

(19) **DANMARK**

(10) **DK/EP 3011037 T4**



Patent- og
Varemærkestyrelsen

(12) **Oversættelse af ændret
europæisk patentskrift**

-
- (51) Int.Cl.: **C 12 N 15/82 (2006.01)** **C 07 K 14/415 (2006.01)** **C 12 N 15/113 (2010.01)**
- (45) Oversættelsen bekendtgjort den: **2024-09-09**
- (80) Dato for Den Europæiske Patentmyndigheds bekendtgørelse om opretholdelse af patentet i ændret form: **2024-06-26**
- (86) Europæisk ansøgning nr.: **14750139.9**
- (86) Europæisk indleveringsdag: **2014-06-06**
- (87) Den europæiske ansøgnings publiceringsdag: **2016-04-27**
- (86) International ansøgning nr.: **DE2014000310**
- (87) Internationalt publikationsnr.: **WO2014202044**
- (30) Prioritet: **2013-06-17 DE 102013010026**
- (84) Designerede stater: **AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR**
- (73) Patenthaver: **KWS SAAT SE & Co. KGaA, Grimsehlstraße 31, 37574 Einbeck, Tyskland**
- (72) Opfinder: **TÖRJËK, Otto, Gebrüder-Grimm-Weg 11, 37574 Einbeck, Tyskland**
BORCHARDT, Dietrich, Oehlmannstr. 4, 37574 Einbeck, Tyskland
MECHELKE, Wolfgang, Alter Stadtweg 3, 37574 Einbeck, Tyskland
LEIN, Jens Christoph, Schillerstr. 40, 37083 Göttingen, Tyskland
- (74) Fuldmægtig i Danmark: **AWA Denmark A/S, Strandgade 56, 1401 København K, Danmark**
- (54) Benævnelse: **RESISTENSGEN MOD RIZOMANIA**
- (56) Fremdragne publikationer:
GIDNER SARA ET AL: "QTL mapping of BNYVV resistance from the WB41 source in sugar beet", GENOME, Bd. 48cl, Nr. 2, April 2005 (2005-04), Seiten 279-285, XP009180707, ISSN: 0831-2796 in der Anmeldung erwähnt
LEIN JENS CHRISTOPH ET AL: "Resistance gene analogues are clustered on chromosome 3 of sugar beet and cosegregate with QTL for rhizomania resistance", GENOME, Bd. 50, Nr. 1, Januar 2007 (2007-01), Seiten 61-71, XP001526559, ISSN: 0831-2796
TIAN YANYAN ET AL: "The absence of TIR-type resistance gene analogues in the sugar beet (Beta vulgaris L.) genome.", JOURNAL OF MOLECULAR EVOLUTION, Bd. 58, Nr. 1, Januar 2004 (2004-01), Seiten 40-53, XP002731101, ISSN: 0022-2844
GRIMMER M K ET AL: "An anchored linkage map for sugar beet based on AFLP, SNP and RAPD markers and QTL mapping of a new source of resistance to Beet necrotic yellow vein virus", THEORETICAL AND APPLIED GENETICS ; INTERNATIONAL JOURNAL OF PLANT BREEDING RESEARCH, SPRINGER, BERLIN, DE, Bd. 114, Nr. 7, 9. Februar 2007 (2007-02-09), Seiten 1151-1160, XP019510491, ISSN: 1432-2242, DOI: 10.1007/S00122-007-0507-3
BUTORINA A K ET AL: "Molecular genetic investigation of sugar beet (L.)", RUSSIAN JOURNAL OF GENETICS, NAUKA/INTERPERIODICA, MO, Bd. 47, Nr. 10, 8. Oktober 2011 (2011-10-08), Seiten 1141-1150, XP019962364, ISSN: 1608-3369, DOI: 10.1134/S102279541110005X

Fortsættes ...

