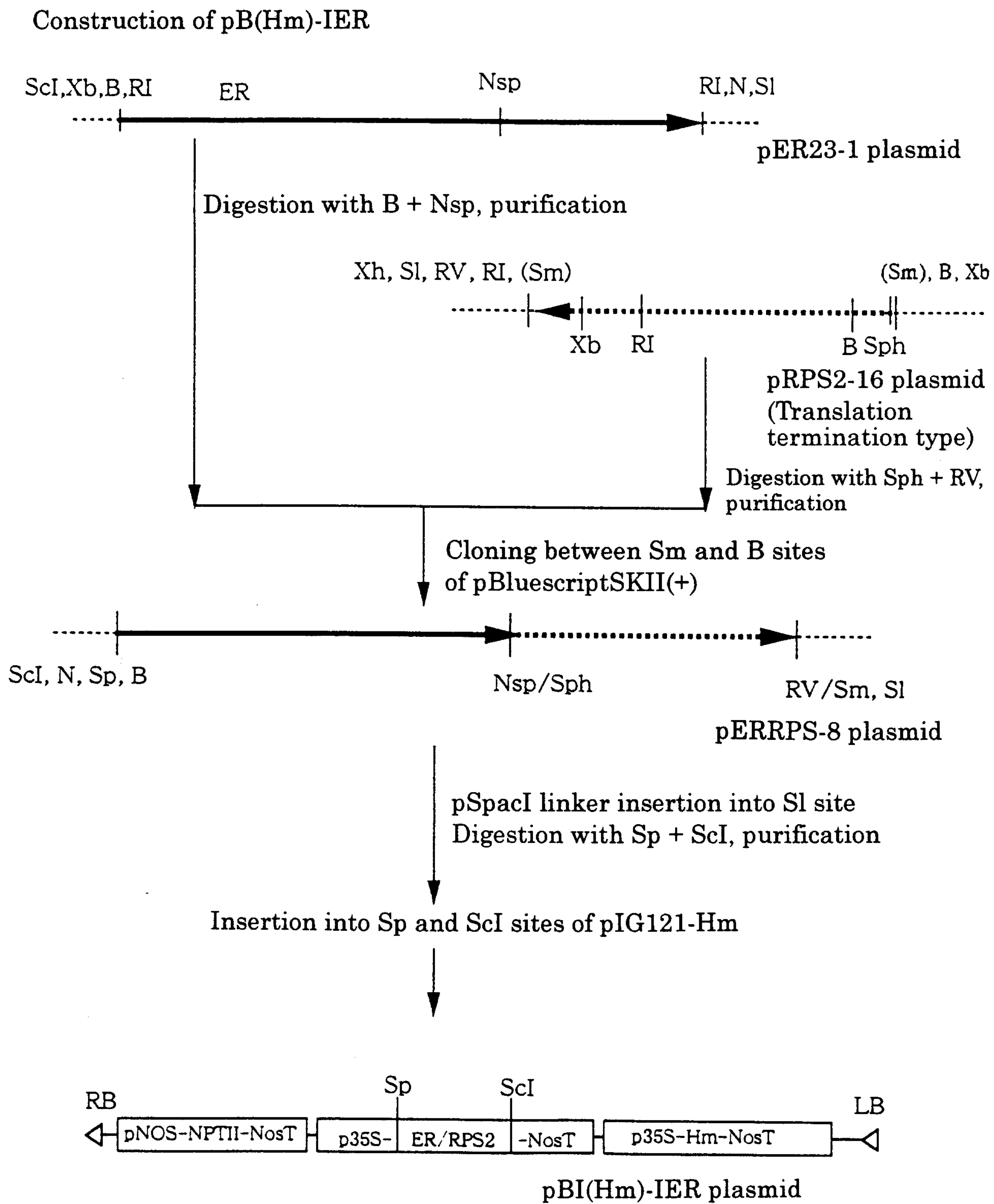


FIG. 7

