



(86) **Date de dépôt PCT/PCT Filing Date:** 2014/06/06
(87) **Date publication PCT/PCT Publication Date:** 2014/12/24
(85) **Entrée phase nationale/National Entry:** 2015/12/17
(86) **N° demande PCT/PCT Application No.:** DE 2014/000310
(87) **N° publication PCT/PCT Publication No.:** 2014/202044
(30) **Priorité/Priority:** 2013/06/17 (DE10 2013 010 026.7)

(51) **Cl.Int./Int.Cl.** **C12N 15/82** (2006.01),
C07K 14/415 (2006.01), **C12N 15/113** (2010.01)

(71) **Demandeur/Applicant:**
KWS SAAT SE, DE

(72) **Inventeurs/Inventors:**
TORJEK, OTTO, DE;
BORCHARDT, DIETRICH, DE;
MECHELKE, WOLFGANG, DE;
LEIN, JENS CHRISTOPH, DE

(74) **Agent:** MARKS & CLERK

(54) **Titre : GENE DE RESISTANCE A LA RHIZOMANIE**
(54) **Title: RESISTANCE GENE AGAINST RHIZOMANIA**

(57) **Abrégé/Abstract:**

The invention relates to a novel nucleic acid molecule, which encodes a polypeptide that can give a plant, in particular a plant of the genus Beta, in which the polypeptide is expressed, resistance to a pathogen, in particular to "Beet Necrotic Yellow Vein Virus", and a preferred nucleic acid molecule that encodes the RZ-3 gene of Beta maritima, derivatives and homologues thereof. The invention further relates to vectors, to transgenic plant cells, to transgenic plants, to methods for the production thereof and to methods for identifying a nucleic acid molecule that provides resistance.



Abstract

The present invention provides a new nucleic acid molecule which encodes a polypeptide that is able to convey a resistance to a pathogen, in particular to “Beet Necrotic Yellow Vein Virus” in a plant, in particular a plant of the Beta genus, in which the polypeptide is expressed, and also a preferred nucleic acid molecule encoding the RZ-3 gene of *Beta maritima*, derivatives and homologues thereof. Further aspects of the invention include vectors, transgenic plant cells, transgenic plants, methods for production thereof, and methods for identifying a resistance-conveying nucleic acid molecule.

- 1 -

RESISTANCE GENE AGAINST RHIZOMANIA

Field of the invention

The present invention relates to a nucleic acid molecule that encodes a polypeptide that is able to convey a resistance to a pathogen, in particular to “Beet Necrotic Yellow Vein Virus” in a plant, in particular of the Beta genus in which the polypeptide is expressed. The invention also relates to a polypeptide that is able to convey a resistance to a pathogen in a plant, in particular a resistance to BNYYV in a plant of the Beta genus, in which the polypeptide is expressed and that is encoded by the nucleic acid molecule according to the invention. The invention also relates to a transgenic plant, plant cell, plant organ, plant tissue, plant part or a seed of a plant that comprise the nucleic acid molecule or parts thereof, and also to methods for producing a transgenic plant or plant cell of this type. The invention also includes methods for detecting the resistance-conveying nucleic acid molecule and methods for the selection of plants or plant cells that have the resistance-conveying nucleic acid molecule.

Background of the invention

Rhizomania is the most serious sugar beet disease worldwide in terms of profitability and may cause earnings losses of 50 % and more. The disease, which is also referred to as “root madness”, is caused by the “Beet Necrotic Yellow Vein Virus” (BNYYV) and is transmitted by the soil-borne protozoa *Polymyxa betae*. A BNYYV infection manifests itself in an increased proliferation of the thin roots and secondary roots and in the formation of a greatly reduced root body with reduced sugar content. Infected plants demonstrate a reduced water uptake and are thus more sensitive to dry stress. When the infection spreads to the entire plant, this results in a yellowing of the leaf veins, in necrotic lesions, and yellow flecks on the leaves. Since a curative combatting of the disease is not possible, as is the case with other viral diseases, damage can be prevented only via the cultivation of resistant species. Three major genes against rhizomania are currently being examined in essence: RZ-1 (also referred to as “Holly”), RZ-2 and RZ-3. In addition, further rhizomania resistance genes are described in the literature, although these are of lesser significance. Here, the resistance gene RZ-1 is already incorporated into most breeding lines (seed parent and/or pollinator parent components). It has been found, however, that the resistance conveyed by RZ-1 is insufficient in heavily infected regions or in regions

- 2 -

having diverse BNYVV pathotypes (for example Sohi & Maleki, 2004). For this reason, it was already proposed some time ago to combine RZ-1 with, for example, RZ-2 or RZ-3. RZ-2 and RZ-3 originate from *Beta vulgaris* subsp. *maritima* sources (WB42, WB41) and genetically map in the same region on chromosome 3 of the sugar beet genome, whereas RZ-1 likewise maps on chromosome 3, but south of RZ-2 and RZ-3. Scholten et al. (1999) determined a distance of 20-25 cM between the RZ major genes RZ-1 and RZ-2. Gidner et al. (2005) found a shorter distance of 5 cM between RZ-1 and RZ-2 and did not conclude that RZ-2 and RZ-3 map on the same locus. Schmidlin et al. (2008) identified differently induced genes by means of expression analysis in infected beets, however these did not correspond to RZ-2 or RZ-3. In the study by Larson et al. (2008), some BNYVV-induced proteins were detected in the sugar beet using the MALDI-TOF-MS method, however the proteins which were encoded by RZ-1, RZ-2 or RZ-3 could not be identified by the scientists. In addition, the sequence region, in particular around this resistance gene, is repetitive, which makes the development of diagnostic markers particularly difficult. Until now, neither high-resolution marker maps nor verified candidate genes were publicly available for the specified rhizomania resistance genes. In addition, the functional background of these resistance genes, i.e. the genetic structure, previously was not fully known.

For sustainable cultivation against rhizomania intended to counteract the risk of resistance-breaking BNYVV isolates, it is necessary to constantly identify new resistance genes and to integrate these into the gene pool of the crop plants such as sugar beet.

Summary of the invention

The present invention has been developed on the basis of the above-described prior art, wherein one object of the present invention is to provide a nucleic acid molecule and/or a polypeptide that is able to convey resistance against rhizomania in a plant. A further object is to provide a transgenic rhizomania-resistant plant and a method for the production thereof. A further object of the present invention is to provide methods for using and developing molecular markers that enable an efficient cultivation against rhizomania and the development of new resistant plant lines.

Embodiments of the present invention that achieve the objects are based on the genetic fine-mapping, identification, isolation and characterisation of a gene originating from the donor *Beta vulgaris* subsp. *maritima* and encoding for a polypeptide or protein able to convey resistance against a pathogen in a plant in which the polypeptide is expressed.

Some of the terms used in this application will first be explained in greater detail hereinafter:

The term “approximately” in conjunction with the specification of a length of a nucleotide sequence means a deviation by ± 200 base pairs, preferably by ± 100 base pairs and particularly preferably by ± 50 base pairs.

A “plant of the Beta genus” belongs to the foxtail family (Amaranthaceae). These plants include plants of the species *Beta macrocarpa*, *Beta vulgaris*, *Beta lomatogona*, *Beta macrorhiza*, *Beta corolliflora*, *Beta trigyna* and *Beta nana*. A plant of the species *Beta vulgaris* is in particular a plant of the sub-species *Beta vulgaris* subsp. *maritima* (Seemangold) or *Beta vulgaris* subsp. *vulgaris*. These include, for example, *Beta vulgaris* subsp. *vulgaris* var. *altissima* (sugar beet in the narrower sense), *Beta vulgaris* ssp. *vulgaris* var. *vulgaris* (Mangold), *Beta vulgaris* ssp. *vulgaris* var. *conditiva* (beetroot), *Beta vulgaris* ssp. *vulgaris* var. *crassa/alba* (fodder beet).

The term “hybridise” or “hybridisation” is understood to mean a process in which a single-strand nucleic acid molecule attaches to a nucleic acid strand that is complementary to the greatest possible extent, i.e. forms base pairs. Standard methods for hybridisation are described for example in Sambrook et al. 2001. This is preferably understood to mean that at least 60%, more preferably at least 65%, 70%, 75%, 80% or 85%, particularly preferably 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% of the bases of the nucleic acid molecule form a base pairing with the nucleic acid strand that is complementary to the greatest possible extent. The possibility of such an annealing is dependent on the stringency of the hybridisation conditions. The term “stringency” relates to the hybridisation conditions. High stringency is then given when a base pairing is hindered, and low stringency is given when a base pairing is facilitated. The stringency of the hybridisation conditions is dependent for example on the salt concentration or ion strength and temperature. Generally, the stringency can be increased by an increase of the temperature and/or a lowering of the salt content. “Stringent hybridisation conditions” are understood to mean conditions under which a hybridisation takes place predominantly only between homologous nucleic acid molecules. The term “hybridisation conditions” here does not relate only to the conditions prevailing during the actual attachment of the nucleic acids, but also to the conditions prevailing during the subsequent washing steps. Stringent hybridisation conditions are, for example, conditions under which predominantly only those nucleic acid molecules hybridise that have at least 70%, preferably at least 75%, at least 80%, at least 85% or at least 90%, particularly preferably at least 95%, at least 96%, at least 97%, at least 98% or at least 99% sequence identity. Stringent hybridisation conditions for example are: hybridisation in 4 x SSC at 65 °C and subsequent multiple washing in 0.1 x SSC at 65 °C for a total of

- 4 -

approximately 1 hour. The term “stringent hybridisation conditions” used here may also mean: hybridisation at 68 °C in 0.25 M sodium phosphate, pH 7.2, 7 % SDS, 1 mM EDTA and 1 % BSA for 16 hours and subsequent washing twice with 2 x SSC and 0.1 % SDS at 68 °C. Hybridisation preferably takes place under stringent conditions.

An “isolated nucleic acid molecule” is understood to mean a nucleic acid molecule dissolved out from its natural or original environment. The term also includes a synthetically produced nucleic acid molecule. An “isolated polypeptide” is understood to mean a polypeptide dissolved out from its natural or original environment. The term also includes a synthetically produced polypeptide.

A “molecular marker” is a nucleic acid that is polymorphous in a plant population. Such a marker is thus able to detect and differentiate different allelic states (alleles). Known analytical methods used for this purpose are RFLP, AFLP, SNP, SSR or KASP, for example. The term “molecular marker” relates to nucleotide sequences that are complementary or at least complementary to the greatest possible extent or homologous to genomic sequences, for example nucleic acids used as probes or primers. Markers that describe polymorphisms can be detected with use of well-established methods. These include, for example, PCR-based sequence-specific amplification, a detection of ‘restriction fragment length polymorphisms’ (RFLPs), a detection of polynucleotide polymorphisms by means of ‘allele specific hybridisation’ (ASH), a detection of amplified variable sequences of the plant genome, a detection of a ‘self-sustained sequence replication’, a detection of ‘simple sequence repeats’ (SSRs), a detection of ‘single nucleotide polymorphisms’ (SNPs), or a detection of ‘amplified fragment length polymorphisms’ (AFLPs). Furthermore, the methods for detection of ‘expressed sequence tags’ (ESTs) and SSR markers derived from EST sequences and ‘randomly amplified polymorphic DNA’ (RAPD) are also known.

A “promoter” means a non-translated regulatory DNA sequence, typically upstream of an encoding region, which contains the binding point for the RNA polymerase and initiates the transcription of the DNA.

A “pathogen” means an organism that in interaction with a plant leads to disease symptoms at one or more organs of the plant. These pathogens include, for example, animal, fungal, bacterial or viral organisms or oomycetes.

A “pathogen infection” is understood to mean the earliest moment at which a pathogen interacts with a plant host tissue. By way of example in the case of the viral pathogen BNYVV, this is transmitted by the protozoa *Polymyxa betae*. *Polymyxa* forms spores that can survive in the ground for many decades. The virus also lives in these spores. When these dormant spores

- 5 -

germinate to form mobile zoospores, the virus can pass via these spores into cells of the plant host tissue and can interact there with the host (Esser 2000).

Plant "organs" for example mean leaves, shoot axis, stem, roots, hypocotyl, vegetative buds, meristems, embryos, anthers, ovules or fruits. Plant "parts" mean a combination of a number of organs, for example a flower or a seed, or part of an organ, for example a cross section through the stem. Plant "tissues" are, for example, callus tissue, storage tissue, meristematic tissue, leaf tissue, stem tissue, root tissue, plant tumour tissue or reproductive tissue. Plant "cells" for example are to be understood to mean isolated cells with a cell wall or aggregates thereof or protoplasts.

The term "resistance" is to be understood broadly and covers the scope of protection from a delay to complete inhibition of the development of the disease. An example of a pathogen of significance is Beet Necrotic Yellow Vein Virus (BNYVV). A resistant plant cell of the invention or resistant plant of the invention preferably achieves a resistance to BNYVV. A resistance to a pathogen is to be equated with a resistance to the disease caused by this pathogen, for example a resistance to BNYVV and a resistance to rhizomania.

"Transgenic plant" relates here to a plant in the genome of which at least one nucleic acid has been integrated. This may be a heterologous nucleic acid here. The nucleic acid is preferably integrated in a stable manner, which means that the integrated nucleic acid is retained in the plant in a stable manner, can be expressed, and also can be passed on in a stable manner to the descendant.

The present invention discloses a nucleic acid molecule which includes a polypeptide that is able to convey a resistance to a pathogen in a plant in which the polypeptide is expressed. The nucleic acid molecule comprises a nucleotide sequence selected from

- a) a nucleotide sequence that encodes a polypeptide having an amino acid sequence according to SEQ ID NO: 2 or SEQ ID NO: 3,
- b) a nucleotide sequence comprising the encoding sequence of the DNA sequence according to SEQ ID NO: 1,
- c) a nucleotide sequence that hybridises with the complementary sequence of nucleotide sequence according to a) or b) under stringent conditions,

- 6 -

- d) a nucleotide sequence that encodes a polypeptide which is derived by substitution, deletion and/or addition of one or more amino acids of the amino acid sequence encoded by the nucleotide sequence according to a) or b) from a polypeptide encoded by the nucleotide sequence according to a) or b),
- e) a nucleotide sequence that encodes a polypeptide which has an amino acid sequence at least 60% identical to an amino acid sequence encoded by the nucleotide sequence according to a) or b), or
- f) a nucleotide sequence that encodes at least one nucleotide-binding domain (NBS or NB-ARC) corresponding to amino acid positions 168-227 of SEQ ID NO: 2 or corresponding to amino acid positions 182-241 of SEQ ID NO: 3, at least one leucine-rich domain (LRR) corresponding to amino acid positions 591-613 of SEQ ID NO: 2 or corresponding to amino acid positions 605-627 of SEQ ID NO: 3 and/or at least one internal repetitive domain (IR) corresponding to amino acid positions 1013-1072 of SEQ ID NO: 2 or corresponding to amino acid positions 1027-1086 of SEQ ID NO: 3.

The nucleic acid molecule may be an isolated nucleic acid molecule. It is preferably DNA, and particularly preferably cDNA (encoding DNA). The polypeptide encoded by the nucleic acid molecule according to the invention preferably conveys a resistance to the viral pathogen "Beet Necrotic Yellow Vein Virus" (BNYVV), which causes the plant disease rhizomania. Furthermore, the polypeptide encoded by the nucleic acid molecule according to the invention, in particular a plant of the *Beta* genus, conveys a resistance to a pathogen. The plant is preferably a plant of the *Beta vulgaris* species, particularly preferably a plant of the sub-species *Beta vulgaris* subsp. *maritima* or *Beta vulgaris* subsp. *vulgaris*; these include, for example, the crop types constituted by sugar beet, beetroot, fodder beet, leaf beet, Swiss chard.

In one embodiment of the nucleic acid molecule according to the invention the nucleic acid molecule comprises the nucleotide sequence according to a). The amino acid sequence according to SEQ ID NO: 2 of the encoded polypeptide and/or according to SEQ ID NO: 3 of the encoded polypeptide constitutes the resistance protein of the RZ-3 gene. Here, this is a resistance gene protein of the NBS-LRR type, which is characterised by a certain structural motif. The general structure of such resistance proteins in plants has already been well examined (Martin *et al.* 2003). However, the principle of the structural formation in particular of what is known as the LRR domain, which is a potential identification domain for mostly unknown pathogenic effectors, is not foreseeable. Consequently, the identification of a BNYVV-resistance-conveying gene or protein on the basis purely of the known structural motifs is not

- 7 -

possible. The identification of the RZ-3 resistance gene took place over the course of a map-based cloning process, which required intensive genetic mapping and fine-mapping of the target region in which the RZ-3 resistance gene was initially suspected. The development work will be described in greater detail further below.

The identified resistance protein belongs to the NBS-LRR type and has a nucleotide-binding domain (NBS, also known as NB-ARC) (nucleotide-binding adaptor shared by APAF-1, R proteins, and CED-4)) corresponding to amino acid positions 168-227 of SEQ ID NO: 2 or corresponding to amino acid positions 182-241 of SEQ ID NO: 3, a leucine-rich domain (LRR) corresponding to amino acid positions 591-613 of SEQ ID NO: 2 or corresponding to amino acid positions 605-627 of SEQ ID NO: 3 and/or at least one internal repetitive domain (IR; internal repeat domain) corresponding to amino acid positions 1013-1072 of SEQ ID NO: 2 or corresponding to amino acid positions 1027-1086 of SEQ ID NO: 3. The NBS domain is encoded by nucleotides 2019-2882 of SEQ ID NO: 1, the LRR domain is encoded by nucleotides 3288-3356 of SEQ ID NO: 1 and the IR domain is encoded by nucleotides 4554-4871 of SEQ ID NO: 1. The NB-ARC domain is a central nucleotide-binding domain. It is likely a functional ATPase domain, which expectedly regulates the activity of a resistance protein. The NB-ARC domain consists of three sub-domains: NB, ARC1 und ARC2. Characteristic motifs of the NB-ARC domains are APAF-1 (apoptotic protease-activating factor-1), which is supposedly responsible for the hypersensitive reaction, hhGRExE, Walker-A- or P-loop, Walker-B, GxP, RNBS-A to D and MHD (Ooijen et al., 2008). Some of the specified motifs can already be identified. In a further embodiment of the nucleic acid molecule according to the invention the nucleic acid molecule comprises the nucleotide sequence according to b). The nucleotide sequence comprises the encoding sequences of the DNA sequence according to SEQ ID NO: 1, which encode for the amino acid sequences according to SEQ ID NO: 2 and 3.

In a further embodiment of the nucleic acid molecule the nucleic acid molecule comprises the nucleotide sequence according to d). This nucleotide sequence encodes a polypeptide constituting a derivative of the polypeptide encoded by the nucleotide sequence according to a) or b). A derivative of the polypeptide constitutes a derived amino acid sequence, which has at least one substitution, deletion or addition of one or more amino acids, wherein the functionality of the encoded polypeptide/protein is maintained. In the case of the substitution of an amino acid by another amino acid having the same or equivalent or similar chemical-physical properties, reference is made to a "conservative exchange" or "semi-conservative exchange". Examples of physical-chemical properties of an amino acid are, for example, the hydrophobicity or the charge. It is known to a person skilled in the art which amino acid substitution constitutes a conservative or semi-conservative exchange. The common general knowledge in the art

- 8 -

additionally allows a person skilled in the art to recognise, identify and detect the amino acid deletions and additions harmful to the functionality of the resistance protein RZ-3 and also those positions at which these are possible. It is known to a person skilled in the art that in the case of the present NBS-LRR protein for modifications of the amino acid sequence (substitutions, deletions or additions of one or more amino acids), the functionality of the above-defined preserved domains must be retained in particular and therefore only limited modifications of the above-mentioned type are possible in these domains. The nucleotide sequence of this embodiment then encodes for a derivative or for a derived amino acid sequence when the nucleotide sequence is homologous or identical at least to an extent of 60%, 65%, 70%, 75%, 80%, 85%, 90%, 92%, 94%, 96%, 97%, 98%, 99% to the nucleotide sequence according to a) or b). Such nucleotide sequences, which encode for a derivative or for a derived amino acid sequence, preferably can be produced either directly or indirectly (for example via amplification or replication steps) form a starting nucleotide sequence corresponding over the entire length or at least partially to SEQ ID NO: 1 or another sequence disclosed here.