AMIRI R ET AL: "A new RAPD marker for beet necrotic yellow vein virus resistance gene in *Beta vulgaris*", *BIOLOGIA PLANTARUM*, KLUWER ACADEMIC PUBLISHERS, DO, Bd. 53, Nr. 1, 21. März 2009 (2009-03-21), Seiten 112-119, XP019670809, ISSN: 1573-8264

PAVLI OURANIA I ET AL: "Achievements and prospects in breeding for rhizomania resistance in sugar beet", *FIELD CROPS RESEARCH*, Bd. 122, Nr. 3, Juni 2011 (2011-06), Seiten 165-172, XP002731102, ISSN: 0378-4290 WO-A1-2020/053313

DE-A1-102013 010 026

THIEL et al.: "Identification of Beet necrotic yellow vein virus P25 Pathogenicity Factor - Interacting Sugar Beet Proteins That Represent Putative Virus Targets or Components of Plant Resistance", *Molecular Plant-Microbe Interactions*, vol. 22, no. 8, 2009, pages 999-1010,

DATABASE Nucleotide 2006, TIAN et al.: "Beta vulgaris clone cZR-7(f) NBS-LRR type resistance protein mRNA, complete cds", retrieved from NCBI Database accession no. DQ907612.1

Tandem Technology: "Magistral leaflet", Ses VanderHave Technical Bulletin, 8 August 2007 (2007-08-08),

Anonymous: "Abstract brochure Part A:Papers", Fourth Meeting of the ECPGR Working Group on Beta and the World Beta Network, 20 June 2012 (2012-06-20), - 22 June 2012 (2012-06-22), page 9,

MEULEMANS et al.: "Interaction between different genes and influence of the genetic background in the expression of Rhizomania resistance", 1st joint IIRB-ASSBT Congress, 2003, pages 161-173,

Wan-Song NI et al.: "Construction of a Plant Transformation-ready Expression cDNA Library for *Theilungella halophila* Using Recombination Cloning", *Journal of Integrative Plant Biology*, vol. 49, no. 9, 2007, pages 1313-1319,

KWS Saat SE: "KWS sees the Breeding Reservation weakened by a patent decision of the EPO", *AgroPages*, 13 December 2018 (2018-12-13), Retrieved from the Internet: URL:<http://news.agropages.com/News/NewsDetail---28699.htm>

Friedrich J. Kopisch-Obuch et al.: "Maintaining High Levels of Rhizomania Resistance in Sugar Beet Cultivars", 76th IIRB Congress, 6 June 2018 (2018-06-06), pages 1-13,

Gina G Capistrano-Gossmann; Et Al: "Crop wild relative populations of *Beta vulgaris* allow direct mapping of agronomically important genes", *Nature Communications*, vol. 8, 2017, pages 1-8, & SES VanderHave, "Sequence element from Capistrano et al around probe CAU38880", [cited 25.01.2019] & SES VanderHave. "Rz2 Capistrano vs Rz3 KWS". [cited 25.01.2019] & SES VanderHave. "SEQ ID NO:1 vs Capistrano", [cited 14.02.2019]

SES VanderHave. "Rz2/3 vs shotgun", [cited 28.01.2019] & Dohm et al, "GenBank AYZW01060923.1" 17.12.2013, [cited 14.02.2019]

SES VanderHave, "SEQ ID NO:1 vs ARYA01032019.1" 13.05.2013, [cited 04.02.2019] & SES VanderHave. "GenBank ARYA01032019.1" 13.06.2013, [cited 14.02.2019]

anonymous: "2013 Sugarbeet Research and Extension Reports", North Dakota State University, vol. 44, 1 January 2014 (2014-01-01), page 2pp, 249, 263, 272, 273,

CUI, Y. et al.: "Fine Mapping of the Bsr1 Barley Stripe Mosaic Virus Resistance Gene in the Model Grass *Brachypodium distachyon*", *PLOS One*, vol. 7, no. 6, June 2012 (2012-06), pages e38333-1-e38333-10,

Declaration of Dr. Harlene Hatterman-Valenti & Declaration of Dr. Hendrik Tschoep & Scheme of the DNA extraction procedure performed at SESVANDERHAVE

Data sequencing of the variety 36812 RR (cfr D15) from the external provider Eurofins Genomics and explanations on how the sequencing has been carried out

Pferdmenges, Friederike et al., "Beet necrotic yellow vein virus (BNYW) P-type overcomes resistance in Rz1 x Rz2sugar beet hybrids after mechanical inoculation? 2006: poster

Pferdmenges F.: "Occurrence, spread and pathogenicity of different Beet necrotic yellow vein virus (BNYVV) isolates", Dissertation Georg-August Universität Göttingen, 1 September 2007 (2007-09-01), pages 1-111,

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

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 application 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

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 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

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 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 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 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 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).

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) 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 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 BNVYY 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 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.

- 12 -

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.

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 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 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

- 15 -

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:

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.

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

- 17 -

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:

- 18 -

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

- 19 -

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 (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

- 20 -

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

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

- 22 -

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 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*

- 23 -

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.

P A T E N T K R A V

1. Nukleinsyremolekyle, som koder for et polypeptid, der er i stand til at bibringe resistens over for et patogen i en plante, hvor polypeptidet udtrykkes, *k e n d e t e g n e t* ved, at nukleinsyremolekylet omfatter en nukleotidsekvens, som er udvalgt blandt
 - 5 a) en nukleotidsekvens, som koder for et polypeptid med en aminosyresekvens ifølge SEQ ID NO:2 eller SEQ ID NO:3,
 - b) en nukleotidsekvens, som omfatter den kodende sekvens af DNA-sekvensen ifølge SEQ ID NO:1,
 - c) en nukleotidsekvens, som koder for et polypeptid, der kan afledes af et polypeptid, som kodes for af nukleotidsekvensen ifølge a) eller b), ved substitution, deletering og/eller tilføjelse af en aminosyre i aminosyresekvensen, som kodes for af nukleotidsekvensen ifølge a) eller b),
 - d) en nukleotidsekvens, som koder for et polypeptid, der frembyder en aminosyresekvens, som har mindst 80 % similaritet med en aminosyresekvens, som kodes for af
 - 15 nukleotidsekvensen ifølge a) eller b), eller
 - e) en nukleotidsekvens, som koder for mindst aminosyrepositionerne 168-227 i SEQ ID NO:2 og mindst aminosyrepositionerne 591-613 i SEQ ID NO:2 og mindst aminosyrepositionerne 1013-1072 i SEQ ID NO:2, eller som koder for mindst aminosyrepositionerne 182-241 i SEQ ID NO:3, mindst aminosyrepositionerne 605-627 i SEQ ID NO:3 og mindst
 - 20 aminosyrepositionerne 1027-1086 i SEQ ID NO:3.
2. Vektor, som omfatter nukleinsyremolekylet ifølge krav 1.
3. Prokaryot celle eller gærcele, som omfatter nukleinsyremolekylet ifølge krav 1 eller vektoren ifølge krav 2.
4. Polypeptid, som er i stand til at bibringe resistens mod et patogen i en plante,
 - 25 hvori polypeptidet udtrykkes, og som kodes for af nukleinsyremolekylet ifølge krav 1.
5. Transgen plantecelle, som omfatter nukleinsyremolekylet ifølge krav 1 som transgen eller vektoren ifølge krav 2.
6. Transgen plante eller del deraf, omfattende en plantecelle ifølge krav 5.
7. Frø af planten ifølge krav 6, hvor frøene omfatter nukleinsyremolekylet ifølge
 - 30 krav 1 som transgen.
8. Fremgangsmåde til fremstilling af en transgen plantecelle, *k e n d e t e g n e t* ved,

at fremgangsmåden omfatter skridtet at indbringe nukleinsyremolekylet ifølge krav 1 eller vektoren ifølge krav 2 i plantecellen.

9. Fremgangsmåde til fremstilling af en transgen plante, k e n d e t e g n e t ved, at fremgangsmåden omfatter følgende trin:

5 a) at indbringe nukleinsyremolekylet ifølge krav 1 eller vektoren ifølge krav 2 i en plantecelle og

b) at regenerere den transgene plante fra den transgene plantecelle fra trin a).