In a further embodiment of the nucleic acid molecule according to the invention the nucleic acid molecule comprises the nucleotide sequence according to e). This nucleotide sequence codes a polypeptide which has an amino acid sequence identical to an extent of at least 60%, 65%, 70%, 75%, 80%, 85%, 90%, 92%, 94%, 96%, 97%, 98%, 99% to an amino acid sequence encoded by the nucleotide sequence according to a) or b).

In a further embodiment of the nucleic acid molecule according to the invention, the nucleic acid molecule comprises the nucleotide sequence according to f). The nucleotide sequence here encodes at least one nucleotide-binding domain (NBS) corresponding to amino acid positions 168-227 of SEQ ID NO: 2 or corresponding to amino acid positions 182-241 of SEQ ID NO: 3, at least one leucine-rich domain (LRR) corresponding to amino acid positions 591-613 of SEQ ID NO: 2 or corresponding to amino acid positions 605-627 of SEQ ID NO: 3 and/or at least one internal repetitive domain (IR) corresponding to amino acid positions 1013-1072 of SEQ ID NO: 2 or corresponding to amino acid positions 1027-1086 of SEQ ID NO: 3. The nucleotide sequence preferably encodes for a polypeptide comprising at least one nucleotide-binding domain (NBS) corresponding to amino acid positions 168-227 of SEQ ID NO: 2 or corresponding to amino acid positions 182-241 of SEQ ID NO: 3, at least one leucine-rich domain (LRR) corresponding to amino acid positions 591-613 of SEQ ID NO: 2 or corresponding to amino acid positions 605-627 of SEQ ID NO: 3 and at least one internal repetitive domain (IR) corresponding to amino acid positions 1013-1072 of SEQ ID NO: 2 or corresponding to amino acid positions 1027-1086 of SEQ ID NO: 3. These domains are particularly preferably arranged in the polypeptide sequentially from the N- to the C-terminus in

- 9 -

the order NBS-LRR-IR, wherein in each case one or more further amino acids may be present between domains.

The present invention also relates to a polypeptide that is able to convey a resistance to a pathogen in a plant in which the polypeptide is expressed and that is encoded by the nucleic acid molecule according to the invention, wherein the pathogen is preferably BNYVV and/or the plant is preferably a plant of the Beta genus, in particular a plant of the species *Beta vulgaris*. The polypeptide particularly preferably has an amino acid sequence according to SEQ ID NO: 2 or according to SEQ ID NO: 3. The polypeptide may be an isolated polypeptide.

In a further aspect the present invention relates to a vector comprising the nucleic acid molecule according to the invention. The vector may be a plasmid, a cosmid, a phage or expression vector, a transformation vector, shuttle vector or cloning vector, it may be double-stranded or single-stranded, linear or circular, or may transform a prokaryotic or eukaryotic host either by integration into the genome thereof or extrachromosomally. The nucleic acid molecule according to the invention is preferably operatively linked in an expression vector to one or more regulatory sequences allowing the transcription and optionally the expression in a prokaryotic or eukaryotic host cell. By way of example, the nucleic acid molecule is under the control of a suitable promoter or a terminator. Suitable promoters may be promoters that are constitutively induced (Ex.: 35S promoter from the "Cauliflower mosaic virus" (Odell et al. 1985), and promoters that are pathogen-inducible are particularly suitable (Ex.: PR1 promoter from *Petersilie* (Rushton et al., 1996). Particularly suitable pathogen-inducible promoters are synthetic or chimeric promoters, which are not present in nature, are formed from a number of elements, and contain a minimal promoter and also have, upstream of the minimal promoter, at least one cis-regulatory element serving as binding point for special transcription factors. Chimeric promoters are designed in accordance with the desired requirements and are induced or repressed by different factors. Examples of such promoters can be found in WO 00/29592, WO 2007/147395 and WO 2013/091612. A suitable terminator is for example the nos terminator (Depicker et al., 1982).

In addition to the above-described vectors, the present invention also provides a method comprising the introduction of a described vector into a host cell. The vector may be introduced for example by conjugation, mobilisation, biolistic transformation, agrobacterium-conveyed transformation, transfection, transduction, vacuum infiltration or electroporation. A person skilled in the art is familiar with such methods and also methods for the preparation of described vectors (Sambrook et al. 2001).

- 10 -

In a further aspect the present invention relates to a host cell comprising the nucleic acid molecule according to the invention or the vector of the present invention. A host cell in the sense of the invention may be a prokaryotic (for example bacterial) or eukaryotic cell (for example a plant cell or a yeast cell). The host cell is preferably an agrobacterium such as *Agrobacterium tumefaciens* or *Agrobacterium rhizogenes* or a plant cell comprising the nucleic acid molecule according to the invention or the vector of the present invention. Numerous methods such as conjugation or electroporation are known to a person skilled in the art, by means of which said person can introduce the nucleic acid molecule according to the invention or the vector of the present invention into an agrobacterium, and methods such as diverse transformation methods (biolistic transformation, agrobacterium-mediated transformation) are also known to such a person, by means of which said person can introduce the nucleic acid molecule according to the invention or the vector of the present invention into a plant cell (Sambrook et al. 2001).

In a further aspect the present invention relates to a transgenic plant cell comprising the nucleic acid molecule according to the invention as transgene or comprising the vector of the present invention. A transgenic plant cell of this type by way of example is a plant cell that is transformed, preferably in a stable manner, with the nucleic acid molecule according to the invention or with the vector of the present invention. In a preferred embodiment of the transgenic plant cell, the nucleic acid molecule is operatively linked to one or more regulatory sequences allowing the transcription and optionally the expression in the plant cell. The overall construct from the nucleic acid molecule according to the invention and the regulatory sequence(s) then constitutes the transgene. Such regulatory sequences by way of example are a promoter or a terminator. Numerous functional promoters and terminators applicable in plants are known to a person skilled in the art. A transgenic plant cell of the present invention, in particular a cell of a plant of the Beta genus, preferably demonstrates a higher resistance to a pathogen, in particular BNYVV, than a corresponding non-transformed plant cell (the plant cell without the transgene). The level of the resistance by way of example to BNYVV can be defined qualitatively in plants of the Beta genus by determination of rating scores (rating score schemes for plants of the Beta genus are known from the prior art, for example for sugar beet Mechelke (1997)). A higher resistance presents itself in an improvement of the resistance by at least one rating score, by at least two rating scores, or by at least three or more rating scores. The present invention furthermore also relates to a method for producing a transgenic plant cell of the present invention comprising a step of introducing the nucleic acid molecule according to the invention or the vector of the present invention into a plant cell. By way of example, the introduction may take place by transformation, preferably by stable transformation. These suitable introduction techniques, such as biolistic transformation, agrobacterium-mediated

- 11 -

transformation or electroporation, are known to a person skilled in the art (Sambrook et al. 2001).

In a further aspect the present invention relates to a transgenic plant or a part thereof, comprising a transgenic plant cell as described above. Here, a part may be a cell, a tissue, an organ or a combination of a number of cells, tissues or organs. A combination of a number of organs is, for example, a flower or a seed. In a particular embodiment the invention relates to a seed from the transgenic plant, wherein the seed comprises the nucleic acid molecule according to the invention as transgene. A transgenic plant of the present invention, in particular a plant of the Beta genus, preferably has a higher resistance to a pathogen, in particular BNYVV, than a corresponding non-transformed plant (plant without the transgene). The level of resistance for example to BNYVV may be defined qualitatively in plants of the Beta genus by determination of rating scores (rating score schemes for plants of the Beta genus are known from the prior art, for example for sugar beet Mechelke (1997)). A higher resistance presents itself in an improvement of the resistance by at least one rating score, by at least two rating scores, or by at least three or more rating scores. The invention also relates to a method for producing a transgenic plant comprising a step of introducing the nucleic acid molecule according to the invention or the vector of the present invention into a plant cell and optionally a step of selecting a transgenic plant cell. Furthermore, such a method for producing a transgenic plant is characterised by a subsequent step that includes regenerating the transgenic plant from the transgenic plant cell produced in the first step. Methods for regeneration are known from the prior art to a person skilled in the art.

In a further aspect the present invention also relates to a method for conveying or increasing a resistance to a pathogen, in particular BNYVV, in a plant, preferably a plant of the Beta genus, comprising a step of transforming a plant cell with the nucleic acid molecule according to the invention or the vector of the present invention. This method preferably leads to an improvement of the resistance by at least one rating score, particularly preferably to an improvement of the resistance by at least two, three or more rating scores. Rating score schemes for plants of the beta genus are known from the prior art, for example for sugar beet Mechelke (1997).

In a further aspect the present invention relates to a regulatory sequence of a promoter which controls the expression of a gene comprising the nucleic acid molecule according to the invention, characterised in that the regulatory sequence is able to convey or to modulate the expression of a heterologous DNA sequence as a result of a pathogen infection, and the regulatory sequence comprises a nucleic acid molecule having a nucleotide sequence

- 12 -

according to SEQ ID NO: 1 of nucleotides 1-1403. The heterologous DNA sequence is preferably a nucleotide sequence which encodes for a component of the plant pathogen defence (Ex.: resistance gene (R-gene) or gene encoding for enzymes involved in the signal transfer, such as kinases or phosphatases and for G-protein) or which encodes for a pathogenic effector (what are known as avirulence genes (avr)). Furthermore, the present invention includes a recombinant DNA molecule which comprises the above-described regulatory sequence. The recombinant DNA molecule is preferably operatively linked to a heterologous DNA sequence.

In a further aspect the present invention relates to a host cell transformed with the above-described regulatory sequence or with the specified recombinant DNA molecule and to a transgenic plant, plant tissue or plant cell comprising the regulatory sequence or the recombinant DNA molecule as transgene. The invention also provides a method for producing a transgenic plant cell comprising a step of introducing the regulatory sequence of the invention or the recombinant DNA molecule and optionally a step of selecting a transgenic plant cell. The invention also provides a method for producing a transgenic plant comprising a step of introducing the regulatory sequence of the invention or the recombinant DNA molecule into a plant cell and optionally a step of selecting a transgenic plant cell. Such a method for producing a transgenic plant is furthermore characterised by a subsequent step that includes regenerating the transgenic plant from the transgenic plant cell produced in the first step.

As already mentioned above, the RZ-3 resistance gene was identified during the course of a map-based cloning process. The process performed for example comprised the following steps: genetic fine-mapping, physical mapping, construction of a very large splicing population of more than 8000 F2 splicing descendants, recombinant screening, marker development in the target region, comparative BAC sequencing in resistant and sensitive genotypes, bioinformatic analyses, protein predictions and comparison of the proteins. Such laborious development work is extremely costly and it is unknown whether it is actually successful in identifying the gene. Following integration of the RZ-3 locus from *Beta vulgaris* subsp. *maritima* in a plant of the Beta genus, specifically in sugar beet (*Beta vulgaris* subsp. *vulgaris* var. *altissima*), markers having a good diagnostic value were developed for the tracking of the RZ-3 genome segment in the fine-mapping, which proved to be particularly difficult, since the target region is repetitive over wide areas. Surprisingly, however, it was possible to successfully develop a few diagnostic markers that in part also functioned only with a certain marker technique, such as pyrosequences, i.e. as PSQ markers, or that were zero allelic.

- 13 -

In spite of these technical difficulties, it was possible to delimit the RZ-3 locus to a genomic region of 0.67 cM by comprehensive analysis with use of these markers. This corresponds to a physical length of approximately 340,000 bp. In spite of intensive developments, it was only possible to a limited extent to further reduce the *Beta vulgaris* subsp. *maritima* introgression around the gene in a marker-assisted manner and to identify candidate genes for the RZ-3 gene. A further shortening of the introgression, however, is desirable in any case from a cultivation viewpoint in order to eliminate any potential “linkage drag”, closely coupled to the RZ-3 gene. Lastly, a target region could be limited to just approximately 0.07 cM in a number of steps by means of fine-mapping and with incorporation of sequence information from physical maps. This, however, was only possible since a total of 8004 were examined, including informative recombinant BC2S1 or BC2S2 plants, which were analysed intensively with 90-180 descendants in each case. This was necessary since the resistance expression was not always clear for unknown reasons. These descendants were genotyped into individual plants and phenotyped in parallel. By means of statistical methods (t-test, power analysis), the phenotypes of the informative recombinants (homozygote resistant – RR; heterozygote resistant – Rs; homozygote susceptible – ss) were detected and therefore conclusions could be made regarding the genotype of the informative recombinants.

In the relatively small target region of approximately 38,000 bp, ten genes could be annotated in the susceptible genotype. Overlapping clones from a resistant BAC library were identified for this target region with the aid of new markers, which describe specifically the target region, and were then sequenced. Due to the repetitiveness of the target region, the sequence of the susceptible genotype presented numerous small portions with unknown sequence content. For this reason, the assembling of the RR and ss sequences was particularly demanding. However, it was possible to identify a putative resistance gene. This contained, in practically all ss genotypes, a retrotransposon having a length of approximately 8000 bp between the LRR domain and the IR domain, which could not be detected in RR genotypes. An amino acid sequence predicted from the putative resistance gene sequence showed that the gene presumably encodes for an NB-ARC-LRR protein. It can be assumed that this insertion of the retrotransposon destroys the function of the gene in susceptible ss genotypes, since it separates the internal repeat domain (IR) from the two other domains (NB-ARC and LRR).

The comparison of the NBS-LRR gene in ss genotypes with that in RR genotypes also showed diagnostic polymorphisms, which can be inferred from Figures 1, 2 and 3. Based on these polymorphisms in the NBS-LRR gene, markers were developed and tested in a wide set of approximately 100 ss and RR genotypes. The marker patterns, but also the comparative sequencing in the target gene, confirmed that the insertion is practically always coupled with the

- 14 -

sensitivity. However, a few ss genotypes were found that did not have the retrotransposon insertion and were still susceptible. These ss genotypes, however, could be clearly distinguished from the RR genotypes by means of markers that describe the diagnostic polymorphisms according to Figures 1, 2 and/or 3.

In the analysed population, recombinants were identified in the target region that show a recombination between the NBS-LRR gene and the downstream, adjacent annotated putative gene, which could code for an ankyrin repeat protein. In the case of two plants, the recombinations can be found between the NBS-LRR gene and the upstream, adjacent annotated putative gene, which could code for a DUF565 protein (protein with unknown function). By means of the resistance analysis of the descendants of all these recombinant plants (single-gene removal upstream and downstream of the NBS-LRR gene), it was possible to demonstrate quite clearly that the gene between the ankyrin repeat gene and the DUF565 gene, specifically the NBS-LLR gene characterised here, is responsible for the resistance in the RR genotype. Figure 4 shows the physical map of the RZ-3 target region with the developed markers. The genotype data of eight dense recombinant lines, and also the statistical analysis of the descendants thereof are illustrated in Figure 5.

In a further aspect the present invention relates to a method for identifying a nucleic acid molecule that encodes a protein able to convey a resistance to the pathogen BNYVV in a plant of the Beta genus in which the protein is expressed. The method comprises the detection of the absence of an insertion in the encoding nucleotide sequence of the nucleic acid molecule. The method preferably comprises the detection of the absence of an insertion, in particular of a retrotransposon, in the coding nucleotide sequence of the nucleic acid molecule. The retrotransposon may be, for example, approximately 500 bp, approximately 1000 bp, approximately 2000 bp, approximately 4000 bp, approximately 8000 bp, or more than approximately 8000 bp long. In a particular embodiment of the method the nucleic acid molecule is the nucleic acid molecule according to the invention as described above and encodes the resistance-conveying RZ-3 gene or a functional homologue of RZ-3. The plant of the Beta genus is preferably *Beta vulgaris* subsp. *maritima* or *Beta vulgaris* subsp. *vulgaris* var. *altissima* (sugar beet). A person skilled in the art knows which methods are suitable for detecting the absence of the insertion. By way of example, a person skilled in the art in the knowledge of the nucleic acid molecule according to the invention disclosed here can develop molecular markers which detect the presence or the absence of an insertion in the above-described region in the NBS-LLR gene (see the examples for an exemplary approach). The present invention includes such markers and use thereof for the detection of the presence or absence of the insertion for the selection of resistant, in particular BNYVV-resistant, plants, in

- 15 -

particular *Beta vulgaris* subsp. *maritima* or *Beta vulgaris* subsp. *vulgaris* var. *altissima* (sugar beet). Such markers preferably describe loci at the insertion points of the retrotransposon. Insertion points mean transition points between genomic DNA and retrotransposon on 5' and/or 3' side of the insertion. Transition points are to be defined broadly, and marker loci may be arranged on the DNA at a distance of less than 1000 nucleotides, preferably less than 800 or 600 nucleotides, particularly preferably less than 400, 200, 150, 100, 50, 40, 30, 20 or 10 nucleotides upstream or downstream of an insertion point. Alternatively or in addition to the step of detecting the presence or absence of an insertion in the encoding nucleotide sequence of the nucleic acid molecule, the method may also comprise the detection of at least one polymorphism according to Figures 1, 2 and/or 3, preferably of at least 2 or 3 polymorphisms according to Figures 1, 2 and/or 3, particularly preferably of at least four, five or more polymorphisms according to Figures 1, 2 and/or 3 in the encoding nucleotide sequence of the nucleic acid molecule according to the invention with use of molecular markers which identify the polymorphisms, in particular diagnostic polymorphisms. This detection preferably takes place with use of at least one molecular marker per polymorphism, in particular per diagnostic polymorphism. A person skilled in the art knows which marker techniques are to be applied for the detection of a corresponding polymorphism, and how to construct molecular markers for this (literature). Furthermore, the present invention includes molecular markers which describe or detect a polymorphism according to Figures 1, 2 and/or 3, and also the use of a molecular marker for the detection of a polymorphism according to Figures 1, 2 and/or 3. Furthermore, the above identification methods also constitute methods for the selection of a plant which has a resistance to BNYVV. The selection method comprises a terminating step of selecting a resistant plant.

Furthermore, it was also possible to demonstrate that in the examined RR genotypes there were a genomic DNA sequence portion according to SEQ ID NO: 4 upstream of and adjacent to RZ-3 (SEQ ID NO: 1) as well as a genomic DNA sequence portion according to SEQ ID NO: 5 downstream of and adjacent to RZ-3 (SEQ ID NO: 1), which are closely coupled to the RZ-3 gene and are therefore extremely suitable as DNA regions for the development of diagnostic markers for RZ-3. The present invention therefore relates to a method for selecting a plant which has a resistance to BNYVV. The selection method comprises the use of a molecular marker on a DNA sequence according to SEQ ID NO: 4 and/or on a DNA sequence according to SEQ ID NO: 5 and a terminating step of selecting a resistant plant. A person skilled in the art knows how to develop and use markers on the basis of the disclosed sequence information.

By means of the present invention, the further following advantages can be obtained for the cultivation and development of new resistant plant lines of the Beta genus:

- 16 -

Sequence information and also the identified polymorphisms, which allow a distinction between resistant RR and susceptible ss alleles of the disclosed gene, make the marker development possible directly in the gene, which constitutes a significant facilitation for the plant grower in particular in view of the development of optimised elite lines without "linkage drag". In addition, the knowledge of the sequential structure can be used for the identification of further resistance genes, in particular against rhizomania, which for example are partially homologous.

The use disclosed here of the resistant gene allele in cis- or trans-genetic approaches opens up the possibility of developing new resistant species of the Beta genus, which on account of the dose effect have a higher resistance or in which, as a result of the stacking of the disclosed gene with other resistance genes, a resistance interruption can be avoided and the resistance expression can be optimised. Modifications of the gene by means of tilling or selective engineering are also conceivable for the development of new resistance alleles.

The present invention also relates to the use of the identified resistant RZ3 gene allele in a genetic or molecular stack with other genetic elements, which can convey agronomically advantageous properties, in a plant. As a result, the economical value of crop plants can be significantly increased, for example by increasing the yield performance or by developing new cultivation areas for a plant that were not accessible previously for the cultivation of these plants, *inter alia* due to biotic factors such as heavy pathogen pressure or abiotic factors such as dryness. An agronomically advantageous property is, for example, a tolerance to a herbicide, such as glyphosate, glufosinate or ALS inhibitors. Numerous further herbicides and the applicability thereof are known from the prior art to a person skilled in the art. Said person can refer to the prior art in order to gain knowledge as to which genetic elements are to be used and in which way in order to implement a corresponding tolerance in plants. A further example of an agronomically advantageous property is an additional pathogen resistance, wherein pathogens may be, for example, insects, viruses, nematodes, bacteria or fungi. By way of example, by combining different pathogen resistances/tolerances, a broad pathogen defence for a plant can be achieved, since genetic elements may have effects that supplement one another. Numerous resistance genes for example are known for this purpose to a person skilled in the art as genetic elements. A further example of an agronomically advantageous property is a tolerance of cool temperatures or frost. Plants that have this property could be sown earlier in the year or for example or could remain longer in the field, even during periods of frost, which for example may lead to increased revenue. Here too, a person skilled in the art could refer to the prior art in order to find suitable genetic elements. Further examples for agronomically advantageous properties are water utilisation efficiency, nitrogen utilisation efficiency and harvest. Genetic elements that can be used to convey such properties could be found in the prior art.

- 17 -

Numerous modifications for pathogen defence are also known to a person skilled in the art. Besides the frequently described families of the R genes, the Avr/R approach, the Avr gene complementation (WO 2013/127379), the auto-activation of an R-gene (WO 2006/128444), the HIGS (host induced gene silencing) approach (for example WO2013/050024) or the VIGS (virus induced gene silencing) approach could be used advantageously. In particular, the auto-activation of an R gene could be of significance for the present invention. A nucleic acid that encodes for an auto-activated resistance protein for producing a resistance to pathogens in plants is to be created for this purpose. This nucleic acid then has only a limited part of an NBS-LRR resistance gene, such as the RZ3 gene, which extends downstream from the 5'-end of the encoding region of the NBS-LRR resistance gene to the start of the NBS domain of the NBS-LRR resistance gene, wherein the NBS-LRR resistance gene is not a TIR-NBS-LRR resistance gene.

Furthermore, the invention also includes the use of the resistant RZ3 gene allele, identified with an above-described method, for combination with one of the above modifications or with an above-described genetic element which can convey one or more or agronomically advantageous properties in a plant.

Variants and embodiments of the present invention will be described in an exemplary manner with reference to the accompanying figures and sequences:

Sequences:

SEQ ID NO: 1	genomic DNA sequence of the resistance gene RZ-3. The sequence comprises nucleotide 1 to 1403 of the regulatory region of the promoter
SEQ ID NO: 2	predicted protein sequence of the resistance protein RZ-3_1
SEQ ID NO: 3	predicted protein sequence of the resistance protein RZ-3_2
SEQ ID NO: 4	adjacent chromosomal region upstream of RZ-3 (SEQ ID NO: 1)
SEQ ID NO: 5	adjacent chromosomal region downstream of RZ-3 (SEQ ID NO: 1)
SEQ ID NO: 6	consensus sequence of the genomic sequence of the RZ-3 gene in ss genotypes
SEQ ID NO: 7	target sequence in the RZ3 gene of the RNAi construct in the vector pZFN-C48-RNAi.

Figures:

Fig. 1 A-I: nucleotide sequence comparison between consensus sequence of the genomic sequence of the RZ-3 gene in ss genotypes (SEQ ID NO: 6) and the RZ-3 gene of the RR genotypes (SEQ ID NO: 1). Diagnostic polymorphisms are shown in grey and bold. Polymorphisms that are not diagnostic are underlined. The potential transcription start points in the gene are characterised by arrows. They lead to two polypeptide variants RZ-3_1 and RZ-3_2. The position of the retrotransposon is characterised by a black triangle at the top.

Fig. 2 A-L: amino acid sequence comparison of the predicted polypeptide from the RR genotype (RZ-3_1; SEQ ID NO: 2) and polypeptides from 22 different ss genotypes. Diagnostic polymorphisms are shown in grey and bold. Polymorphisms that are not diagnostic are underlined.

Fig. 3 A-L: amino acid sequence comparison of the predicted polypeptide from the RR genotypes (RZ-3_2; SEQ ID NO: 3) and polypeptides from 22 different ss genotypes. Diagnostic polymorphisms are shown in grey and bold. Polymorphisms that are not diagnostic are underlined.

Fig. 4: physical map of the RZ-3 target region. Five genes were annotated in the illustrated target region of the sensitive reference genotype: ("2" (DUF565), "3" (hypothetical protein), "4" (NBS-LRR candidate gene), "5" (retrotransposon) and "6" (ankyrin repeat)). The NBS-LRR candidate gene ("4") contains a retrotransposon ("5") in the sensitive reference sequence. This retrotransposon is missing completely in the resistance sequence, and therefore only four genes can still be annotated in the resistant genotype "2", "3", "4" and "6"). The positions of the densest recombinations (recombinants: 111T_3515/ ZR11007_03075 with numeral "7" and 111PB3645/ZR08093_05621 with numeral "8") are illustrated at the top. With the aid thereof, the shorter target region "1" could be limited. The markers developed for this purpose from the recombinant analysis are reproduced as black dashes in the lower part of the figure. A gene segment ("9") selected in part from the domain region "10" as target sequence was used for the

- 19 -

validation of the gene in the RNAi approach for the *gene* splicing of the resistant RZ-3 gene allele.

Fig. 5: marker analysis of the densest recombinants in the RZ-3 target region (small letters in bold-framed region are marker data generated *in silico*). The eight recombinant lines were phenotyped and genotyped with a total of 1051 descendants. The descendants were divided into 3 groups (RR resistant homozygote, Rs heterozygote, ss sensitive homozygote) on the basis of the marker data in the NBS-LRR candidate gene or the splicing flanking region in the event that the NBS-LRR candidate gene is homozygote RR or ss. In addition, the corresponding ELISA values are reproduced. A splicing or non-splicing of the descendants was examined by t-test and Wilcoxon statistic. The candidate gene could be quite clearly delimited between the markers s3e5800s01 and s3e5873s01 on the basis of the results.

Fig. 6: transformation vector pZFN-C48-RNAi: d35S-promoter; C48 s: C48 sequence sense orientation; AtAAP6 intron2: Arabidopsis thaliana amino acid permease 6 intron; C48 as: C48 sequence antisense orientation; Nos-T: nos terminator; LB flanking site: Left border flanking site; ZFN site: Zinc-finger nuclease recognition site (complementary); Pnos: Nos promoter; NPT: coding sequence; neomycin phosphotransferase (npt) gene; pAG7: pAG7 terminator; Bvpal3'UTR: 3' untranslated region of the Beta vulgaris Pal gene; LB: Left border; aadA: coding sequence; aminoglycoside-3"-adenylyltransferases (AAD); pVS1-REP: pVS1 replication origin; ColE1 ori: ColE1 replication origin; RB: right border.

Examples:

Mapping and fine-mapping of the RZ-3 gene/genetic physical map

The RZ-3 resistance (also referred to as C48 resistance or C48) was mapped in a number of steps by means of mapping and fine-mapping on chromosome 3 between 57.1 and 57.8 cM

- 20 -

(internal reference map), i.e. at a genetic distance between two flanking markers of 0.0714 cM in the genetic map. A total of 8004 plants for the crossing S504 (sensitive genotype) x T74 (resistant genotype) were examined for the mapping. Parallel to the C48 QTL mapping new informative markers were developed in a target-oriented manner following each mapping step and were used for the limitation of the C48 target region.

The fine-mapping coordinates were additionally confirmed with the analysis of the descendants of the informative recombinants. Informative recombinant BC2S1 or BC2S2 plants were analysed intensively for this purpose, in each case with 90-180 descendants. These descendants were genotyped and phenotyped in parallel on an individual plant basis. By means of static methods (t-test, power analysis), the phenotypes of the informative recombinants (homozygote resistant RR/heterozygote Rs/homozygote susceptible ss) were detected and conclusions could therefore be made regarding the genotype of the informative recombinants. Provided the homozygote classes of the descendants (RR versus ss) differed in terms of resistance, the gene was present in the heterozygote region (Rs) of the parent plant; otherwise it was present in the homozygote region (RR or ss) of the parent plant.

A physical map was generated for a rhizomania-sensitive genotype by projecting markers and genetic positions thereof onto the chromosome sequences. With the limitation of the C48 QTL region, new informative markers were developed on the basis of the reference sequence and additional comparative sequencings in resistant genotypes (next generation sequencing and Sanger sequencing).

The region identified by the fine-mapping comprises a sequence length of 37996 base pairs (positions of flanking SNP markers) in the sensitive reference sequence. The collinearity between the genetic and the physical map in the target region is consistent (sequence of 12 markers in the target region).

Identification and sequencing of resistant BAC clones

A BAC library has been developed for a selected RZ-3 (C48) resistant genotype. This BAC bank was sampled with the used markers in the C48 QTL region. A number of BAC clones were found for the above-identified target region. Of these, three BAC clones of different length, which detected the target region completely, were selected for sequencing. The BAC clones were sequenced and a "de novo" assembling was carried out on the basis of the resultant sequence-reads. Among the resultant resistant sequence contigs, the greatest sequence had a length of 110909 bp (34537 reads) and comprised the target region completely.

Comparison of sensitive and resistant sequences - sequence evaluation

The collinearity of the two ss and RR sequences was compared with use of different software tools. For both resistant and sensitive sequences, a gene annotation was performed using Maker and Pedant softwares. The gene annotation on both sequences demonstrated the same sequence of putative genes. Surprisingly, however, a significant difference in one of these genes could be determined, specifically in the gene of the present invention (RZ-3). A retrotransposon could be annotated in the sensitive genotype in this identified NBS-LRR gene. The insertion of the retrotransposon passed in the gene between the two domains of the LRR domain and the IR domain. The resistant genotype does not have this insertion and is reproduced in SEQ ID NO: 1. Furthermore, the predicted polypeptide sequences were then compared and evaluated (illustrated partially in Figures 2 and 3).

Comparative sequencing of the NB-ARC-LRR candidate gene

The NB-ARC-LRR candidate gene was sequenced comparatively in two steps. The retrotransposon insertion point was verified in a genotype set having a total of 92 resistant and sensitive genotypes. This analysis showed that none of the resistant genotypes had a retrotransposon insertion. Of these sensitive genotypes, the insertion could be detected in more than 90% of cases. The detection of the insertion therefore appears to be coupled with the susceptible genotype. Due to the found inconsistencies (approximately 10% of the remaining susceptible genotypes without insertion) however, the sequencing was extended with promoter region in the second step for the entire gene before the insertion point (SEQ ID NO: 1). On the whole, 31 selected resistant and susceptible genotypes inclusive of the inconsistent genotypes were sequenced and compared. As a result, all resistant genotypes, which are to be attributed to seven different resistance sources, were 100% identical to the compared approximately 4100 base pairs. In addition, completely diagnostic polymorphisms were found in the nucleotide sequence, of which a number lead to amino acid substitutions in the protein sequence (see Figures 1, 2 and 3). Some of these substitutions, particularly in the domain regions, could cause the functional loss of the identified resistance protein in the ss genotypes. Furthermore, three INDELs coupled fully with the resistance (Linkage disequilibrium = 1) were also found in the promoter region (Figure 1). These INDELs are also to be considered as potential candidates for the functional loss.

Verification of the gene by means of dense recombinants

- 22 -

In the analysed population with 8004 plants, 16 recombinants were identified in the target region (fine-mapped region with 37996 base pairs). Of these 16 genotypes, 9 plants contained the recombination between the NB-ARC-LRR protein and the adjacent ankyrin repeat protein on the right-hand side. In the case of two plants the recombinations are between the NB-ARC-LRR protein and the adjacent DUF565 proteins to the left (protein with unknown function). By means of the analysis of the progeny of all of these recombinant plants (a gene distance to the left and right) it was possible to demonstrate quite clearly that the gene lies between DUF565 and the ankyrin repeat protein, specifically that only the NB-ARC protein is responsible for the resistance.

Exemplary detection of the absence of the transposon insertion

For the detection of the retrotransposon insertion, 3 special dominant primer combinations were developed. The first and the second primer combinations are able to detect the insertion, since in each case one primer of the two primer pairs sits in the retrotransposon (left or right flank of the retrotransposon) and the second primer binds directly before or after the retrotransposon. A third primer pair detects the absence of the retrotransposon in that the primers find a binding point before and after the retrotransposon. A PCR product may then be produced under standard conditions only when the retrotransposon is missing, otherwise, with the retrotransposon, the PCR product would be too large and no amplicon would be created in this case.

Verification of the gene by means of RNAi approach

Besides the above-described verification of the gene by means of dense recombinants, a further detection of the resistance effect of the gene was also performed by means of RNA interference. For this purpose, a resistant standard sugar beet genotype was transformed with a DNA construct, which encodes a double-stranded hairpin RNA. This dsRNA was able to effect a gene silencing post-transcriptionally, which would reduce or switch off the effect of the resistant RZ-3 gene allele, whereby the previously resistant sugar beet genotype should become sensitive to rhizomania.

In order to provide a suitable DNA construct, a defined target sequence region of the resistant RZ3 gene allele of 434 base pair length (SEQ ID NO: 7; Figure 4) was selected, amplified by PCR and cloned both in sense and antisense direction in the vector pZFN, which is suitable for

- 23 -

the synthesis of hairpin structures (Figure 6). This vector has a doubled CaMV 35S promoter, a multiple cloning point, an intron from the gene AtAAP6, which encodes in *Arabidopsis thaliana* for an amino acid permease, a further multiple cloning point, and a nos terminator. The transformation of the sugar beet with the provided vector was performed in accordance with the protocol of Lindsey & Gallois (1990) with use of the antibiotic kanamycin as selection marker. Following a number of selection steps, successful transformation was examined on transgenic shoots via PCR by means of detection of the presence of the nptII gene, the AAP6 intron and the two t-DNA border sequences (LB/RB) and the absence of *vir*. Positive shoots were clonally multiplied *in vitro* to 30 shoots in each case, rooted, and transferred into earth in the greenhouse. Approximately 2 weeks later the transgenic sugar beet plants were planted in earth contaminated by rhizomania, in which they were cultivated for 8 to 10 weeks. As control, non-transformed plants of the same resistant genetic standard transformation background were used under the same conditions. In order to detect the spread of rhizomania, the roots of the sugar beet plants were harvested and quantified by means of ELISA test of the BNYVV attack, wherein a low ELISA value indicated a resistance and a high value indicated a sensitivity (Mechelke 1997, Clark & Adams 1977). The ELISA value of the transformed sugar beet, with a mean value of 3.55, was significantly higher than the ELISA value of the control, which was also resistant, with a mean value of 1.27 and was compatible to the sensitive standard D108_ss (Table 1). The results of the ELISA test accordingly showed that a previously resistant plant was sensitive to BNYVV as a result of the specific gene silencing of the resistant RZ-3 allele in the transformation background. Consequently, the gene of the present invention could be clearly verified as the resistance gene RZ3.

- 24 -

Table 1: Results of the ELISA- test following statistical analyses (D108_ss = sensitive standard; 6921_RR = resistant transformation background; 6921_RNAi = resistant transformation background with dsRNA directed against RZ3 gene).

	D108_ss	6921_RR	6921_RNAi
n	6	25	64
mean value	3.98	1.27	3.55
standard error	0.02	0.25	0.11
standard deviation	0.06	1.24	0.87
<i>T-test (significance level): $p < 0.0001$</i>			

References

- Clark, M.F.; Adams, A.N. (1977): Characteristics of the microplate method of enzyme-linked immunosorbent assay for the detection of plant viruses. *J. Gen. Virol.* 34, 475–483
- Depicker A, Stachel S, Dhaese P, Zambryski P, Goodman HM (1982) Nopaline synthase: transcript mapping and DNA sequence. *J Mol Appl Genet.* 1(6): 561-73.
- Esser K (2000) Kryptogamen 1: Cyanobakterien Algen Pilze Flechten Praktikum und Lehrbuch. Springer Publishing House, Berlin, Heidelberg, 3rd edition. 2000.
- Gidner S, Lennefors BL, Nilsson NO, Bensefelt J, Johansson E, Gyllenspetz U, Kraft T (2005) QTL mapping of BNYVV resistance from the WB41 source in sugar beet. *Genome* 48: 279-285.
- Larson RL, Wintermantel WM, Hill A, Fortis L, Nunez A (2008) Proteome changes in sugar beet in response to Beet necrotic yellow vein virus. *Physiological and Mol. Pl. Pathol.* 72: 62-72.
- Lindsey, K., and P. Gallois (1990) "Transformation of sugar beet (*Beta vulgaris*) by *Agrobacterium tumefaciens*." *Journal of experimental botany* 41.5: 529-536.
- Martin GB, Bogdanove AJ; Sessa G (2003) Understanding the functions of plant disease resistance proteins. *Annual Review of Plant Biology* 54: 23-61.
- Mechelke W (1997) Probleme in der Rizomaniaresistenzzüchtung, Vorträge für Pflanzenzüchtung, Resistenzzüchtung bei Zuckerrüben, Gesellschaft für Pflanzenzüchtung e.V., 113-123.
- Odell JT, Nagy F, Chua N-H (1985) Identification of DNA sequences required for activity of the cauliflower mosaic virus 35S promoter. *Nature* 313, 810 - 812
- Rushton PJ, Torres JT, Parniske M, Wernert P, Hahlbrock K, and Somssich IE (1996) Interaction of elicitor-induced DNA-binding proteins with elicitor response elements in the promoters of parsley PR1 genes. *EMBO J.* 15(20): 5690–5700.
- Sambrook J, Russell DW (2001) *Molecular Cloning. A laboratory manual*, Cold Spring Harbor Laboratory Press, 3. Aufl. 2001.
- Schmidlin LEDEB, Weyens G, Lefebvre M, Gilmer D (2008) Identification of differentially expressed root genes upon rhizomania disease. *Mol. Plant Pathol.* 9(6):741-51.

- 26 -

Scholten OE, Bock TSMD, Klein-Lankhorst RM, Lange W (1999) Inheritance of resistance to Beet necrotic yellow vein virus in Beta vulgaris conferred by a second gene for resistance. Theor. Appl. Genet. 99:740-746.

Sohi HH, Maleki M(2004) Evidence for presence of types A and B of beet necrotic yellow vein virus (BNYVV) in Iran. Virus Genes 29(3): 353-8.

Van Ooijen G, Mayr G, Kasiem MMA, Albrecht M, Cornelissen BJC, Takken FLW (2008) Structure–function analysis of the NB-ARC domain of plant disease resistance proteins. Journal of Experimental Botany, 59(6): 1383-1397

WO/2000/29592 (Max-Planck-Gesellschaft zur Förderung der Wissenschaften e.V.). Chimeric promoters capable of mediating gene expression in plants upon pathogen infection and uses thereof.

WO/2006/128444 (KWS SAAT AG). AUTOACTIVATED RESISTANCE PROTEIN.

WO/2007/147395 (KWS SAAT AG). Pathogen induzierbarer synthetischer Promotor.

WO/2013/127379 (KWS SAAT AG). PATHOGEN-RESISTANT TRANSGENIC PLANT.

WO/2013/050024 (KWS SAAT AG). TRANSGENIC PLANT OF THE SPECIES BETA VULGARIS HAVING ENHANCED RESISTANCE TO CERCOSPORA.

WO/2013/091612 (KWS SAAT AG). NOVEL PLANT-DERIVED CIS-REGULATORY ELEMENTS FOR THE DEVELOPMENT OF PATHOGEN-RESPONSIVE CHIMERIC PROMOTORS.

Claims:

1. A nucleic acid molecule which encodes a polypeptide that is able to convey a resistance to a pathogen in a plant in which the polypeptide is expressed, characterised in that the nucleic acid molecule comprises a nucleotide sequence selected from
 - a) a nucleotide sequence that encodes a polypeptide having an amino acid sequence according to SEQ ID NO: 2 or SEQ ID NO: 3,
 - b) a nucleotide sequence comprising the encoding sequence of the DNA sequence according to SEQ ID NO: 1,
 - c) a nucleotide sequence that hybridises with the complementary sequence of the nucleotide sequence according to a) or b) under stringent conditions,
 - d) a nucleotide sequence that encodes a polypeptide which is derived by substitution, deletion and/or addition of one or more amino acids of the amino acid sequence encoded by the nucleotide sequence according to a) or b) from a polypeptide encoded by the nucleotide sequence according to a) or b),
 - e) a nucleotide sequence that encodes a polypeptide which has an amino acid sequence at least 60% identical to an amino acid sequence encoded by the nucleotide sequence according to a) or b), or
 - f) a nucleotide sequence that encodes at least one nucleotide-binding domain (NBS) corresponding to amino acid positions 168-227 of SEQ ID NO: 2 or corresponding to amino acid positions 182-241 of SEQ ID NO: 3, at least one leucine-rich domain (LRR) corresponding to amino acid positions 591-613 of SEQ ID NO: 2 or corresponding to amino acid positions 605-627 of SEQ ID NO: 3 and/or at least one internal repetitive domain (IR) corresponding to amino acid positions 1013-1072 of SEQ ID NO: 2 or corresponding to amino acid positions 1027-1086 of SEQ ID NO: 3
2. A vector that comprises the nucleic acid molecule according to claim 1.
3. A host cell that comprises the nucleic acid molecule according to claim 1 or the vector according to claim 2.