10. Fremgangsmåde til bestemmelse af et nukleinsyremolekyle, som koder for et protein, der er i stand til at bibringe resistens over for patogenet BNYVV i en plante af slægten Beta, hvori proteinet udtrykkes, k e n d e t e g n e t ved, at fremgangsmåden omfatter følgende trin:

i. at påvise fraværet af en indsættelse i den kodende nukleotidsekvens af nukleinsyremolekylet ifølge 1 eller

15 ii. at påvise mindst en polymorfisme ifølge Figur 5 1, 2 og/eller 3 i den kodende nukleotidsekvens af nukleinsyremolekylet ifølge krav 1 under anvendelse af molekulære markører, som genkender polymorfismen.

11. Molekulær markør til selektion af BNYVV-resistente planter, som påviser en polymorfisme ifølge Figur 1, 2 eller 3.

20 12. Anvendelse af en molekulær markør ifølge krav 12 til selektion af BNYVV15 resistente planter, omfattende

i) at påvise en polymorfisme ifølge Figur 1, 2 eller 3 eller

ii) at bestemme tilstedeværelsen eller fraværet af en indsættelse i den kodende nukleotidsekvens af et nukleinsyremolekyle ifølge krav 1.

25 13. Anvendelse af en med nukleinsyremolekylet ifølge krav 1 nært forbunden molekulær markør på en DNA-sekvens ifølge SEQ ID NO:4 eller på en DNA-sekvens ifølge SEQ ID NO:5 i en fremgangsmåde til selektion af en plante, som udviser resistens over for BNYVV.

FIG 1 A

consensus-sensitive resistant sequence	1 70 (1) CAAATCTTCTGGCATCAATGGCGGTGTTGCCGTTCA~~W~~TTTAACATCAATGGAGGTAAGAGTCATGTT (1) CAAATCTTCTGGCATCAATGGCGGTGTTGCCGTTCA~~A~~TTTAACATCAATGGAGGTAAGAGTCATGTT
consensus-sensitive resistant sequence	71 140 (71) TTTTCAACAATATAAAACTTATA-----TGATTTTTCGTGTTTTC~~CCC~~----- (71) TTTTCAACAATATAAAACTTATACTTCCTCTGTTCTGTTTAAATGA~~AACG~~TTGTTTCT~~CA~~CGCAACC
consensus-sensitive resistant sequence	141 210 (113) ----- (141) CAACCCACTTTTTTAATAATAAATATTTT~~TTAGTTGTGTGCACGTAAAAA~~TATAAAAAAGTTATAATTG
consensus-sensitive resistant sequence	211 280 (114) ---GTATCTTGTTTGAGATTGTGATTATTAAGAGAGTCAAGTCTCACAATATTCGAAAGTCTACGTAATC (211) ~~ATA~~GTATCTTGTTTGAGATTGTGATTATTAAGAGAGTCAAGTCTCACAATATTCGAAAGTCTACGTAATC
consensus-sensitive resistant sequence	281 350 (181) CACCTCAAATTGACGAAGAAAACAAGCAGGAAAGGATTAAGTAAGTTCGTGGAACCA~~Y~~TAGAATT~~A~~ATTT (281) CACCTCAAATTGACGAAGAAAACAAGCAGGAAAGGATTAAGTAAGTTCGTGGAACCA~~C~~TAGAATT~~G~~ATTT
consensus-sensitive resistant sequence	351 420 (251) TCAAATATAGCTCTACCTAATATATGGCCTACTTTTAATTTTAAATAAG~~A~~AGAAGGTAATGTGATTAGAA (351) TCAAATATAGCTCTACCTAATATATGGCCTACTTTTAATTTTAAATAAG~~G~~AGAAGGTAATGTGATTAGAA
consensus-sensitive resistant sequence	421 490 (321) ACAAATTGGTCTTAAATTATTCATTAAGCTTAATAATGCATAAACTTTATCAAGTGCTATCTTCTTTTCA (421) ACAAATTGGTCTTAAATTATTCATTAAGCTTAATAATGTATAAAC~~ATA~~AATCAAGTGCTATCTTCTTTTCA
consensus-sensitive resistant sequence	491 560 (391) GGGCCGTCTTGAAGATTTTKGGKCCCRGTTCTATTATGAAAATTGRGCCCTAAATTTATASAAAATAAA (491) GGGCCGTCTTGAAGATTTTGGGCCC~~G~~GTTCTATTATGAAAATTGGGCCCTAAATTTATAG~~A~~AAAATAAA

FIG 1 B

[illegible]

FIG 1 C

		1121		1190
consensus-sensitive	(779)	AGATGAATTGTTTGTCTTTGATGGTCTCCAATGCATATTTTGTAT	WMTTAGGAATTYTAATTATGTACTA	
resistant sequence	(1096)	AGATGAATTGTTTGTCTTTGATGGTCTCCAATGCATATTTTGTAT	ACTTAGGAATTCTAATTATGTACTA	
		1191		1260
consensus-sensitive	(849)	TTAGTASAYATYGAGAYGAATACARAATYGCCATAATGAAGTATGATTATTTTARTTATATACTTTCTCC		
resistant sequence	(1166)	TTAGTAGACATTGAGATGAATACATAATYGCCATAATGAAGTATGATTATTTTAGTTATATACTTTCTCC		
		1261		1330
consensus-sensitive	(919)	GTTCCAAATATATAAARTGTAACACTTGTGTACTTTATGCGTACTAATGCATAAAYAACGTGCACTCTCCY	G	
resistant sequence	(1236)	ATTCCAAATATATAAAATGTAACACTTGTGTACTTTATGCGTACTAATGCATAACAACGTGCACTCTCA	ATG	
		1331		1400
consensus-sensitive	(989)	TGTTTAATTATATACTTTTTGAGAGAAGTGWTACATTGGGGACCATGGGACTGTGTATAATTTGACCGCA		
resistant sequence	(1306)	TGTTTAATTATATACTTTTTGAGAGAAGTGTTACATTGGGGACCATGGGACTGTGTATAATTTGACCGCA		
		1401	↓Start1	1470
consensus-sensitive	(1059)	AAATYGAAGTGTYGCATTTGATTGAAAAYGGAGARRGTAGTATATAGRTGGAACACNNNNNNNNNTGCTG		
resistant sequence	(1376)	AAATCGAAGTGTCGCATTTGATTGAAAATGGAGAGAGTAGTATATAGATGGAACACAGCAGAGACTGCTG		
		1471		1540
consensus-sensitive	(1129)	RTCATCTTTGGCCAACAAACCMAAAATTGATATTAATCCYYTWWTYMRGKYWNNTTTCATCTTTTGGACA		
resistant sequence	(1446)	GTCATCTTTGGCCAACAAACCMAAAATTGATATTAATCCCTTATTCAGGTCA--TTTCATCTTTTGGACA		
		↓Start2		1610
consensus-sensitive	(1199)	CAAAATGGATGTTGTAGGCWCTGCGCTATCTGCTGCCCAATCTCTGTTTGCAGCCCTGCAAAGTTCTGAG		
resistant sequence	(1514)	CAAAATGGATGTTGTAGGCACCTGCGCTATCTGCTGCCCAATCTCTGTTTGCAGCCCTGCAAAGTTCTGAG		
		1611		1680
consensus-sensitive	(1269)	CTCAAAGAGATCCTCTCGATCTTTGGCTACAAATCCC	A	ACTTGATGACCTCCAACGCAYTGMTCTACCA
resistant sequence	(1584)	CTCAAAGAGATCCTCTCGATCTTTGGCTACAAATCCC	G	ACTTGATGACCTCCAACGCACTGTCTCTACCA