4. A polypeptide that is able to convey a resistance to a pathogen in a plant in which the polypeptide is expressed and that is encoded by the nucleic acid molecule according to claim 1.
5. A transgenic plant cell that comprises the nucleic acid molecule according to claim 1 as transgene or that comprises the vector according to claim 2.
6. A transgenic plant or a part thereof, comprising a plant cell according to claim 5.
7. A seed of the plant according to claim 6, wherein the seed comprises the nucleic acid molecule according to claim 1 as transgene.
8. A method for producing a transgenic plant cell, characterised in that the method comprises a step of introducing the nucleic acid molecule according to claim 1 or the vector according to claim 2 into the plant cell.
9. A method for producing a transgenic plant, characterised in that the method comprises the following steps
 - a) introducing the nucleic acid molecule according to claim 1 or the vector according to claim 2 into a plant cell, and
 - b) regenerating the transgenic plant from the transgenic plant cell from step a).
10. A regulatory sequence of a promoter which controls the expression of a gene comprising the nucleic acid molecule according to the invention, characterised in that the regulatory sequence is able to convey or to modulate the expression of a heterologous DNA sequence as a result of a pathogen infection, and the regulatory sequence comprises a nucleic acid molecule having a nucleotide sequence according to SEQ ID NO: 1 of nucleotides 1-1403.
11. A recombinant DNA molecule that comprises the regulatory sequence according to claim 10.
12. A host cell that is transformed with the regulatory sequence according to claim 10 or that is transformed with the recombinant DNA molecule according to claim 11.
13. A transgenic plant, plant tissue or plant cell that comprises the regulatory sequence according to claim 10 or the recombinant DNA molecule according to claim 11 as transgene.

14. A method for identifying a nucleic acid molecule which encodes a protein that is able to convey a resistance to the pathogen BNYVV in a plant of the Beta genus in which the protein is expressed, characterised in that the method comprises the following step
- i. detecting the absence of an insertion in the encoding nucleotide sequence of the nucleic acid molecule according to claim 1, or
 - ii. detecting at least one polymorphism according to Figures 1, 2 and/or 3 in the encoding nucleotide sequence of the nucleic acid molecule according to claim 1 with use of molecular markers that identify the polymorphism.

FIG 1 A

consensus-sensitive resistant sequence	1	CAAATCTTCTGGCATCAATGGCGGTGTTGCCGTTCAATTTAAACATCAATGGAGGTAAGATCATGTT	70
	(1)	CAAATCTTCTGGCATCAATGGCGGTGTTGCCGTTCAATTTAAACATCAATGGAGGTAAGATCATGTT	
consensus-sensitive resistant sequence	71	TTTTCAACAATATAAAACTTATA-----TGATTTTTC TGTTTTC	140
	(71)	TTTTCAACAATATAAAACTTATACTTCCTCTGTTTAAATGAACGTTTGTTTCTCACAACC	
consensus-sensitive resistant sequence	141	-----CAACCCACTTTTAAATAAATAATTTTAGTTGTCGACGTAATAATAATAAATTG	210
	(113)	-----CAACCCACTTTTAAATAAATAATTTTAGTTGTCGACGTAATAATAATAAATTG	
consensus-sensitive resistant sequence	211	--GTATCTTGTGAGATTGTGATTATTAAGAGAGTCAAGTCTCACAAATATTCGAAAGTCTACGTAATC	280
	(114)	ATAGTATCTTGTGAGATTGTGATTATTAAGAGAGTCAAGTCTCACAAATATTCGAAAGTCTACGTAATC	
consensus-sensitive resistant sequence	281	CACCTCAAATTGACGAAGAAACAAGCAGGAAAGGATTAAGTAAGTTCGTGGAAACCACTAGAAATTAATTT	350
	(181)	CACCTCAAATTGACGAAGAAACAAGCAGGAAAGGATTAAGTAAGTTCGTGGAAACCACTAGAAATTAATTT	
consensus-sensitive resistant sequence	351	TCAAATATAGCTCTACCTAATAATATGGCCTACTTTTAAATTTTAAATAAGAAAGGTAATGTGATTAGAA	420
	(251)	TCAAATATAGCTCTACCTAATAATATGGCCTACTTTTAAATTTTAAATAAGAAAGGTAATGTGATTAGAA	
consensus-sensitive resistant sequence	421	ACAAATTGGTCTTAAATTATTCATTAAAGCTTAATAATGCATAAACTTTATCAAGTGCTATCTTTTCA	490
	(321)	ACAAATTGGTCTTAAATTATTCATTAAAGCTTAATAATGCATAAACTTTATCAAGTGCTATCTTTTCA	
consensus-sensitive resistant sequence	491	GGCCCGTCTTGAAGATTTTKGGKCCCRGTTCTATTATGAAAAATTGRGCCCTAAATTTATASAAAATAAA	560
	(391)	GGCCCGTCTTGAAGATTTTGGGCCCGGTTCTATTATGAAAAATTGGGCCCTAAATTTATAGAAAATAAA	

FIG 1 C

consensus-sensitive resistant sequence	(779) (1096)	1121 AGATGAATTGTTGTCTTTGATGGTCTCCAATGCATATTTTGATATWMTTAGGAATTYTAATTATGTACTA AGATGAATTGTTGTCTTTGATGGTCTCCAATGCATATTTTGATATCTTAGGAATTCTAATTATGTACTA	1190
consensus-sensitive resistant sequence	(849) (1166)	1191 TTAGTASAYATYGAGAYGAATACARAATYGCCATAATGAAGTATGATTATTTTARTTATATACTTTCTCC TTAGTAGACATTTGAGATGAATACATAATYGCCATAATGAAGTATGATTATTTTAGTTATATACTTTCTCC	1260
consensus-sensitive resistant sequence	(919) (1236)	1261 GTTCCAAATATATAAARTGTAAACACTTGTGTACTTTATGCGTACTAATGCATAAYAACGTGCACCTCTC Y G ATTCCAAATATATAAATGTAAACACTTGTGTACTTTATGCGTACTAATGCATAACAACGTGCACCTCTC A TG	1330
consensus-sensitive resistant sequence	(989) (1306)	1331 TGTTTAATTATATACTTTTGTAGAGAAGTGTACATTGGGACCATGGGACTGTGTATAATTGACCGCA TGTTTAATTATATACTTTTGTAGAGAAGTGTTACATTGGGACCATGGGACTGTGTATAATTGACCGCA	1400
consensus-sensitive resistant sequence	(1059) (1376)	1401 ↓Start1 AAATYGAAGTGTYGCATTTGATTGAAAAYGGAGARRGTAGTATATAGRTGGAACACNNNNNNNNNTGCTG AAATCGAAGTGTCGCATTTGATTGAAAATGGAGAGAGTAGTATATAGATGGAACACAGACGAGACTGCTG	1470
consensus-sensitive resistant sequence	(1129) (1446)	1471 RTCATCTTTGGCCAAACAAACCMAAAATTGATATTAATCCYYTWWTYMRGKYWNNNTTTCATCTTTTGGACA GTCATCTTTGGCCAAACAAACCCTTAATCCCTTATTTCAGGTCA--TTTCATCTTTTGGACA	1540
consensus-sensitive resistant sequence	(1199) (1514)	↓Start2 CAAAATGGATGTTGTAGGCWCTGCGCTATCTGCTGCCCAATCTCTGTTGCGAGCCCTGCAAAGTTCTGAG CAAAATGGATGTTGTAGGCACCTGCGCTATCTGCTGCCCAATCTCTGTTGCGAGCCCTGCAAAGTTCTGAG	1610
consensus-sensitive resistant sequence	(1269) (1584)	1611 CTCAAAGAGATCCTCTCGATCTTTGGCTACAAATCCCC A ACTTGATGACCTCCAACGCAYTGTMTCTACCA CTCAAAGAGATCCTCTCGATCTTTGGCTACAAATCCCC G ACTTGATGACCTCCAACGCAC T GTCTCTACCA	1680

FIG 1 D

consensus-sensitive resistant sequence	(1339) (1654)	1681 TCAAYGCTGTATTCCGTGATGCTGAGACCAAAACAGGAGCTCACTCATGAAGCACARCATTTGGCTCGAGGA TCAACGGCTGTATTCCGTGATGCTGAGACCAAAACAGGAGCTCACTCATGAAGCACAGCATTTGGCTCGAGGA	1750
consensus-sensitive resistant sequence	(1409) (1724)	1751 ACTCAAGGATGCTGTCCTTTGAAGCAGATGATCTGTTGACGAGTTTGTCACTCTTGCCGAGCAGAAAGCAA ACTCAAGGATGCTGTCCTTTGAAGCAGATGATCTGTTGACGAGTTTGTCACTCTTGCCGAGCAGAAAGCAA	1820
consensus-sensitive resistant sequence	(1479) (1794)	1821 CTTGTAAGGCTGGTGGCAGTCTTTCCAAAAAGATGCGCCCAATTCTTTTCTGATTCCAACCCCTTGGCA CTTGTAAGGCTGGTGGCAGTCTTTCCAAAAAGATGCGCCCAATTCTTTTCTGATTCCAACCCCTTGGCA	1890
consensus-sensitive resistant sequence	(1549) (1864)	1891 TYGCTTATARGATGTCACRAGGGGTTAAGAAGATCAAGAAGAGTTGGATGYTATYGCCTTACAATCATCA TTGCTTATAGGATGTCACGAGGGGTTAAGAAGATCAAGAAGAGTTGGATGCTATCGCTTACAATCATCA	1960
consensus-sensitive resistant sequence	(1619) (1934)	1961 ATTTAGCTTTAAGATTGATCTTGAGCCCTAT AA AGAGAGAGAGGCT CG AGACTGGTTCTGTCTGAACGCA ATTTAGCTTTAAGATTGATCTTGAGCCCTAT G AAAGAGAGAGAGGCT AG AGACTGGTTCTGTCTGAACGCA	2030
consensus-sensitive resistant sequence	(1689) (2004)	2031 GGTGATATCATTTGGAAGAGAGGAT TG ACTTGGAGAAAGAT CG TAGGTTTGTTKCTTGATTTCTAACATCCAGC GGTGATATCATTTGGAAGAGAGGAT CG ACTTGGAGAAAGAT TG TAGGTTTGTTGCTTGATTTCTAACATCCAA C	2100
consensus-sensitive resistant sequence	(1759) (2074)	2101 GTGATGTGCTTTCCCTTACKATWGTGGGAATGGAGGGTTGGGTAAAACTGCTCTTGCCCAACTCGTGTA GTGATGTGCTTTCCCTTACTAT T GTGGGAATGGAGGGTTGGGTAAAACTGCTCTTGCCCAACTCGTGTA	2170
consensus-sensitive resistant sequence	(1829) (2144)	2171 CAATGATCCAAGGGTCAGAACTGCTTTTCCATTGAGATGTTGGAATTGTNN CT GATCAAGATCAAA AAK CAATGATCCAAGGGTCAGAACTGCTTTTCCATTGAGATGTTGGAATTGT GT CTGATCAAGATCAAA AAG	2240

FIG 1 E

consensus-sensitive resistant sequence	(1899) (2214)	2241 MAACTAGATGTGAAAGAAATTTTGGGTAAGATTCTGTCTACAGCTACTGGTAAGAAATCAYRAGGGTTCAA CAACTAGATGTGAAAGAAATTTTGGGTAAGATTCTGTCTACAGCTACTGGTAAGAAATCATGAGGGGTTCAA	2310
consensus-sensitive resistant sequence	(1969) (2284)	2311 CCATGGATCAKGTGCAAACCYAACTACRRGAACAACCTATGTGGCAAGAGATACTTGCTTGTTTTGGATGA CCATGGATCAGGTGCAAACCCAACTACGAGAAACAACCTATGTGGCAAGAGATACTTGCTTGTTTTGGATGA	2380
consensus-sensitive resistant sequence	(2039) (2354)	2381 TGTATGGAATGAGAAATCCTAATCAATTGCGTDWYYTKGKWRAATTCTTTCATGGGAGGTC AAAGGGG AAAT TGTATGGAATGAGAAATCCTAATCAATTGCGTGATCTGGTAGAAATTCTTTCATGGGAGGTC GAAAGCA GAAAT	2450
consensus-sensitive resistant sequence	(2109) (2424)	2451 TGGATTSTGGTAACTACGCCGTTTCGCAYGAGACARCGAGAAATTATAAGAGATGGTCCATTGCACAAGCTSC TGGATTSTGGTAACTACGCCGTTTCGCACGAGACAGCGAGAAATTATAAGAGATGGTCCATTGCACAAGCTCC	2520
consensus-sensitive resistant sequence	(2179) (2494)	2521 AAGGTTTGTCTGARRNAAACTYTTGGCGTTTATYTGTAAGGTGGACCTTCGGATCAGTGCAACCAAAAT AAGGTTTGTCTGAGG-AAAAC CTCT TGGCGTTTATTTGT CA GGTGGACCTTCGGATCAGTGCAAGCAAAAT	2590
consensus-sensitive resistant sequence	(2249) (2563)	2591 TCCCTAATGACTTT GT CATGATTGCACGAGATATAGTYGACAAATGTGCTCGAAACCCCTYTGGCTATAAG TCCCTAATGACTTT AT CATGATTGCACGAGATATAGTTGACAAATGTGCTCGAAACCCCTCTGGCTATAAG	2660
consensus-sensitive resistant sequence	(2319) (2633)	2661 AKTGGTAGGAAGCTTTTGTGTGGTCAAGACAAGAGTAAGTGGCTTTCATTTCATGAGATM TG TTTAGCC AGTGGTAGGAAGCTTTTGTGTGGTCAAGACAAGAGTAAGTGGCTTTCATTTCATGAGATC GAT TTTAGGC	2730
consensus-sensitive resistant sequence	(2389) (2703)	2731 AACATTAGAAAGAGYCATAAATGATATCATGCYAATACTGAACCTAAGTTACCATCATCTTGAACCTCCAA AACATTAGAAAGAGCC CA TAAATGATATCATGCCAATACTGAACCTAAGTTACCATCATCTTGAACCTCCAA	2800

FIG 1 F

consensus-sensitive resistant sequence	(2459) (2773)	2801 T CA KGAGATGCTTTAGTTATTGTGCARTGTTCCAAAGGATTTCCCTTATAGGGAAG A AGACG TT GTGATAAA T TA AGAGATGCTTTAGTTATTGTGCAGTGTTC CA AAAGATTTCCCTTATAGGGAAG CA GACG CT GTGATAAA	2870
consensus-sensitive resistant sequence	(2529) (2843)	2871 CCT TT TGGATGGCACAAAGTTATATTGTTCC A TTAGACAAAGATCAAAGCATAGATGAYGCTAGTGAGGAA CCT CT TGGATGGCACAAAGTTATATTGTTCC G TTAGACAAAGATCAAAGCATAGATGATGATGCTAGTGAGGAA	2940
consensus-sensitive resistant sequence	(2599) (2913)	2941 TACATATCAATTTTGYTG CA GAGATGTTTTTTTCGAAAA CA TCGGAA CA CAGAAAAAGAT K ATGTTATTAAAGA TACATATCAATTTTGT CG GAGATGTTTTTTTCGAAAA TG TCGGAG CA CAGAAAAAGAT G GTTGTTATTAAAGA	3010
consensus-sensitive resistant sequence	(2669) (2983)	3011 T A CATGATCTCATGCGATGATATTGCTCAAAATGTCATGGGGAAGGAGCTTTGTACGAC AA AAAAACATTAG T CC ATGATCTCATGCGATGATATTGCTCAAAATGTCATGGGGAAGGAGCTTTGTACGAC TA AAAAACATTAG	3080
consensus-sensitive resistant sequence	(2739) (3053)	3081 TGGCAGCTTGGATAAA A ATGTTCCGCATCTATCTCTTGCCAGAACTAGTTTTGCAAGATACTCTTTYAAT TGGCAGCTTGGATAAA G TGTTCCGCATCTATCTCTTGCCAGAACTAGTTTTGCAAGATACTCTTTC CA AT	3150
consensus-sensitive resistant sequence	(2809) (3123)	3151 GCAACTCATATTGCTCCYATTTCTRTGCTGGCTACTGGTGTCAGGAWKCTGAGATAAMCCAGTTYTCAG GCAACTCATATTGCTCC T ATTTCTGTGCTGGCTACTGGTGTCAGGATGCTGAGATA AA ACCAGTTTTCAG	3220
consensus-sensitive resistant sequence	(2879) (3193)	3221 TTGAGGCATTAGTACCAAAYTGT TC CCCTAAGGGCA TT GKACCTSGCTTGGTCGAAGATA AAAA AGTKT TTGAGGCATTAGTACCAA CT GT TT GT AC CTAAGGGCA TT GACCTCGCTTGGTCGAAGATA AAAA AGTTT	3290
consensus-sensitive resistant sequence	(2949) (3263)	3291 ACCAGACTCRATTGGTGGATTGTGCATTTGAGGTACTTAGATCTTTCRTATAASGAAGATYTGGAAGTA ACCAGACTCGATTGGTGGATTGTGCATTTGAGGTACTTAGATCTTTCGTATA AA CGAAGAT CT GGAAAGTA	3360

FIG 1 G

consensus-sensitive resistant sequence	(3019) (3333)	3361 CTTCCGAACTCAATTGCYAACTATATAATCTRCAAACCTTACAATTGAAGGGTTGCAAGAGATTGGAAG CTTCCAAACTCAATTGCTAAACTATATAATCTACAAACCTTACAATTGAAGGGTTGCAAGAGATTGGAAG	3430
consensus-sensitive resistant sequence	(3089) (3403)	3431 GGTTAYCAAAACATTTGAGCAGGCTGGTTAAGCTTCAAACCTTTRGATATAAATGGTTGCAAYAAATGTAAC GGTTA ~~C~~ CAAAACATTTGAGCAGGCTGGTTAAGCTTCAAACCTTTGGATATAAATGGTTGCAACAATGTAAC	3500
consensus-sensitive resistant sequence	(3159) (3473)	3501 TTATATGCCCAAGGCATGGGTAAGATGACTTGCCCTTCACACTCTCAGTAAGTTTATAGTGGTGGAGAA TTATATGCCCAAGGCATGGGTAAGTTGACTTGCCCTTCACACTCTCAGTAAGTTTATAGTGGTGGAGAA	3570
consensus-sensitive resistant sequence	(3229) (3543)	3571 GGGARTTGTCAAGTTGGAAGMAAYGGTTTGATGGG ~~GW~~ GGAAGATCTAAAGGCTCTCAAGAAACCTAAAGG GGGAGTTGTTC ~~A~~ AGTTGGAAGCAATGTTTGGATGGGTTGGAAGATCTAAAGGCTCTCAAGAAACCTAAAGG	3640
consensus-sensitive resistant sequence	(3299) (3613)	3641 GTCATCTGGRAATCCAAATCAGGTGGCCCCGAAAATACTACAGATGCTGTCAAGGAAGATGTTAAGAGGGA GTCATCTGGAAATCCAAATCAGGTGGCCCCAAAATACTACAGATGCTGTCAAGGAAGATGTTAAGAGGGA	3710
consensus-sensitive resistant sequence	(3369) (3683)	3711 AGGATTATACYTGAATCATAAGGAACATCTCAATCACATTGTGGTTGATTTAGATGTGAGGAGGGTGGT AGGATTATACCTGAATCATAAGGAACATCTCAATCACATTGTGGTTGATTTAGATGTGAGGAGGGTGGT	3780
consensus-sensitive resistant sequence	(3439) (3753)	3781 GGAAGAAATGGATGATGAGGAAGCAAGAAGATTGATGGAAGAGYTGCGGCCACATCCTTATCTTGAAAAATT GGAAGAAATGGATGATGAGGAAGCAAGAAGATTGATGGAAGAGTTGCGGCCACATCCTTATCTTGAAAAATT	3850
consensus-sensitive resistant sequence	(3509) (3823)	3851 TGGCTGTGAAAGCATAATATGGTGYGAAAAYGCCCTGRTTGGGYAACCCCTTCTYCCAAATCTTACAGAGCT TGGCTGTGAAAGCATAATATGGTGTGAAAATGCCCTGGTTGGGCAACCCCTTCTCCCAAATCTTACAGAGCT	3920

FIG 1 H

consensus-sensitive	3921	(3579)	TTWTC	3990
resistant sequence		(3893)	TTTCTTTCTGATTGTGGGAAAYGGAGWRCCTTCCATGCMTGGGAAACTTGGWTYDTCTRAAMGTYCTC	
consensus-sensitive	3991	(3649)	CGRCTTTCGCATTTGGCRAAATTGGAGTAYATTGRAGAAAGATAGCWCATCAGCTMWTTTTCAGKTKTAGGC	4060
resistant sequence		(3963)	CGACTTTCGCATTTGGCAAAATTGGAGTACATTGAAGAAAGATAGCTCATCAGCTAAATTTTCAGGTGTAGGC	
consensus-sensitive	4061	(3719)	CTGGACCCRGAAAGTGCAGGACTATCATTTATACTTCCCCCTCCCTTGAAACKCCTTGAGTTGAAGCRTTTGYR	4130
resistant sequence		(4033)	CTGGACCAAGAAAGTGCAGGACTATCATTTATACTTCCCCCTCCCTTGAAACGCCCTTGAGTTGAAGCGTTTGTG	
consensus-sensitive	4131	(3789)	YAAGTAAAGGATGGAGGAGARGGGAGGGTTAGGAGATGATCACCAGCCCTTTTAATGAAAGCAGCAGC	4200
resistant sequence		(4103)	TAAGTTAAAGGATGGAGGAGAGAGGGGTAAGGAGATGATCACCAGCCCTTTTAATGAAAGCAGCAGC	
			▼Retrotransposon Insertion	
consensus-sensitive		(3859)	AAT-----	4270
resistant sequence		(4173)	AATACACAAGTCCAATTACAATTATGTCTTCCCTCAATTGAAGTCATTGAGAAATAGAAAGATGCCCATTGC	
consensus-sensitive	4271	(3893)	TGACATTTATGCCGCTGTGTCCCAAGACAGAAAAACDGCATTTAGTTGTATTAAATGAAYGACTCCGGAT	4340
resistant sequence		(4243)	TGACATTTATGCCGCTGTGTCCCAAGACAGAAAAACTGCATTTAGTTGTATTAAATGAACGACTCCGGAT	
consensus-sensitive	4341	(3963)	AGTGCATCTAAAGGAGATGAGAATTTCTATGCTCCATTACATTCATCATCTGATCCTGAAACCCG	4410
resistant sequence		(4313)	AGTGCATCTAAAGGAGATGAGAATTTCTATGCTCCATTACATTCATCATCTGATCCTGAAACCCG	
consensus-sensitive	4411	(4033)	AGGAGCACTATTTCCCATTTCCCATGTTAAGAGAGGTATACATAAAACAATGTGGCA	4480
resistant sequence		(4383)	AGGAGCACTATTTCCCATTTCCCATGTTAAGAGAGGTATACATAAAACAATGTGGCA	

FIG 1 I

consensus-sensitive resistant sequence	(4103) (4453)	4481 CTATGGAGGCTTTTAGGTGTCCTCACTCATATGACAATAAAAAACGACAAAGGTAGAGAGTTTGGGAGAAAGT CTATGGAGGCTTTTAGGTGTCCTCACTCATATGACAATAAAAAACGACGAGGTAGAGAGTTTGGGAGAAAGT	4550
consensus-sensitive resistant sequence	(4173) (4523)	4551 TGGGAGGTGTTTCGGAGCTRCTCATCTTCTTTGCGATCCTTGAATATCACAGGTTGCTCCAACCTTAAGA TGGAGAGGTGTTTCGGAGCTGCTCATCTTCTTTGCGATCCTTGAATATCACAGGTTGCTCCAACCTTAAGA	4620
consensus-sensitive resistant sequence	(4243) (4593)	4621 AGTGTTTCTGGAGGGCTGGAGCATCTCACTRCCTTTGGAGATKTTAGAAATATACGACACCCATAAGCTGA AGTGTTTCTGGAGGGCTGGAGCATCTCACTGCTTTGGAGATGTTAGAAATATACGACACCCATAAGCTGA	4690
consensus-sensitive resistant sequence	(4313) (4663)	4691 GTCTWTCAGAAAGACCCAGAAAGGTGTTGTGCCATGGAAATCCCTTCATCACTCCCTCAGCTACTTGMAAAT GTCTATCAGAAAGACCCAGAAAGGTGTTGTGCCATGGAAATCCCTTCATCACTCCCTCAGCTACTTGCAAGT	4760
consensus-sensitive resistant sequence	(4383) (4733)	4761 GATGAATCTCCCWACAGCTGGTCAACCTGCCTGATTCGATGCAGTTCCTTGGYCCCTCCAAACCTTTCA GATGAATCTCCCAACAGCTGGTCAACCTGCCTGATTCGATGCAGTTCCTTGGCTCCCTCCAAACCTTTCA	4830
consensus-sensitive resistant sequence	(4453) (4803)	4831 ATGTTGCATTGCAGTAAACTGGAATCAGTGCCAGATTGGATGCCCMGACTCACTTCYCTCAGGAAGCTTA ATGTTGCATTGCAGTAAACTGGAATCAGTGCCAGATTGGATGCCCACTCACTTCCTCAGGAAGCTTA	4900
consensus-sensitive resistant sequence	(4523) (4873)	4901 TGGTTTCATTCTGTTCCGCACATCTGGAGAGAAGATGYCAAAATCCAACCTGGGGTGGACTGGCCTAACAT TGGTTTCATTCTGTTCCGCACATCTGGAGAGAAGATGCCAAATCCAACCTGGGGTGGACTGGCCTAACAT	4970
consensus-sensitive resistant sequence	(4593) (4943)	4971 TCAACACATCCCCCTSCATTGATGTCACCTCTAGCCCTCCTAAGTTTTTAGTGTGCCGTATGAATAG TCAACACATCCCCCTCCATTGATGTCACCTCTAGCCCTCCTAAGTTTTTAGTGTGCCGTATGAATAG	5037

FIG 2 A

1																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																											
---	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

FIG 2 B

Res. sequence	(71)	KDAVFEADDLDFEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYMRSGVKKIKKKLDAIAYNHQF	71	140
Sen. sequence1	(71)	KDAVFEADDLDFEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYMRSGVKKIKKKLDAIAYNHQF		
Sen. sequence2	(71)	KDAVFEADDLDFEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYMRSGVKKIKKKLDAIAYNHQF		
Sen. sequence3	(71)	KDAVFEADDLDFEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYMRSGVKKIKKKLDAIAYNHQF		
Sen. sequence4	(71)	KDAVFEADDLDFEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYMRSGVKKIKKKLDAIAYNHQF		
Sen. sequence5	(71)	KDAVFEADDLDFEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYMRSGVKKIKKKLDAIAYNHQF		
Sen. sequence6	(71)	KDAVFEADDLDFEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYMRSGVKKIKKKLDAIAYNHQF		
Sen. sequence7	(28)	KDAVFEADDLDFEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYMRSGVKKIKKKLDAIAYNHQF		
Sen. sequence8	(71)	KDAVFEADDLDFEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYMRSGVKKIKKKLDAIAYNHQF		
Sen. sequence9	(71)	KDAVFEADDLDFEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYMRSGVKKIKKKLDAIAYNHQF		
Sen. sequence10	(71)	KDAVFEADDLDFEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYMRSGVKKIKKKLDAIAYNHQF		
Sen. sequence11	(71)	KDAVFEADDLDFEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYMRSGVKKIKKKLDAIAYNHQF		
Sen. sequence12	(71)	KDAVFEADDLDFEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYMRSGVKKIKKKLDAIAYNHQF		
Sen. sequence13	(28)	KDAVFEADDLDFEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYMRSGVKKIKKKLDAIAYNHQF		
Sen. sequence14	(71)	KDAVFEADDLDFEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYMRSGVKKIKKKLDAIAYNHQF		
Sen. sequence15	(71)	KDAVFEADDLDFEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYMRSGVKKIKKKLDAIAYNHQF		
Sen. sequence16	(71)	KDAVFEADDLDFEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYMRSGVKKIKKKLDAIAYNHQF		
Sen. sequence17	(71)	KDAVFEADDLDFEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYMRSGVKKIKKKLDAIAYNHQF		
Sen. sequence18	(71)	KDAVFEADDLDFEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYMRSGVKKIKKKLDAIAYNHQF		
Sen. sequence19	(71)	KDAVFEADDLDFEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYMRSGVKKIKKKLDAIAYNHQF		
Sen. sequence20	(71)	KDAVFEADDLDFEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYMRSGVKKIKKKLDAIAYNHQF		
Sen. sequence21	(71)	KDAVFEADDLDFEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYKMSQGVKKIKKKLDVIAYNH		
Sen. sequence22	(71)	KDAVFEADDLDFEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYKMSQGVKKIKKKLDVIAYNH		

FIG 2 C

Res.	sequence	(141)	SFKIDLEP	210
Sen.	sequence1	(141)	IKERRLETG	
Sen.	sequence2	(141)	SVVNAGDI	
Sen.	sequence3	(141)	IGREDDLE	
Sen.	sequence4	(141)	KIVGLLLD	
Sen.	sequence5	(141)	SNIQRDVS	
Sen.	sequence6	(141)	SFLTIVGM	
Sen.	sequence7	(98)	GGLGKTAL	
Sen.	sequence8	(141)	AVYN	
Sen.	sequence9	(141)	SFKIDLEP	
Sen.	sequence10	(141)	IKERRLETG	
Sen.	sequence11	(141)	SVVNAGDI	
Sen.	sequence12	(141)	IGREDDLE	
Sen.	sequence13	(98)	KIVGLLLD	
Sen.	sequence14	(141)	SNIQRDVS	
Sen.	sequence15	(141)	SFLTIVGM	
Sen.	sequence16	(141)	GGLGKTAL	
Sen.	sequence17	(141)	AVYN	
Sen.	sequence18	(141)	SFKIDLEP	
Sen.	sequence19	(141)	IKERRLETG	
Sen.	sequence20	(141)	SVVNAGDI	
Sen.	sequence21	(141)	IGREDDLE	
Sen.	sequence22	(141)	KIVGLLLD	

FIG 2 D

Res. sequence	(211)	DPRVRTAFPLRCWNCVSDQDQKQLDVKEILGKIL	211	280
Sen. sequence1	(211)	DPRVRTAFPLRCWNCVSDQDQKQLDVKEILGKIL		
Sen. sequence2	(211)	DPRVRTAFPLRCWNCVSDQDQKQLDVKEILGKIL		
Sen. sequence3	(211)	DPRVRTAFPLRCWNCVSDQDQKQLDVKEILGKIL		
Sen. sequence4	(211)	DPRVRTAFPLRCWNCVSDQDQKQLDVKEILGKIL		
Sen. sequence5	(211)	DPRVRTAFPLRCWNCVSDQDQKQLDVKEILGKIL		
Sen. sequence6	(211)	DPRVRTAFPLRCWNCVSDQDQKQLDVKEILGKIL		
Sen. sequence7	(168)	DPRVRTAFPLRCWNCI*-----		
Sen. sequence8	(211)	DPRVRTAFPLRCWNCVSDQDQKQLDVKEILGKIL		
Sen. sequence9	(211)	DPRVRTAFPLRCWNCVSDQDQKQLDVKEILGKIL		
Sen. sequence10	(211)	DPRVRTAFPLRCWNCVSDQDQKQLDVKEILGKIL		
Sen. sequence11	(211)	DPRVRTAFPLRCWNCVSDQDQKQLDVKEILGKIL		
Sen. sequence12	(211)	DPRVRTAFPLRCWNCVSDQDQKQLDVKEILGKIL		
Sen. sequence13	(168)	DPRVRTAFPLRCWNCI*-----		
Sen. sequence14	(211)	DPRVRTAFPLRCWNCVSDQDQKQLDVKEILGKIL		
Sen. sequence15	(211)	DPRVRTAFPLRCWNCVSDQDQKQLDVKEILGKIL		
Sen. sequence16	(211)	DPRVRTAFPLRCWNCVSDQDQKQLDVKEILGKIL		
Sen. sequence17	(211)	DPRVRTAFPLRCWNCVSDQDQKQLDVKEILGKIL		
Sen. sequence18	(211)	DPRVRTAFPLRCWNCVSDQDQKQLDVKEILGKIL		
Sen. sequence19	(211)	DPRVRTAFPLRCWNCVSDQDQKQLDVKEILGKIL		
Sen. sequence20	(211)	DPRVRTAFPLRCWNCVSDQDQKQLDVKEILGKIL		
Sen. sequence21	(211)	DPRVRTAFPLRCWNCVSDQDQKQLDVKEILGKIL		
Sen. sequence22	(211)	DPRVRTAFPLRCWNCVSDQDQKQLDVKEILGKIL		

FIG 2 F

Res. sequence	(351)	NDFVMIARDIVDKCARNPLAIRVVGSLLCGQDKSKWLSFHEICLANIRKSHNDIMPIILNLSYHHLEPPIR	420
Sen. sequence1	(351)	NDFVMIARDIVDKCARNPLAIRVVGSLLCGQDKSKWLSFHEICLANIRKSHNDIMPIILNLSYHHLEPPIR	
Sen. sequence2	(351)	NDFVMIARDIVDKCARNPLAIRVVGSLLCGQDKSKWLSFHEICLANIRKSHNDIMPIILNLSYHHLEPPIR	
Sen. sequence3	(351)	NDFVMIARDIVDKCARNPLAIRVVGSLLCGQDKSKWLSFHEICLANIRKSHNDIMPIILNLSYHHLEPPIR	
Sen. sequence4	(351)	NDFVMIARDIVDKCARNPLAIRVVGSLLCGQDKSKWLSFHEICLANIRKSHNDIMPIILNLSYHHLEPPIR	
Sen. sequence5	(351)	NDFVMIARDIVDKCARNPLAIRVVGSLLCGQDKSKWLSFHEICLANIRKSHNDIMPIILNLSYHHLEPPIR	
Sen. sequence6	(351)	NDFVMIARDIVDKCARNPLAIRVVGSLLCGQDKSKWLSFHEICLANIRKSHNDIMPIILNLSYHHLEPPIR	
Sen. sequence8	(351)	NDFVMIARDIVDKCARNPLAIRVVGSLLCGQDKSKWLSFHEICLANIRKSHNDIMPIILNLSYHHLEPPIR	
Sen. sequence9	(351)	NDFVMIARDIVDKCARNPLAIRVVGSLLCGQDKSKWLSFHEICLANIRKSHNDIMPIILNLSYHHLEPPIR	
Sen. sequence10	(351)	NDFVMIARDIVDKCARNPLAIRVVGSLLCGQDKSKWLSFHEICLANIRKSHNDIMPIILNLSYHHLEPPIR	
Sen. sequence11	(351)	NDFVMIARDIVDKCARNPLAIRVVGSLLCGQDKSKWLSFHEICLANIRKSHNDIMPIILNLSYHHLEPPIR	
Sen. sequence12	(351)	NDFVMIARDIVDKCARNPLAIRVVGSLLCGQDKSKWLSFHEICLANIRKSHNDIMPIILNLSYHHLEPPIR	
Sen. sequence14	(351)	NDFVMIARDIVDKCARNPLAIRVVGSLLCGQDKSKWLSFHEICLANIRKSHNDIMPIILNLSYHHLEPPIR	
Sen. sequence15	(351)	NDFVMIARDIVDKCARNPLAIRVVGSLLCGQDKSKWLSFHEICLANIRKSHNDIMPIILNLSYHHLEPPIR	
Sen. sequence16	(351)	NDFVMIARDIVDKCARNPLAIRVVGSLLCGQDKSKWLSFHEICLANIRKSHNDIMPIILNLSYHHLEPPIR	
Sen. sequence17	(351)	NDFVMIARDIVDKCARNPLAIRVVGSLLCGQDKSKWLSFHEICLANIRKSHNDIMPIILNLSYHHLEPPIR	
Sen. sequence18	(350)	NDFVMIARDIVDKCARNPLAIRVVGSLLCGQDKSKWLSFHEICLANIRKSHNDIMPIILNLSYHHLEPPIR	
Sen. sequence19	(351)	NDFVMIARDIVDKCARNPLAIRVVGSLLCGQDKSKWLSFHEICLANIRKSHNDIMPIILNLSYHHLEPPIR	
Sen. sequence20	(351)	NDFVMIARDIVDKCARNPLAIRVVGSLLCGQDKSKWLSFHEICLANIRKSHNDIMPIILNLSYHHLEPPIR	
Sen. sequence21	(333)	XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX	
Sen. sequence22	(351)	XXXXXXXXXXXXCARNPLAIRVVGSLLCGQDKSKWLSFHEICLANIRKSHNDIMPIILNLSYHHLEPPIR	

FIG 2 G

Res. sequence	(421)	RCFSYCA	FFPKDFLIGK	Q	TLINLWMAQGYIVPLDKDQSIDDASEEYISILL	RCFFEN	GA	EKDGVIKIH	490
Sen. sequence1	(421)	RCFSYCA	FFPKDFLIGK	KTTLINLWMAQGYIVPLDKDQSIDDASEEYISILL	QRCFFEN	IGTEKDYVIKIH			
Sen. sequence2	(421)	RCFSYCA	FFPKDFLIGK	KTTLINLWMAQGYIVPLDKDQSIDDASEEYISILL	QRCFFEN	IGTEKDYVIKIH			
Sen. sequence3	(421)	RCFSYCA	FFPKDFLIGK	KTTLINLWMAQGYIVPLDKDQSIDDASEEYISILL	QRCFFEN	IGTEKDYVIKIH			
Sen. sequence4	(421)	RCFSYCA	FFPKDFLIGK	KTTLINLWMAQGYIVPLDKDQSIDDASEEYISILL	QRCFFEN	IGTEKDYVIKIH			
Sen. sequence5	(421)	RCFSYCA	FFPKDFLIGK	KTTLINLWMAQGYIVPLDKDQSIDDASEEYISILL	QRCFFEN	IGTEKDYVIKIH			
Sen. sequence6	(421)	RCFSYCA	FFPKDFLIGK	KTTLINLWMAQGYIVPLDKDQSIDDASEEYISILL	QRCFFEN	IGTEKDYVIKIH			
Sen. sequence8	(421)	RCFSYCA	FFPKDFLIGK	KTTLINLWMAQGYIVPLDKDQSIDDASEEYISILL	QRCFFEN	IGTEKDYVIKIH			
Sen. sequence9	(421)	RCFSYCA	FFPKDFLIGK	KTTLINLWMAQGYIVPLDKDQSIDDASEEYISILL	QRCFFEN	IGTEKDYVIKIH			
Sen. sequence10	(421)	RCFSYCA	FFPKDFLIGK	KTTLINLWMAQGYIVPLDKDQSIDDASEEYISILL	QRCFFEN	IGTEKDYVIKIH			
Sen. sequence11	(421)	RCFSYCA	FFPKDFLIGK	KTTLINLWMAQGYIVPLDKDQSIDDASEEYISILL	QRCFFEN	IGTEKDYVIKIH			
Sen. sequence12	(421)	RCFSYCA	FFPKDFLIGK	KTTLINLWMAQGYIVPLDKDQSIDDASEEYISILL	QRCFFEN	IGTEKDYVIKIH			
Sen. sequence14	(421)	RCFSYCA	FFPKDFLIGK	KTTLINLWMAQGYIVPLDKDQSIDDASEEYISILL	QRCFFEN	IGTEKDYVIKIH			
Sen. sequence15	(421)	RCFSYCA	FFPKDFLIGK	KTTLINLWMAQGYIVPLDKDQSIDDASEEYISILL	QRCFFEN	IGTEKDYVIKIH			
Sen. sequence16	(421)	RCFSYCA	FFPKDFLIGK	KTTLINLWMAQGYIVPLDKDQSIDDASEEYISILL	QRCFFEN	IGTEKDYVIKIH			
Sen. sequence17	(421)	RCFSYCA	FFPKDFLIGK	KTTLINLWMAQGYIVPLDKDQSIDDASEEYISILL	QRCFFEN	IGTEKDYVIKIH			
Sen. sequence18	(420)	RCFSYCV	VF	FPKDFLIGK	KTTLINLWMAQGYIVPLDKDQSIDDASEEYISILL	QRCFFEN	IGTEKDDVIKIH		
Sen. sequence19	(421)	RCFSYCA	FFPKDFLIGK	KTTLINLWMAQGYIVPLDKDQSIDDASEEYISILL	QRCFFEN	IGTEKDYVIKIH			
Sen. sequence20	(421)	RCFSYCA	FFPKDFLIGK	KTTLINLWMAQGYIVPLDKDQSIDDASEEYISILL	QRCFFEN	IGTEKDYVIKIH			
Sen. sequence21	(378)	XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX							
Sen. sequence22	(421)	RCFSYCA	VF	FPKDFLIGK	KTTLINLWMAQGYIVPLDKDQSIDDASEEYISIL	XQRCFFEN	IGTEKDDVIKIH		