FIG 1 D

		1681		1750
consensus-sensitive	(1339)	TCAAY <u>G</u> CTGTATTCCGTGATGCTGAGACCAAACAGGAGCTCACTCATGAAGCACAR <u>C</u> ATTGGCTCGAGGA		
resistant sequence	(1654)	TCAAC <u>G</u> CTGTATTCCGTGATGCTGAGACCAAACAGGAGCTCACTCATGAAGCACAG <u>C</u> ATTGGCTCGAGGA		
		1751		1820
consensus-sensitive	(1409)	ACTCAAGGATGCTGTCTTTGAAGCAGATGATCTGTTGACGAGTTTGTCACTCTTGCCGAGCAGAAGCAA		
resistant sequence	(1724)	ACTCAAGGATGCTGTCTTTGAAGCAGATGATCTGTTGACGAGTTTGTCACTCTTGCCGAGCAGAAGCAA		
		1821		1890
consensus-sensitive	(1479)	CTTGTAGAGGCTGGTGGCAGTCTTTCCAAAAAGATGCGCCAATTCTTTTCTGATTCCAACCCCCTTGGCA		
resistant sequence	(1794)	CTTGTAGAGGCTGGTGGCAGTCTTTCCAAAAAGATGCGCCAATTCTTTTCTGATTCCAACCCCCTTGGCA		
		1891		1960
consensus-sensitive	(1549)	TYGCTTATARGATGTCACRAGGGGTAAAGAAGATCAAGAAGAAGTTGGATGYTATYGCCTTACAATCATCA		
resistant sequence	(1864)	TTGCTTATAGGATGTCACGAGGGGTAAAGAAGATCAAGAAGAAGTTGGATGCTATCGCTTACAATCATCA		
		1961		2030
consensus-sensitive	(1619)	ATTTAGCTTTAAGATTGATCTTGAGCCTATAAAAGAGAGAAGGCTCGAGACTGGTTCTGTCGTGAACGCA		
resistant sequence	(1934)	ATTTAGCTTTAAGATTGATCTTGAGCCTATGAAAGAGAGAAGGCTAGAGACTGGTTCTGTCGTGAACGCA		
		2031		2100
consensus-sensitive	(1689)	GGTGATATCATTGGAAGAGAGGATGACTTGGAGAAGATCGTAGGTTTGTTKCTTGATTCTAACATCCAGC		
resistant sequence	(2004)	GGTGATATCATTGGAAGAGAGGACCGACTTGGAGAAGATGCTAGGTTTGTGCTTGATTCTAACATCCAAC		
		2101		2170
consensus-sensitive	(1759)	GTGATGTGTCTTTCCTTACKATWGTGGGAATGGGAGGGTTGGGTAAAACCTGCTCTTGCCCAACTCGTGTA		
resistant sequence	(2074)	GTGATGTGTCTTTCCTTACTATTGTGGGAATGGGAGGGTTGGGTAAAACCTGCTCTTGCCCAACTCGTGTA		
		2171		2240
consensus-sensitive	(1829)	CAATGATCCAAGGGTCAGAACTGCTTTTCCATTGAGATGTTGGAATTGTNNCTCTGATCAAGATCAAAAK		
resistant sequence	(2144)	CAATGATCCAAGGGTCAGAACTGCTTTTCCATTGAGATGTTGGAATTGTGTCTCTGATCAAGATCAAAAG		

FIG 1 E

consensus-sensitive resistant sequence	(1899) (2214)	2241	2310
		MAACTAGATGTGAAAGAAATTTTGGGTAAGATTCTG <u>T</u> CTACAGCTACTGGTAAGAATCAYRAGGGTTCAA CAACTAGATGTGAAAGAAATTTTGGGTAAGATTCTG <u>G</u> CTACAGCTACTGGTAAGAATCATGAGGGTTCAA	
consensus-sensitive resistant sequence	(1969) (2284)	2311	2380
		CCATGGATCAKGTGCAAACCYAACTACRRGAACAACTATGTGGCAAGAGATACTTGCTTGTTTTGGATGA CCATGGATCAGGTGCAAACCCAACTACGAGAACAACCTATGTGGCAAGAGATACTTGCTTGTTTTGGATGA	
consensus-sensitive resistant sequence	(2039) (2354)	2381	2450
		TGTATGGAATGAGAATCCTAATCAATTGCGTDWYYTKGKWRAATTCTTCATGGGAGGTCAAGGGGAAAT TGTATGGAATGAGAATCCTAATCAATTGCGTGATCTGGTAGAATTCTTCATGGGAGGTCAAGCAAAAT	
consensus-sensitive resistant sequence	(2109) (2424)	2451	2520
		TGGATTSTGGTAACTACGCGTTCGCAYGAGACARCGAGAATTATAAGAGATGGTCCATTGCACAAGCTSC TGGATTGTGGTAACTACGCGTTCGCACGAGACAGCGAGAATTATAAGAGATGGTCCATTGCACAAGCTCC	
consensus-sensitive resistant sequence	(2179) (2494)	2521	2590
		AAGGTTTGTCTGARRNAAAACYTTGGCGTTTATYTGTAAGGTGGACCTTCGGATCAGTGCAACCAAAAT AAGGTTTGTCTGAGG-AAAACCTTTGGCGTTTATTTGTCAGGTGGACCTTCGGATCAGTGCAAGCAAAAT	
consensus-sensitive resistant sequence	(2249) (2563)	2591	2660
		TCCCTAATGACTTTGTTCATGATTGCACGAGATATAGTYGACAAATGTGCTCGAAACCCTYTGGCTATAAG TCCCTAATGACTTTATTCATGATTGCACGAGATATAGTTGACAAATGTGCTCGAAACCCTCTGGCTATAAG	
consensus-sensitive resistant sequence	(2319) (2633)	2661	2730
		AKTGGTAGGAAGTCTTTTGTGTGGTCAAGACAAGAGTAAGTGGCTTTTCATTTTCATGAGATMTGTTTAGCC AGTGGTAGGAAGTCTTTTGTGTGGTCAAGACAAGAGTAAGTGGCTTTTCATTTTCATGAGATCGATTTTAGGC	
consensus-sensitive resistant sequence	(2389) (2703)	2731	2800
		AACATTAGAAAGAGYCATAATGATATCATGCYAATACTGAACCTAAGTTACCATCATCTTGAACCTCCAA AACATTAGAAAGAGCCATAATGATATCATGCCAATACTGAACCTAAGTTACCATCATCTTGAACCTCCAA	