FIG 2 H

Res. sequence	(491)	491	560
Sen. sequence1	(491)		
Sen. sequence2	(491)		
Sen. sequence3	(491)		
Sen. sequence4	(491)		
Sen. sequence5	(491)		
Sen. sequence6	(491)		
Sen. sequence8	(491)		
Sen. sequence9	(491)		
Sen. sequence10	(491)		
Sen. sequence11	(491)		
Sen. sequence12	(491)		
Sen. sequence14	(491)		
Sen. sequence15	(491)		
Sen. sequence16	(491)		
Sen. sequence17	(491)		
Sen. sequence18	(490)		
Sen. sequence19	(491)		
Sen. sequence20	(491)		
Sen. sequence21	(448)		
Sen. sequence22	(491)		

FIG 2 I

561																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														
-----	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

FIG 2 J

Res. sequence	(631)	PKHLSRLVKLQTLTDI	631	700
Sen. sequence1	(631)	PKHLSRLVKLQTLTDI		
Sen. sequence2	(631)	PKHLSRLVKLQTLTDI		
Sen. sequence3	(631)	PKHLSRLVKLQTLTDI		
Sen. sequence4	(631)	PKHLSRLVKLQTLTDI		
Sen. sequence5	(631)	PKHLSRLVKLQTLTDI		
Sen. sequence6	(631)	PKHLSRLVKLQTLTDI		
Sen. sequence8	(631)	PKHLSRLVKLQTLTDI		
Sen. sequence9	(631)	PKHLSRLVKLQTLTDI		
Sen. sequence10	(631)	PKHLSRLVKLQTLTDI		
Sen. sequence11	(631)	PKHLSRLVKLQTLTDI		
Sen. sequence12	(631)	PKHLSRLVKLQTLTDI		
Sen. sequence14	(631)	PKHLSRLVKLQTLTDI		
Sen. sequence15	(631)	PKHLSRLVKLQTLTDI		
Sen. sequence16	(631)	PKHLSRLVKLQTLTDI		
Sen. sequence17	(631)	PKHLSRLVKLQTLTDI		
Sen. sequence18	(630)	PKHLSRLVKLQTLTDI		
Sen. sequence19	(631)	PKHLSRLVKLQTLTDI		
Sen. sequence20	(631)	PKHLSRLVKLQTLTDI		
Sen. sequence21	(574)	PKHLSRLVKLQTLTDI		
Sen. sequence22	(599)	PKHLSRLVKLQTLTDI		

FIG 2 K

Res. sequence	(701)	LEIQIRWP	NTTDAVKEDV	REGLYLNHKEHLNHI	VDFRCEE	GGGRMDDEE	ARRLMEELR	PHPYLENLA	770
Sen. sequence1	(701)	LEIQIRWP	NTTDAVKEDV	KREGLYLNHKEHLNHI	VDFRCEE	GGGRMDDEE	ARRLMEELR	PHPYLENLA	
Sen. sequence2	(701)	LEIQIRWP	NTTDAVKEDV	KREGLYLNHKEHLNHI	VVD				
Sen. sequence3	(701)	LEIQIRWP	NTTDAVKEDV	KREGLYLNHKEHLNHI	VVD				
Sen. sequence4	(701)	LEIQIRWP	NTTDAVKEDV	KREGLYLNHKEHLNHI	VVD				
Sen. sequence5	(701)	LEIQIRWP	NTTDAVKEDV	KREGLYLNH					
Sen. sequence6	(701)	LEIQIRWP	NTTDAVKEDV	KREGLYN					
Sen. sequence8	(701)	LEIQIRWP	NTTDAVKEDV	KREGLYLNH					
Sen. sequence9	(701)	LEIQIRWP	NTTDAVKEDV	KREGLYLNH					
Sen. sequence10	(701)	LEIQIRWP	NTTDAVKEDV	KREGLYLNHKEHLNHI	VDFRCEE	GGGRMDDEE	ARRLMEELR	PHPYLENLA	
Sen. sequence11	(701)	LEIQIRWP	NTTDAVKEDV	KREGLYLNHKEHLNHI	VVD				
Sen. sequence12	(701)	LEIQIRWP	NTTDAVKEDV	KREGLYLNHKEHLNHI	VVD				
Sen. sequence14	(701)	LEIQIRWP	NTTDAVKEDV	KREGLYLNHKEHLNHI	VVD				
Sen. sequence15	(701)	LEIQIRWP	NTTDAVKEDV	KREGLYLNHKEHLNHI	VVD				
Sen. sequence16	(701)	LEIQIRWP	NTTDAVKEDV	KREGLYLNHKEHLNHI	VVD				
Sen. sequence17	(701)	LEIQIRWP	NTTDAVKEDV	KREGLYLNHKEHLNHI	VVD				
Sen. sequence18	(700)	LEIQIRWP	NTTDAVKEDV	KREGLYLNHKEHLNHI	VVD				
Sen. sequence19	(701)	LEIQIRWP	NTTDAVKEDV	KREGLYLNHKEHLNHI	VVD				
Sen. sequence20	(701)	LEIQIRWP	NTTDAVKEDV	KREGLYLNHKEHLNHI	VVD				
Sen. sequence21	(644)	LEIQIRWP	NTTDAVKEDV	KREGLYLNH					
Sen. sequence22	(669)	LEIQIRWP	NTTDAVKEDV	KREGLYLNHKEHLNHI	VVD				

FIG 2 L

Res. sequence	(771)	VKAYYGVMKPGWATLLPNLT	ELF	SDCGE	LENLPCLGNLDHLKVLRLSHLAKLEYIEEDSSSANFRCRPG	840
Sen. sequence1	(771)	VKAYYGAKMPGWATLLPNLT	ELY	SDCGE	SECLPCMGNLDCLKVLRLSHLAKLEYIEEDSTSANFSFRPG	
Sen. sequence10	(771)	VKAYYGAKMPGWATLLPNLT	ELY	SDCGE	SECLPCMGNLDCLKVLRLSHLAKLEYIEEDSTSANFSFRPG	
Res. sequence	(841)	PESAGLSLYFPSLERLELKRRLCKLKGWRRGEG	LGDDHQP	FNESS>Ret.	Ins<NTQVQLQLCLPQLKSLRIERCPLLT	910
Sen. sequence1	(841)	PESAGLSLYFPSLELLELKRRLHKLKGWRRREG	LGDDHQP	FNESS>Ret.	Ins<-----	
Sen. sequence10	(841)	PESAGLSLYFPSLELLELKRRLHKLKGWRRREG	LGDDHQP	FNESS>Ret.	Ins<-----	
Res. sequence	(911)	FMPLCPKTEKHLVVFNERLRIVHAKR	DENFYAPLHSSSDPENPRNTIPIPMFREVYINNVAWLNSLPM			980
Res. sequence	(981)	EAFRCLTHMTIKNDEVESLGEVGEVFRSCSSLSR	SINITGCSNLSRSVSGGLEH	LTALEMLEIYDTHKLSL		1050
Res. sequence	(1051)	SEDPEGVVPWKS	LHHSLSYLQLMNLPQLVNL	PDSMQFLAALRTL	SI	1120
Res. sequence	(1121)	SFCSAHLERRCQNPTGVDWPNIQHI	PSIDVTSSLPKFLVLPYE			1163

FIG 3 A

Res. sequence	1	1	70
Sen. sequence1	(1)	MERVVYRWNTAETAGHLWPTNP ~~KL~~ ILIPYSALQSSSELKEILSIFGYKS ~~RL~~ DDDLQRTVSTINAVFRDAETKQELTHEAQHWLEEL	
Sen. sequence2	(1)	MERVVYRWNTAETAGHLWPTNP ~~KL~~ ILIPY-ALQSSSELKEILSIFGYKSQ ~~LD~~ DLQRTVSTINAVFRDAETKQELTHEAQHWLEEL	
Sen. sequence3	(1)	MERVVYRWNTAETAGHLWPTNP ~~KL~~ ILIPYSALQSSSELKEILSIFGYKSQ ~~LD~~ DLQRTVSTINAVFRDAETKQELTHEAQHWLEEL	
Sen. sequence4	(1)	MERVVYRWNTAETAGHLWPTNP ~~KL~~ ILIPYSALQSSSELKEILSIFGYKSQ ~~LD~~ DLQRTVSTINAVFRDAETKQELTHEAQHWLEEL	
Sen. sequence5	(1)	MERVVYRWNTAETAGHLWPTNP ~~KL~~ ILIPYSALQSSSELKEILSIFGYKSQ ~~LD~~ DLQRTVSTINAVFRDAETKQELTHEAQHWLEEL	
Sen. sequence6	(1)	MERVVYRWNTAETAGHLWPTNP ~~KL~~ ILIPYSALQSSSELKEILSIFGYKSQ ~~LD~~ DLQRTVSTINAVFRDAETKQELTHEAQHWLEEL	
Sen. sequence7	(1)	-----STINAVFRDAETKQELTHEAQHWLEEL	
Sen. sequence8	(1)	MERVVYRWNTAETAGHLWPTNP ~~KL~~ ILIPYSALQSSSELKEILSIFGYKSQ ~~LD~~ DLQRTVSTINAVFRDAETKQELTHEAQHWLEEL	
Sen. sequence9	(1)	MERVVYRWNTAETAGHLWPTNP ~~KL~~ ILIPYSALQSSSELKEILSIFGYKSQ ~~LD~~ DLQRTVSTINAVFRDAETKQELTHEAQHWLEEL	
Sen. sequence10	(1)	MERVVYRWNTAETAGHLWPTNP ~~KL~~ ILIPY-ALQSSSELKEILSIFGYKSQ ~~LD~~ DLQRTVSTINAVFRDAETKQELTHEAQHWLEEL	
Sen. sequence11	(1)	MERVVYRWNTAETAGHLWPTNP ~~KL~~ ILIPYSALQSSSELKEILSIFGYKSQ ~~LD~~ DLQRTVSTINAVFRDAETKQELTHEAQHWLEEL	
Sen. sequence12	(1)	MERVVYRWNTAETAGHLWPTNP ~~KL~~ ILIPYSALQSSSELKEILSIFGYKSQ ~~LD~~ DLQRTVSTINAVFRDAETKQELTHEAQHWLEEL	
Sen. sequence13	(1)	-----STINAVFRDAETKQELTHEAQHWLEEL	
Sen. sequence14	(1)	MERVVYRWNTAETAGHLWPTNP ~~KL~~ ILIPYSALQSSSELKEILSIFGYKSQ ~~LD~~ DLQRTVSTINAVFRDAETKQELTHEAQHWLEEL	
Sen. sequence15	(1)	MERVVYRWNTAETAGHLWPTNP ~~KL~~ ILIPYSALQSSSELKEILSIFGYKSQ ~~LD~~ DLQRTVSTINAVFRDAETKQELTHEAQHWLEEL	
Sen. sequence16	(1)	MERVVYRWNTAETAGHLWPTNP ~~KL~~ ILIPYSALQSSSELKEILSIFGYKSQ ~~LD~~ DLQRTVSTINAVFRDAETKQELTHEAQHWLEEL	
Sen. sequence17	(1)	MERVVYRWNTAETAGHLWPTNP ~~KL~~ ILIPYSALQSSSELKEILSIFGYKSQ ~~LD~~ DLQRTVSTINAVFRDAETKQELTHEAQHWLEEL	
Sen. sequence18	(1)	MERVVYRWNTAETAGHLWPTNP ~~KL~~ ILIPYSALQSSSELKEILSIFGYKSQ ~~LD~~ DLQRTVSTINAVFRDAETKQELTHEAQHWLEEL	
Sen. sequence19	(1)	MERVVYRWNTAETAGHLWPTNP ~~KL~~ ILIPYSALQSSSELKEILSIFGYKSQ ~~LD~~ DLQRTVSTINAVFRDAETKQELTHEAQHWLEEL	
Sen. sequence20	(1)	MERVVYRWNTAETAGHLWPTNP ~~KL~~ ILIPYSALQSSSELKEILSIFGYKSQ ~~LD~~ DLQRTVSTINAVFRDAETKQELTHEAQHWLEEL	
Sen. sequence21	(1)	-----ALQSSSELKEILSIFGYKSQ ~~LD~~ DLQRTVSTINAVFRDAETKQELTHEAQHWLEEL	
Sen. sequence22	(1)	-----ALQSSSELKEILSIFGYKSQ ~~LD~~ DLQRTVSTINAVFRDAETKQELTHEAQHWLEEL	

FIG 3 B

Res. sequence	71	140
Sen. sequence1	(71)	KDAVFEADDLDFEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYMRSGVKKIKKKLDAIAYNHQF
Sen. sequence2	(71)	KDAVFEADDLDFEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYMRSGVKKIKKKLDAIAYNHQF
Sen. sequence3	(71)	KDAVFEADDLDFEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYMRSGVKKIKKKLDAIAYNHQF
Sen. sequence4	(71)	KDAVFEADDLDFEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYMRSGVKKIKKKLDAIAYNHQF
Sen. sequence5	(71)	KDAVFEADDLDFEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYMRSGVKKIKKKLDAIAYNHQF
Sen. sequence6	(71)	KDAVFEADDLDFEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYMRSGVKKIKKKLDAIAYNHQF
Sen. sequence7	(28)	KDAVFEADDLDFEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYMRSGVKKIKKKLDAIAYNHQF
Sen. sequence8	(71)	KDAVFEADDLDFEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYMRSGVKKIKKKLDAIAYNHQF
Sen. sequence9	(71)	KDAVFEADDLDFEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYMRSGVKKIKKKLDAIAYNHQF
Sen. sequence10	(71)	KDAVFEADDLDFEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYMRSGVKKIKKKLDAIAYNHQF
Sen. sequence11	(71)	KDAVFEADDLDFEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYMRSGVKKIKKKLDAIAYNHQF
Sen. sequence12	(71)	KDAVFEADDLDFEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYMRSGVKKIKKKLDAIAYNHQF
Sen. sequence13	(28)	KDAVFEADDLDFEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYMRSGVKKIKKKLDAIAYNHQF
Sen. sequence14	(71)	KDAVFEADDLDFEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYMRSGVKKIKKKLDAIAYNHQF
Sen. sequence15	(71)	KDAVFEADDLDFEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYMRSGVKKIKKKLDAIAYNHQF
Sen. sequence16	(71)	KDAVFEADDLDFEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYMRSGVKKIKKKLDAIAYNHQF
Sen. sequence17	(71)	KDAVFEADDLDFEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYMRSGVKKIKKKLDAIAYNHQF
Sen. sequence18	(71)	KDAVFEADDLDFEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYMRSGVKKIKKKLDAIAYNHQF
Sen. sequence19	(71)	KDAVFEADDLDFEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYMRSGVKKIKKKLDAIAYNHQF
Sen. sequence20	(71)	KDAVFEADDLDFEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYMRSGVKKIKKKLDAIAYNHQF
Sen. sequence21	(71)	KDAVFEADDLDFEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYKMSQGVKKIKKKLDVIAYNHQF
Sen. sequence22	(71)	KDAVFEADDLDFEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYKMSQGVKKIKKKLDVIAYNHQF

FIG 3 C

Res. sequence	(141)	SFKIDLEPIKERRLETSVNVNAGDIIIGREDDLEKIVGLLLDSNIQRDVSF	210
Sen. sequence1	(141)	SFKIDLEPIKERRLETSVNVNAGDIIIGREDDLEKIVGLLLDSNIQRDVSF	
Sen. sequence2	(141)	SFKIDLEPIKERRLETSVNVNAGDIIIGREDDLEKIVGLLLDSNIQRDVSF	
Sen. sequence3	(141)	SFKIDLEPIKERRLETSVNVNAGDIIIGREDDLEKIVGLLLDSNIQRDVSF	
Sen. sequence4	(141)	SFKIDLEPIKERRLETSVNVNAGDIIIGREDDLEKIVGLLLDSNIQRDVSF	
Sen. sequence5	(141)	SFKIDLEPIKERRLETSVNVNAGDIIIGREDDLEKIVGLLLDSNIQRDVSF	
Sen. sequence6	(141)	SFKIDLEPIKERRLETSVNVNAGDIIIGREDDLEKIVGLLLDSNIQRDVSF	
Sen. sequence7	(98)	SFKIDLEPIKERRLETSVNVNAGDIIIGREDDLEKIVGLLLDSNIQRDVSF	
Sen. sequence8	(141)	SFKIDLEPIKERRLETSVNVNAGDIIIGREDDLEKIVGLLLDSNIQRDVSF	
Sen. sequence9	(141)	SFKIDLEPIKERRLETSVNVNAGDIIIGREDDLEKIVGLLLDSNIQRDVSF	
Sen. sequence10	(141)	SFKIDLEPIKERRLETSVNVNAGDIIIGREDDLEKIVGLLLDSNIQRDVSF	
Sen. sequence11	(141)	SFKIDLEPIKERRLETSVNVNAGDIIIGREDDLEKIVGLLLDSNIQRDVSF	
Sen. sequence12	(141)	SFKIDLEPIKERRLETSVNVNAGDIIIGREDDLEKIVGLLLDSNIQRDVSF	
Sen. sequence13	(98)	SFKIDLEPIKERRLETSVNVNAGDIIIGREDDLEKIVGLLLDSNIQRDVSF	
Sen. sequence14	(141)	SFKIDLEPIKERRLETSVNVNAGDIIIGREDDLEKIVGLLLDSNIQRDVSF	
Sen. sequence15	(141)	SFKIDLEPIKERRLETSVNVNAGDIIIGREDDLEKIVGLLLDSNIQRDVSF	
Sen. sequence16	(141)	SFKIDLEPIKERRLETSVNVNAGDIIIGREDDLEKIVGLLLDSNIQRDVSF	
Sen. sequence17	(141)	SFKIDLEPIKERRLETSVNVNAGDIIIGREDDLEKIVGLLLDSNIQRDVSF	
Sen. sequence18	(141)	SFKIDLEPIKERRLETSVNVNAGDIIIGREDDLEKIVGLLLDSNIQRDVSF	
Sen. sequence19	(141)	SFKIDLEPIKERRLETSVNVNAGDIIIGREDDLEKIVGLLLDSNIQRDVSF	
Sen. sequence20	(141)	SFKIDLEPIKERRLETSVNVNAGDIIIGREDDLEKIVGLLLDSNIQRDVSF	
Sen. sequence21	(141)	SFKIDLEPIKERRLETSVNVNAGDIIIGREDDLEKIVGLFLDSNIQRDVSF	
Sen. sequence22	(141)	SFKIDLEPIKERRLETSVNVNAGDIIIGREDDLEKIVGLFLDSNIQRDVSF	

	211	280
Res.	(211)	DPRVRTAFPLRCWNCVSDQDQKQLDVKEILGKILATATGKNHEGSTMDQVQTOLREQLCGKRYLLVLDDV
Sen.	sequence1	(211) DPRVRTAFPLRCWNCVSDQDQKKLDVKEILGKILSTATGKNHEGSTMDHVQTOLREQLCGKRYLLVLDDV
Sen.	sequence2	(211) DPRVRTAFPLRCWNCVSDQDQKKLDVKEILGKILSTATGKNHEGSTMDHVQTOLQEQLCGKRYLLVLDDV
Sen.	sequence3	(211) DPRVRTAFPLRCWNCVSDQDQKKLDVKEILGKILSTATGKNHEGSTMDHVQTOLQEQLCGKRYLLVLDDV
Sen.	sequence4	(211) DPRVRTAFPLRCWNCVSDQDQKKLDVKEILGKILSTATGKNHEGSTMDHVQTOLQEQLCGKRYLLVLDDV
Sen.	sequence5	(211) DPRVRTAFPLRCWNCVSDQDQKKLDVKEILGKILSTATGKNHEGSTMDHVQTOLQEQLCGKRYLLVLDDV
Sen.	sequence6	(211) DPRVRTAFPLRCWNCVSDQDQKKLDVKEILGKILSTATGKNHEGSTMDHVQTOLQEQLCGKRYLLVLDDV
Sen.	sequence7	(168) DPRVRTAFPLRCWNCL*-----
Sen.	sequence8	(211) DPRVRTAFPLRCWNCVSDQDQKKLDVKEILGKILSTATGKNHEGSTMDHVQTOLQEQLCGKRYLLVLDDV
Sen.	sequence9	(211) DPRVRTAFPLRCWNCVSDQDQKKLDVKEILGKILSTATGKNHEGSTMDHVQTOLQEQLCGKRYLLVLDDV
Sen.	sequence10	(211) DPRVRTAFPLRCWNCVSDQDQKKLDVKEILGKILSTATGKNHEGSTMDHVQTOLREQLCGKRYLLVLDDV
Sen.	sequence11	(211) DPRVRTAFPLRCWNCVSDQDQKKLDVKEILGKILSTATGKNHEGSTMDHVQTOLREQLCGKRYLLVLDDV
Sen.	sequence12	(211) DPRVRTAFPLRCWNCVSDQDQKKLDVKEILGKILSTATGKNHEGSTMDHVQTOLREQLCGKRYLLVLDDV
Sen.	sequence13	(168) DPRVRTAFPLRCWNCL*-----
Sen.	sequence14	(211) DPRVRTAFPLRCWNCVSDQDQKKLDVKEILGKILSTATGKNHEGSTMDHVQTOLREQLCGKRYLLVLDDV
Sen.	sequence15	(211) DPRVRTAFPLRCWNCVSDQDQKKLDVKEILGKILSTATGKNHEGSTMDHVQTOLREQLCGKRYLLVLDDV
Sen.	sequence16	(211) DPRVRTAFPLRCWNCVSDQDQKKLDVKEILGKILSTATGKNHEGSTMDHVQTOLREQLCGKRYLLVLDDV
Sen.	sequence17	(211) DPRVRTAFPLRCWNCVSDQDQKKLDVKEILGKILSTATGKNHEGSTMDHVQTOLQEQLCGKRYLLVLDDV
Sen.	sequence18	(211) DPRVRTAFPLRCWNCVFDDQDKQLDVKEILGKILSTATGKNHEGSTMDQVQTOLRE-LCGKRYLLVLDDV
Sen.	sequence19	(211) DPRVRTAFPLRCWNCVSDQDQKKLDVKEILGKILSTATGKNHEGSTMDHVQTOLREQLCGKRYLLVLDDV
Sen.	sequence20	(211) DPRVRTAFPLRCWNCVSDQDQKKLDVKEILGKILSTATGKNHEGSTMDHVQTOLREQLCGKRYLLVLDDV
Sen.	sequence21	(211) DPRVRTAFPLRCWNCVSDQDNQNLDVKEILGKILSTATGKNHKGSTMDQVQTOLREQLCGKRYLLVLDDV
Sen.	sequence22	(211) DPRVRTAFPLRCWNCVSDQDNQNLDVKEILGKILSTATGKNHKGSTMDQVQTOLREQLCGKRYLLVLDDV

FIG 3 E

Res. sequence	(281)	WNENPNQLRDLVEFFMGGQ ~~RG~~ GNWIVVTTTRSHETARIIRDGPLHKLQGLSEENSWRLFVRWTFGVSQ ~~AK~~ PKFP	350
Sen. sequence1	(281)	WNENPNQLR ~~YL~~ VEFFMGGQ ~~RG~~ GNWIVVTTTRSHETARIIRDGPLHKLQGLSEENSWRLFVRWTFGVSQPKFP	
Sen. sequence2	(281)	WNENPNQLRXXXXFFMGGQ ~~RG~~ GNWIVVTTTRSHETTRIIRDGPLHKLQGLSEENSWRLFVRWTFGVSQPKFP	
Sen. sequence3	(281)	WNENPNQLRDLGKFFMGGQ ~~RG~~ GNWIVVTTTRSHETTRIIRDGPLHKLQGLSEENSWRLFVRWTFGVSQPKFP	
Sen. sequence4	(281)	WNENPNQLRDXVEFFMGGQ ~~RG~~ GNWIVVTTTRSHETTRIIRDGPLHKLQGLSEENSWRLFVRWTFGVSQPKFP	
Sen. sequence5	(281)	WNENPNQLXXXXFFMGGQ ~~RG~~ GNWIVVTTTRSHETTRIIRDGPLHKLQGLSEENSWRLFVRWTFGVSQPKFP	
Sen. sequence6	(281)	WNENPNQLRIXVEFFMGGQ ~~RG~~ GNWIVVTTTRSHETARIIRDGPLHKLQGLSEENSWRLFVRWTFGVSQPKFP	
Sen. sequence7	(184)	-----	
Sen. sequence8	(281)	WNENPNQLRXXXXFFMGGQ ~~RG~~ GNWIVVTTTRSHETTRIIRDGPLHKLQGLSEENSWRLFVRWTFGVSQPKFP	
Sen. sequence9	(281)	WNENPNQLRDXVEFFMGGQ ~~RG~~ GNWIVVTTTRSHETTRIIRDGPLHKLQGLSEENSWRLFVRWTFGVSQPKFP	
Sen. sequence10	(281)	WNENPNQLR ~~YL~~ VEFFMGGQ ~~RG~~ GNWIVVTTTRSHETARIIRDGPLHKLQGLSEENSWRLFVRWTFGVSQPKFP	
Sen. sequence11	(281)	WNENPNQLR ~~YL~~ VEFFMGGQ ~~RG~~ GNWIVVTTTRSHETARIIRDGPLHKLQGLSEENSWRLFVRWTFGVSQPKFP	
Sen. sequence12	(281)	WNENPNQLXXXXFFMGGQ ~~RG~~ GNWIVVTTTRSHETARIIRDGPLHKLQGLSEENSWRLFVRWTFGVSQPKFP	
Sen. sequence14	(281)	WNENPNQLRILVEFFMGGQ ~~RG~~ GNWIVVTTTRSHETARIIRDGPLHKLQGLSEENSWRLFVRWTFGVSQPKFP	
Sen. sequence15	(281)	WNENPNQLRXXXXFFMGGQ ~~RG~~ GNWIVVTTTRSHETARIIRDGPLHKLQGLSEENSWRLFVRWTFGVSQPKFP	
Sen. sequence16	(281)	WNENPNQLRXXVEFFMGGQ ~~RG~~ GNWIVVTTTRSHETARIIRDGPLHKLQGLSEENSWRLFVRWTFGVSQPKFP	
Sen. sequence17	(281)	WNENPNQLRDLVEFFMGGQ ~~RG~~ GNWIVVTTTRSHETTRIIRDGPLHKLQGLSEENSWRLFVRWTFGVSQPKFP	
Sen. sequence18	(280)	WNENPNQLRDLVEFFMGGQ ~~RG~~ GNWIVVTTTRSHETARIIRDGPLHKLQGLSEENSWRLFVRWTFGVSQPKFP	
Sen. sequence19	(281)	WNENPNQLRILVEFFMXXXXXXXQGLSEENSWRLFVRWTFGVSQPKFP	
Sen. sequence20	(281)	WNENPNQLRDLVEFFMGGQ ~~RG~~ GNWIVVTTTRSHETARIIRDGPLHKLQGLSEENSWRLFVRWTFGVSQPKFP	
Sen. sequence21	(281)	WNENPNQLRDLVEFFMGXXXXXXXQGLSEENSWRLFVRWTFGVSQPKFP	
Sen. sequence22	(281)	WNENPNQLRXXXXXXXQGLSEENSWRLFVRWTFGVSQPKFP	

FIG 3 F

Res. sequence	(351)	NDFFMIARDIVDKCARNPLAIRVVGSLLCGQDKSKWLSFHEICLGNIRKSHNDIMPIILNSYHHLEPPIR	420
Sen. sequence1	(351)	NDFVMIARDIVDKCARNPLAIRVVGSLLCGQDKSKWLSFHEICLANIRKSHNDIMPIILNSYHHLEPPIR	
Sen. sequence2	(351)	NDFVMIARDIVDKCARNPLAIRVVGSLLCGQDKSKWLSFHEICLANIRKSHNDIMLILNSYHHLEPPIR	
Sen. sequence3	(351)	NDFVMIARDIVDKCARNPLAIRVVGSLLCGQDKSKWLSFHEICLANIRKSHNDIMLILNSYHHLEPPIR	
Sen. sequence4	(351)	NDFVMIARDIVDKCARNPLAIRVVGSLLCGQDKSKWLSFHEICLANIRKSHNDIMLILNSYHHLEPPIR	
Sen. sequence5	(351)	NDFVMIARDIVDKCARNPLAIRVVGSLLCGQDKSKWLSFHEICLANIRKSHNDIMLILNSYHHLEPPIR	
Sen. sequence6	(351)	NDFVMIARDIVDKCARNPLAIRVVGSLLCGQDKSKWLSFHEICLANIRKSHNDIMPIILNSYHHLEPPIR	
Sen. sequence8	(351)	NDFVMIARDIVDKCARNPLAIRVVGSLLCGQDKSKWLSFHEICLANIRKSHNDIMLILNSYHHLEPPIR	
Sen. sequence9	(351)	NDFVMIARDIVDKCARNPLAIRVVGSLLCGQDKSKWLSFHEICLANIRKSHNDIMLILNSYHHLEPPIR	
Sen. sequence10	(351)	NDFVMIARDIVDKCARNPLAIRVVGSLLCGQDKSKWLSFHEICLANIRKSHNDIMPIILNSYHHLEPPIR	
Sen. sequence11	(351)	NDFVMIARDIVDKCARNPLAIRVVGSLLCGQDKSKWLSFHEICLANIRKSHNDIMPIILNSYHHLEPPIR	
Sen. sequence12	(351)	NDFVMIARDIVDKCARNPLAIRVVGSLLCGQDKSKWLSFHEICLANIRKSHNDIMPIILNSYHHLEPPIR	
Sen. sequence14	(351)	NDFVMIARDIVDKCARNPLAIRVVGSLLCGQDKSKWLSFHEICLANIRKSHNDIMPIILNSYHHLEPPIR	
Sen. sequence15	(351)	NDFVMIARDIVDKCARNPLAIRVVGSLLCGQDKSKWLSFHEICLANIRKSHNDIMPIILNSYHHLEPPIR	
Sen. sequence16	(351)	NDFVMIARDIVDKCARNPLAIRVVGSLLCGQDKSKWLSFHEICLANIRKSHNDIMPIILNSYHHLEPPIR	
Sen. sequence17	(351)	NDFVMIARDIVDKCARNPLAIRVVGSLLCGQDKSKWLSFHEICLANIRKSHNDIMLILNSYHHLEPPIR	
Sen. sequence18	(350)	NDFVMIARDIVDKCARNPLAIRVVGSLLCGQDKSKWLSFHEICLANIRKSHNDIMPIILNSYHHLEPPIR	
Sen. sequence19	(351)	NDFVMIARDIVDKCARNPLAIRVVGSLLCGQDKSKWLSFHEICLANIRKSHNDIMPIILNSYHHLEPPIR	
Sen. sequence20	(351)	NDFVMIARDIVDKCARNPLAIRVVGSLLCGQDKSKWLSFHEICLANIRKSHNDIMPIILNSYHHLEPPIR	
Sen. sequence21	(333)	XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX	
Sen. sequence22	(351)	XXXXXXXXXXXXKARNPLAIRVVGSLLCGQDKSKWLSFHEXCLANIRKSHNDIMPIILNSYHHLEPPIR	

FIG 3 G

Res. sequence	(421;	RCFSYCA	FFPKDFLIGK	TLINLWMAQGYIVPLDKDQSIDDASEEYISILL	RCFFEN	GA	EKDDGVIKIH	490
Sen. sequence1	(421;	RCFSYCA	FFPKDFLIGKKTTLINLWMAQGYIVPLDKDQSIDDASEEYISILL	QRCFFEN	IGTEKDYVIKIH			
Sen. sequence2	(421;	RCFSYCA	FFPKDFLIGKKTTLINLWMAQGYIVPLDKDQSIDDASEEYISILL	QRCFFEN	IGTEKDYVIKIH			
Sen. sequence3	(421;	RCFSYCA	FFPKDFLIGKKTTLINLWMAQGYIVPLDKDQSIDDASEEYISILL	QRCFFEN	IGTEKDYVIKIH			
Sen. sequence4	(421;	RCFSYCA	FFPKDFLIGKKTTLINLWMAQGYIVPLDKDQSIDDASEEYISILL	QRCFFEN	IGTEKDYVIKIH			
Sen. sequence5	(421;	RCFSYCA	FFPKDFLIGKKTTLINLWMAQGYIVPLDKDQSIDDASEEYISILL	QRCFFEN	IGTEKDYVIKIH			
Sen. sequence6	(421;	RCFSYCA	FFPKDFLIGKKTTLINLWMAQGYIVPLDKDQSIDDASEEYISILL	QRCFFEN	IGTEKDYVIKIH			
Sen. sequence8	(421;	RCFSYCA	FFPKDFLIGKKTTLINLWMAQGYIVPLDKDQSIDDASEEYISILL	QRCFFEN	IGTEKDYVIKIH			
Sen. sequence9	(421;	RCFSYCA	FFPKDFLIGKKTTLINLWMAQGYIVPLDKDQSIDDASEEYISILL	QRCFFEN	IGTEKDYVIKIH			
Sen. sequence10	(421;	RCFSYCA	FFPKDFLIGKKTTLINLWMAQGYIVPLDKDQSIDDASEEYISILL	QRCFFEN	IGTEKDYVIKIH			
Sen. sequence11	(421;	RCFSYCA	FFPKDFLIGKKTTLINLWMAQGYIVPLDKDQSIDDASEEYISILL	QRCFFEN	IGTEKDYVIKIH			
Sen. sequence12	(421;	RCFSYCA	FFPKDFLIGKKTTLINLWMAQGYIVPLDKDQSIDDASEEYISILL	QRCFFEN	IGTEKDYVIKIH			
Sen. sequence14	(421;	RCFSYCA	FFPKDFLIGKKTTLINLWMAQGYIVPLDKDQSIDDASEEYISILL	QRCFFEN	IGTEKDYVIKIH			
Sen. sequence15	(421;	RCFSYCA	FFPKDFLIGKKTTLINLWMAQGYIVPLDKDQSIDDASEEYISILL	QRCFFEN	IGTEKDYVIKIH			
Sen. sequence16	(421;	RCFSYCA	FFPKDFLIGKKTTLINLWMAQGYIVPLDKDQSIDDASEEYISILL	QRCFFEN	IGTEKDYVIKIH			
Sen. sequence17	(421;	RCFSYCA	FFPKDFLIGKKTTLINLWMAQGYIVPLDKDQSIDDASEEYISILL	QRCFFEN	IGTEKDYVIKIH			
Sen. sequence18	(420;	RCFSYCV	VFPKDFLIGKKTTLINLWMAQGYIVPLDKDQSIDDASEEYISILL	QRCFFEN	IGTEKDDVIKIH			
Sen. sequence19	(421;	RCFSYCA	FFPKDFLIGKKTTLINLWMAQGYIVPLDKDQSIDDASEEYISILL	QRCFFEN	IGTEKDYVIKIH			
Sen. sequence20	(421;	RCFSYCA	FFPKDFLIGKKTTLINLWMAQGYIVPLDKDQSIDDASEEYISILL	QRCFFEN	IGTEKDYVIKIH			
Sen. sequence21	(378;	XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX						
Sen. sequence22	(421;	RCFSYCA	VFPKDFLIGKKTTLINLWMAQGYIVPLDKDQSIDDASEEYISILL	QRCFFEN	IGTEKDDVIKIH			

FIG 3H

	Res.	(491)	DLMH	DI	AQ	NV	MG	KELC	T	T	K	N	I	S	G	S	L	D	K	S	V	R	H	L	S	L	A	R	T	S	F	A	R	Y	S	F	N	A	T	H	I	R	S	<u>YF</u>	C	A	G	Y	W	C	D	A	E	I	N	Q	F	S	V	E
	Sen.	sequence1	(491)	DLMH	DI	AQ	NV	MG	KELC	T	T	K	N	I	S	G	S	L	D	K	N	V	R	H	L	S	L	A	R	T	S	F	A	R	Y	S	F	N	A	T	H	I	R	S	<u>YF</u>	A	G	Y	W	C	D	A	E	I	N	Q	F	S	V	E
	Sen.	sequence2	(491)	DLMH	DI	AQ	NV	MG	KELC	T	T	K	N	I	S	G	S	L	D	K	N	V	R	H	L	S	L	A	R	T	S	F	A	R	Y	S	F	N	A	T	H	I	R	S	<u>YF</u>	A	G	Y	W	C	D	A	E	I	X	Q	F	S	V	E
	Sen.	sequence3	(491)	DLMH	DI	AQ	NV	MG	KELC	T	T	K	N	I	S	G	S	L	D	K	N	V	R	H	L	S	L	A	R	T	S	F	A	R	Y	S	F	N	A	T	H	I	R	S	<u>YF</u>	A	G	Y	W	C	D	A	E	I	X	Q	F	S	V	E
	Sen.	sequence4	(491)	DLMH	DI	AQ	NV	MG	KELC	T	T	K	N	I	S	G	S	L	D	K	N	V	R	H	L	S	L	A	R	T	S	F	A	R	Y	S	F	N	A	T	H	I	R	S	<u>YF</u>	A	G	Y	W	C	D	A	E	I	N	Q	F	S	V	E
	Sen.	sequence5	(491)	DLMH	DI	AQ	NV	MG	KELC	T	T	K	N	I	S	G	S	L	D	K	N	V	R	H	L	S	L	A	R	T	S	F	A	R	Y	S	F	N	A	T	H	I	R	S	<u>YF</u>	A	G	Y	W	C	D	A	E	I	N	Q	F	S	V	E
	Sen.	sequence6	(491)	DLMH	DI	AQ	NV	MG	KELC	T	T	K	N	I	S	G	S	L	D	K	N	V	R	H	L	S	L	A	R	T	S	F	A	R	Y	S	F	N	A	T	H	I	R	S	<u>YF</u>	A	G	Y	W	C	D	A	E	I	N	Q	F	S	V	E
	Sen.	sequence8	(491)	DLMH	DI	AQ	NV	MG	KELC	T	T	K	N	I	S	G	S	L	D	K	N	V	R	H	L	S	L	A	R	T	S	F	A	R	Y	S	F	N	A	T	H	I	R	S	<u>YF</u>	A	G	Y	W	C	D	A	E	I	N	Q	F	S	V	E
	Sen.	sequence9	(491)	DLMH	DI	AQ	NV	MG	KELC	T	T	K	N	I	S	G	S	L	D	K	N	V	R	H	L	S	L	A	R	T	S	F	A	R	Y	S	F	N	A	T	H	I	R	S	<u>YF</u>	A	G	Y	W	C	D	A	E	I	N	Q	F	S	V	E
	Sen.	sequence10	(491)	DLMH	DI	AQ	NV	MG	KELC	T	T	K	N	I	S	G	S	L	D	K	N	V	R	H	L	S	L	A	R	T	S	F	A	R	Y	S	F	N	A	T	H	I	R	S	<u>YF</u>	A	G	Y	W	C	D	A	E	I	N	Q	F	S	V	E
	Sen.	sequence11	(491)	DLMH	DI	AQ	NV	MG	KELC	T	T	K	N	I	S	G	S	L	D	K	N	V	R	H	L	S	L	A	R	T	S	F	A	R	Y	S	F	N	A	T	H	I	R	S	<u>YF</u>	A	G	Y	W	C	D	A	E	I	N	Q	F	S	V	E
	Sen.	sequence12	(491)	DLMH	DI	AQ	NV	MG	KELC	T	T	K	N	I	S	G	S	L	D	K	N	V	R	H	L	S	L	A	R	T	S	F	A	R	Y	S	F	N	A	T	H	I	R	S	<u>YF</u>	A	G	Y	W	C	D	A	E	I	N	Q	F	S	V	E
	Sen.	sequence14	(491)	DLMH	DI	AQ	NV	MG	KELC	T	T	K	N	I	S	G	S	L</																																										