FIG 1 F

		2801		2870
consensus-sensitive	(2459)	TCAKAGAGATGCTTTAGTTATTGTGCARTGTTTCCAAAGGATTTCTTATAGGGAAGAAAGACGTTGATAAA		
resistant sequence	(2773)	TTAAGAGATGCTTTAGTTATTGTGCAGTGTTCCTTATAGGGAAGCCAGACGCTGATAAA		
		2871		2940
consensus-sensitive	(2529)	CCTTTGGATGGCACAAGGTTATATTGTTCCATTAGACAAAGATCAAAGCATAGATGAYGCTAGTGAGGAA		
resistant sequence	(2843)	CCTCTGGATGGCACAAGGTTATATTGTTCCGTTAGACAAAGATCAAAGCATAGATGATGCTAGTGAGGAA		
		2941		3010
consensus-sensitive	(2599)	TACATATCAATTTTGYTGCAAGAGATGTTTTTCGAAAAACATCGGAACAGAAAAAGATKATGTTATTAAGA		
resistant sequence	(2913)	TACATATCAATTTTGTGCGGAGATGTTTTTCGAAAAATGTCGGAGCAGAAAAAGATGGTGTATTTAAGA		
		3011		3080
consensus-sensitive	(2669)	TACATGATCTCATGCATGATATTGCTCAAAATGTCATGGGGAAGGAGCTTTGTACGACAAAAACATTAG		
resistant sequence	(2983)	TCCATGATCTCATGCATGATATTGCTCAAAATGTCATGGGGAAGGAGCTTTGTACGACTAAAAACATTAG		
		3081		3150
consensus-sensitive	(2739)	TGGCAGCTTGGATAAAAAATGTTTCGCCATCTATCTCTTGCCAGAACTAGTTTTGCAAGATACTCTTTYAAT		
resistant sequence	(3053)	TGGCAGCTTGGATAAAAAAGTGTTCGCCATCTATCTCTTGCCAGAACTAGTTTTGCAAGATACTCTTTCAAT		
		3151		3220
consensus-sensitive	(2809)	GCAACTCATATTCGCTCCYATTTCTRTGCTGGCTACTGGTGTGAGGAWKCTGAGATAAMCCAGTTYTCAG		
resistant sequence	(3123)	GCAACTCATATTCGCTCCTATTTCTGTGCTGGCTACTGGTGTGAGGATGCTGAGATAAACCCAGTTTTCAG		
		3221		3290
consensus-sensitive	(2879)	TTGAGGCATTAGTACCAAAYTGTTTGTGCCTAAGGGCAATGKACCTSGCTTGGTCGAAGATAAAAAAGTKT		
resistant sequence	(3193)	TTGAGGCATTAGTACCAAAC TGTTTGTGCCTAAGGGCAATGGACCTCGCTTGGTCGAAGATAAAAAAGTTT		
		3291		3360
consensus-sensitive	(2949)	ACCAGACTCRATTGGTGGATTGTTGCATTTGAGGTACTTAGATCTTTCRTATAASGAAGATYTGAAGTA		
resistant sequence	(3263)	ACCAGACTCGATTGGTGGATTGTTGCATTTGAGGTACTTAGATCTTTCGTATAACGAAGATCTGGAAGTA		