FIG 3 I

Res.	sequence	(561)	ALVPNCCLRALDLAWSKI	SLPDS	IGGLHLRLYLD	SYNED	LEVLPNS	IAKLYNLQ	TLQLKGCKRLEGL	630
Sen.	sequence1	(561)	ALVPNCCLRALDLAWSKI	SVPPDS	IGGLHLRLYLD	SYNED	LEVLPNS	IAKLYNLQ	TLQLKGCKRLEGL	
Sen.	sequence2	(561)	ALVPNCCLRALDLAWSKI	SVPPDS	IGGLHLRLYLD	SYNED	LEVLPNS	IAKLYNLQ	TLQLKGCKRLEGL	
Sen.	sequence3	(561)	ALVPNCCLRALDLAWSKI	SVPPDS	IGGLHLRLYLD	SYNED	LEVLPNS	IAKLYNLQ	TLQLKGCKRLEGL	
Sen.	sequence4	(561)	ALVPNCCLRALDLAWSKI	SVPPDS	IGGLHLRLYLD	SYNED	LEVLPNS	IAKLYNLQ	TLQLKGCKRLEGL	
Sen.	sequence5	(561)	ALVPNCCLRALDLAWSKI	SVPPDS	IGGLHLRLYLD	SYNED	LEVLPNS	IAKLYNLQ	TLQLKGCKRLEGL	
Sen.	sequence6	(561)	ALVPNCCLRALDLAWSKI	SVPPDS	IGGLHLRLYLD	SYNED	LEVLPNS	IAKLYNLQ	TLQLKGCKRLEGL	
Sen.	sequence8	(561)	ALVPNCCLRALDLAWSKI	SVPPDS	IGGLHLRLYLD	SYNED	LEVLPNS	IAKLYNLQ	TLQLKGCKRLEGL	
Sen.	sequence9	(561)	ALVPNCCLRALDLAWSKI	SVPPDS	IGGLHLRLYLD	SYNED	LEVLPNS	IAKLYNLQ	TLQLKGCKRLEGL	
Sen.	sequence10	(561)	ALVPNCCLRALDLAWSKI	SVPPDS	IGGLHLRLYLD	SYNED	LEVLPNS	IAKLYNLQ	TLQLKGCKRLEGL	
Sen.	sequence11	(561)	ALVPNCCLRALDLAWSKI	SVPPDS	IGGLHLRLYLD	SYNED	LEVLPNS	IAKLYNLQ	TLQLKGCKRLEGL	
Sen.	sequence12	(561)	ALVPNCCLRALDLAWSKI	SVPPDS	IGGLHLRLYLD	SYNED	LEVLPNS	IAKLYNLQ	TLQLKGCKRLEGL	
Sen.	sequence14	(561)	ALVPNCCLRALDLAWSKI	SVPPDS	IGGLHLRLYLD	SYNED	LEVLPNS	IAKLYNLQ	TLQLKGCKRLEGL	
Sen.	sequence15	(561)	ALVPNCCLRALDLAWSKI	SVPPDS	IGGLHLRLYLD	SYNED	LEVLPNS	IAKLYNLQ	TLQLKGCKRLEGL	
Sen.	sequence16	(561)	ALVPNCCLRALDLAWSKI	SVPPDS	IGGLHLRLYLD	SYNED	LEVLPNS	IAKLYNLQ	TLQLKGCKRLEGL	
Sen.	sequence17	(561)	ALVPNCCLRALDLAWSKI	SVPPDS	IGGLHLRLYLD	SYNED	LEVLPNS	IAKLYNLQ	TLQLKGCKRLEGL	
Sen.	sequence18	(560)	ALVPNCCLRALDLAWSKI	SVPPDS	IGGLHLRLYLD	SYNED	LEVLPNS	IAKLYNLQ	TLQLKGCKRLEGL	
Sen.	sequence19	(561)	ALVPNCCLRALDLAWSKI	SVPPDS	IGGLHLRLYLD	SYNED	LEVLPNS	IAKLYNLQ	TLQLKGCKRLEGL	
Sen.	sequence20	(561)	ALVPNCCLRALDLAWSKI	SVPPDS	IGGLHLRLYLD	SYNED	LEVLPNS	IAKLYNLQ	TLQLKGCKRLEGL	
Sen.	sequence21	(504)	ALVPNCCLRALDLAWSKI	SLPDS	IGGLHLRLYLD	SYNED	LEVLPNS	IAKLYNLQ	TLQLKGCKRLEGL	
Sen.	sequence22	(533)	XXXXNCLRALDLAWSKI	SLPDS	IGGLHLRLYLD	SYNED	LEVLPNS	IAKLYNLQ	TLQLKGCKRLEGL	

FIG 3 J

Res.	sequence	631	700
Sen.	sequence1	(631)	PKHLSRLVKLQTLTDIYGCNNVTYMPKGMGKMTCLHTLSKFI VGGEGSCSSWKQWFDGLEDLKALNNLKGH
Sen.	sequence2	(631)	PKHLSRLVKLQTLTDIYGCNNVTYMPKGMGKMTCLHTLSKFI VGGEGSCSSWKQWFDGQEDLKALNNLKGH
Sen.	sequence3	(631)	PKHLSRLVKLQTLTDIYGCNNVTYMPKGMGKMTCLHTLSKFI VGGEGSCSSWKQWFDGQEDLKALNNLKGH
Sen.	sequence4	(631)	PKHLSRLVKLQTLTDIYGCNNVTYMPKGMGKMTCLHTLSKFI VGGEGSCSSWKQWFDGQEDLKALNNLKGH
Sen.	sequence5	(631)	PKHLSRLVKLQTLTDIYGCNNVTYMPKGMGKMTCLHTLSKFI VGGEGSCSSWKQWFDGQEDLKALNNLKGH
Sen.	sequence6	(631)	PKHLSRLVKLQTLTDIYGCNNVTYMPKGMGKMTCLHTLSKFI VGGEGSCSSWKQWFDGQEDLKALNNLKGH
Sen.	sequence8	(631)	PKHLSRLVKLQTLTDIYGCNNVTYMPKGMGKMTCLHTLSKFI VGGEGSCSSWKQWFDGQEDLKALNNLKGH
Sen.	sequence9	(631)	PKHLSRLVKLQTLTDIYGCNNVTYMPKGMGKMTCLHTLSKFI VGGEGSCSSWKQWFDGQEDLKALNNLKGH
Sen.	sequence10	(631)	PKHLSRLVKLQTLTDIYGCNNVTYMPKGMGKMTCLHTLSKFI VGGEGSCSSWKQWFDGQEDLKALNNLKGH
Sen.	sequence11	(631)	PKHLSRLVKLQTLTDIYGCNNVTYMPKGMGKMTCLHTLSKFI VGGEGSCSSWKQWFDGQEDLKALNNLKGH
Sen.	sequence12	(631)	PKHLSRLVKLQTLTDIYGCNNVTYMPKGMGKMTCLHTLSKFI VGGEGSCSSWKQWFDGQEDLKALNNLKGH
Sen.	sequence14	(631)	PKHLSRLVKLQTLTDIYGCNNVTYMPKGMGKMTCLHTLSKFI VGGEGSCSSWKQWFDGQEDLKALNNLKGH
Sen.	sequence15	(631)	PKHLSRLVKLQTLTDIYGCNNVTYMPKGMGKMTCLHTLSKFI VGGEGSCSSWKQWFDGQEDLKALNNLKGH
Sen.	sequence16	(631)	PKHLSRLVKLQTLTDIYGCNNVTYMPKGMGKMTCLHTLSKFI VGGEGSCSSWKQWFDGQEDLKALNNLKGH
Sen.	sequence17	(631)	PKHLSRLVKLQTLTDIYGCNNVTYMPKGMGKMTCLHTLSKFI VGGEGSCSSWKQWFDGQEDLKALNNLKGH
Sen.	sequence18	(630)	PKHLSRLVKLQTLTDIYGCNNVTYMPKGMGKMTCLHTLSKFI VGGEGSCSSWKQWFDGQEDLKALNNLKGH
Sen.	sequence19	(631)	PKHLSRLVKLQTLTDIYGCNNVTYMPKGMGKMTCLHTLSKFI VGGEGSCSSWKQWFDGQEDLKALNNLKGH
Sen.	sequence20	(631)	PKHLSRLVKLQTLTDIYGCNNVTYMPKGMGKMTCLHTLSKFI VGGEGSCSSWKQWFDGQEDLKALNNLKGH
Sen.	sequence21	(574)	PKHLSRLVKLQTLTDIYGCNNVTYMPKGMGKMTCLHTLSKFI VGGEGSCSSWKKRFDGLEDLKALNNLKGH
Sen.	sequence22	(599)	PKHLSRLVKLQTLTDIYGCNNVTYMPKGMGKMTCLHTLSKFI VGGEGSCSSWKKRFDGLEDLKALNNLKGH

FIG 3 K

Res. sequence	(701)	LEIQIRWPENTTDAVKEDV	REGLYLNHKEHLNHI	VDFRCEEGGGRMDDEEARRLMEELRPHPYLENLA	770
Sen. sequence1	(701)	LEIQIRWPENTTDAVKEDV	KREGLYLNHKEHLNHI	VDFRCEEGGGRMDDEEARRLMEELRPHPYLENLA	
Sen. sequence2	(701)	LEIQIRWPENTTDAVKEDV	KREGLYLNHKEHLNHI	VDFRCEEGGGRMDDEEARRLMEELRPHPYLENLA	
Sen. sequence3	(701)	LEIQIRWPENTTDAVKEDV	KREGLYLNHKEHLNHI	VDFRCEEGGGRMDDEEARRLMEELRPHPYLENLA	
Sen. sequence4	(701)	LEIQIRWPENTTDAVKEDV	KREGLYLNHKEHLNHI	VDFRCEEGGGRMDDEEARRLMEELRPHPYLENLA	
Sen. sequence5	(701)	LEIQIRWPENTTDAVKEDV	KREGLYLNHKEHLNHI	VDFRCEEGGGRMDDEEARRLMEELRPHPYLENLA	
Sen. sequence6	(701)	LEIQIRWPENTTDAVKEDV	KREGLYLNHKEHLNHI	VDFRCEEGGGRMDDEEARRLMEELRPHPYLENLA	
Sen. sequence8	(701)	LEIQIRWPENTTDAVKEDV	KREGLYLNHKEHLNHI	VDFRCEEGGGRMDDEEARRLMEELRPHPYLENLA	
Sen. sequence9	(701)	LEIQIRWPENTTDAVKEDV	KREGLYLNHKEHLNHI	VDFRCEEGGGRMDDEEARRLMEELRPHPYLENLA	
Sen. sequence10	(701)	LEIQIRWPENTTDAVKEDV	KREGLYLNHKEHLNHI	VDFRCEEGGGRMDDEEARRLMEELRPHPYLENLA	
Sen. sequence11	(701)	LEIQIRWPENTTDAVKEDV	KREGLYLNHKEHLNHI	VDFRCEEGGGRMDDEEARRLMEELRPHPYLENLA	
Sen. sequence12	(701)	LEIQIRWPENTTDAVKEDV	KREGLYLNHKEHLNHI	VDFRCEEGGGRMDDEEARRLMEELRPHPYLENLA	
Sen. sequence14	(701)	LEIQIRWPENTTDAVKEDV	KREGLYLNHKEHLNHI	VDFRCEEGGGRMDDEEARRLMEELRPHPYLENLA	
Sen. sequence15	(701)	LEIQIRWPENTTDAVKEDV	KREGLYLNHKEHLNHI	VDFRCEEGGGRMDDEEARRLMEELRPHPYLENLA	
Sen. sequence16	(701)	LEIQIRWPENTTDAVKEDV	KREGLYLNHKEHLNHI	VDFRCEEGGGRMDDEEARRLMEELRPHPYLENLA	
Sen. sequence17	(701)	LEIQIRWPENTTDAVKEDV	KREGLYLNHKEHLNHI	VDFRCEEGGGRMDDEEARRLMEELRPHPYLENLA	
Sen. sequence18	(700)	LEIQIRWPENTTDAVKEDV	KREGLYLNHKEHLNHI	VDFRCEEGGGRMDDEEARRLMEELRPHPYLENLA	
Sen. sequence19	(701)	LEIQIRWPENTTDAVKEDV	KREGLYLNHKEHLNHI	VDFRCEEGGGRMDDEEARRLMEELRPHPYLENLA	
Sen. sequence20	(701)	LEIQIRWPENTTDAVKEDV	KREGLYLNHKEHLNHI	VDFRCEEGGGRMDDEEARRLMEELRPHPYLENLA	
Sen. sequence21	(644)	LEIQIRWPENTTDAVKEDV	KREGLYLNHKEHLNHI	VDFRCEEGGGRMDDEEARRLMEELRPHPYLENLA	
Sen. sequence22	(669)	LEIQIRWPENTTDAVKEDV	KREGLYLNHKEHLNHI	VDFRCEEGGGRMDDEEARRLMEELRPHPYLENLA	

FIG 3 L

Res. sequence	(771)	VKAYYGVMKMPGWATLLPNLTFLSDCGELENLPCLGNDHLLKVLRLSHLAKLEYIEEDSSSANFCRPG	771	840
Sen. sequence1	(771)	VKAYYGAKMPGWATLLPNLTLYLSDCGESECLPCMGNDCLKVLRLSHLAKLEYIEEDSTSANFSFRPG		
Sen. sequence10	(771)	VKAYYGAKMPGWATLLPNLTLYLSDCGESECLPCMGNDCLKVLRLSHLAKLEYIEEDSTSANFSFRPG		
Res. sequence	(841)	PESAGLSLYFPSLERLELKRLCKLKGWRRGEGLGDDHQPFNESS>Ret.Ins<NTQVQLQLCLPQLKSLRIERCPLLT	841	910
Sen. sequence1	(841)	PESAGLSLYFPSLELLELKRLHKLKGWRRREGLGDDHQPFNESS>Ret.Ins<-----		
Sen. sequence10	(841)	PESAGLSLYFPSLELLELKRLHKLKGWRRREGLGDDHQPFNESS>Ret.Ins<-----		
Res. sequence	(911)	FMPLCPKTEKHLVVFNERLRIVHAKRDENFYAPLHSSSDPENPRNTIPIPMFREYVYINNVAWLNSLPM	911	980
Res. sequence	(981)	EAFRCLTHMTIKNDEVESLGEVGEVFRSCSSSLRSLNITGCSNLRSVSGGLEHLTALEMLEIYDTHKLSL	981	1050
Res. sequence	(1051)	SEDPEGVVPWKSLHHSLSYLQLMNLPQLVNLPDSMQFLAALRTLIVHCTKLQSVPDWMPRLTSLRKLMV	1051	1120
Res. sequence	(1121)	SFCSAHLERRCQNPTGVDWPNIQHIPSIDVTSSLPKFLVLPYE	1121	1163

FIG 4

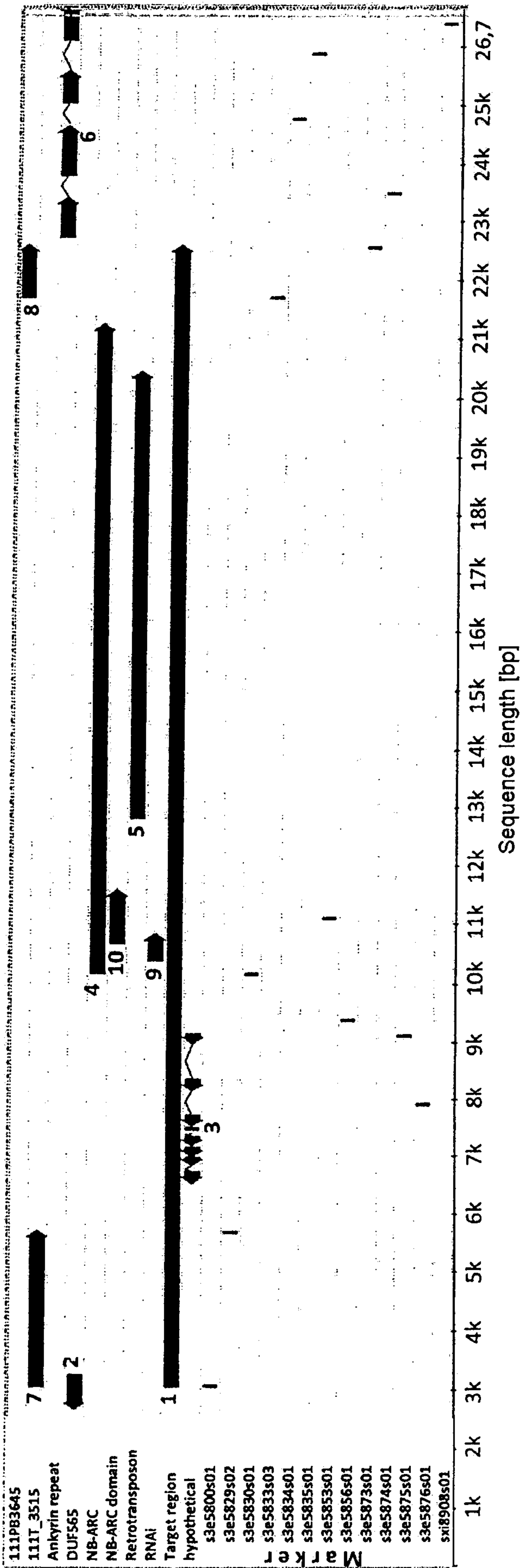


FIG 5

Target Region		Marker										Plants		ELISA values				Splicing			Assertion											
Recombinant Line	Population	Physical Position (bp)										Number of descendants			Number			Average				Population splices	DIFF_RR-ss	T-Test	Milcoxon	Position of the gene						
		9354172	9354296	9404779	9405037	9411766	9414333	9416551	9417771	9418031	9418784	9419705	9430745	9431596	9432522	9433793	9434936	9435438	9442775	Number RR	Number RS						Number ss	Average all	Average RR	Average RS	Average ss	
ZR11007_03075	111T_3515	h	h	h	h	h	h	h	h	h	h	h	h	h	h	h	h	h	h	152	45	61	36	0.161	0.282	0.107	0.117	no	0.165	0.96	0.54	right of s3e5800s01
ZR08093_02718	111PB3553	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	177	40	94	43	2.686	2.608	2.736	2.652	no	-0.04	0.39	0.43	right of s3e5800s01
ZR08093_04549	111PB3857	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	50	14	30	7	2.00	1.94	2.09	1.89	no	0.05	0.54	0.51	left of s3e5873s01
ZR08093_05621	111PB3645	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	73	27	28	17	2.334	1.777	2.397	3.091	yes	-1.313	0.00	0.00	left of s3e5874s01
ZR08093_03547	111PB3579	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	177	41	97	39	1.703	1.121	1.683	2.474	yes	-1.35	0.00	0.00	left of s3e5834s01
ZR08093_02513	111PB3546	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	174	45	87	42	1.848	1.146	1.845	2.630	yes	-1.48	0.00	0.00	left of s3e5834s01
ZR08093_03702	111PB3585	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	79	13	45	20	1.776	0.699	1.455	3.183	yes	-2.484	0.00	0.00	left of s3e5834s01
ZR08093_01718	111PB3523	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	169	50	76	43	1.736	1.157	1.786	2.466	yes	-1.31	0.00	0.00	left of s3e5834s01

1051

FIG 6