FIG 1 G

		3361		3430
consensus-sensitive	(3019)	CTTCC <u>CA</u> ACTCAATTGCYAACTATATAATCTRCAAACCTTACAATTGAAGGGTTGCAAGAGATTGGAAG		
resistant sequence	(3333)	CTTCC <u>CA</u> ACTCAATTGC <u>T</u> AACTATATAATCT <u>A</u> CAAACCTTACAATTGAAGGGTTGCAAGAGATTGGAAG		
		3431		3500
consensus-sensitive	(3089)	GGTTAYCAAAACATTTGAGCAGGCTGGTTAAGCTTCAAACCTTTRGATATA <u>AT</u> GGTTGCAAYAATGTAAC		
resistant sequence	(3403)	GGTTA <u>C</u> CAAAACATTTGAGCAGGCTGGTTAAGCTTCAAACCTTGGATATA <u>AT</u> GGTTGCAACAATGTAAC		
		3501		3570
consensus-sensitive	(3159)	TTATATGCCCAAAGGCATGGGTAAG <u>AT</u> GACTTGCCTTCACACTCTCAGTAAGTTTATAGTGGGTGGAGAA		
resistant sequence	(3473)	TTATATGCCCAAAGGCATGGGTAAG <u>AT</u> GACTTGCCTTCACACTCTCAGTAAGTTTATAGTGGGTGGAGAA		
		3571		3640
consensus-sensitive	(3229)	GGG <u>AR</u> TTGTTCAAGTTGGAAGMAAYG <u>CT</u> TTTGATGGG <u>CG</u> WGAAGATCTAAAGGCTCTCA <u>AG</u> AACCTAAAGG		
resistant sequence	(3543)	GGGAG <u>TT</u> GTTCAAGTTGGAAGCAAT <u>CT</u> TTTGATGGG <u>CT</u> TGGAAGATCTAAAGGCTCTCA <u>AG</u> AACCTAAAGG		
		3641		3710
consensus-sensitive	(3299)	GTCATCTGGRAATCCAAATCAGGTGGCCC <u>AAA</u> AATACTACAGATGCTGTCA <u>AG</u> GAAGATGTTA <u>AG</u> GAGGGA		
resistant sequence	(3613)	GTCATCTGG <u>AA</u> ATCCAAATCAGGTGGCCC <u>AAA</u> AATACTACAGATGCTGTCA <u>AG</u> GAAGATGTTA <u>AG</u> GAGGGA		
		3711		3780
consensus-sensitive	(3369)	AGGATTATAC <u>Y</u> TGAATCATAAGGAACATCTCAATCACATTGTGGTTGATTT <u>AG</u> ATGTGAGGAGGGTGGT		
resistant sequence	(3683)	AGGATTATAC <u>C</u> TGAATCATAAGGAACATCTCAATCACATTGTGGTTGATTT <u>AG</u> ATGTGAGGAGGGTGGT		
		3781		3850
consensus-sensitive	(3439)	GGAAGAATGGATGATGAGGAAGCAAGAAGATTGATGGAAGAG <u>Y</u> TGCGGCCACATCCTTATCTTGAAAATT		
resistant sequence	(3753)	GGAAGAATGGATGATGAGGAAGCAAGAAGATTGATGGAAGAG <u>T</u> TGCGGCCACATCCTTATCTTGAAAATT		
		3851		3920
consensus-sensitive	(3509)	TGGCTGTGAAAGCATAYTATGGTGYGAAAAYGCCTGR <u>TT</u> GGGYAACCTTCTYCCAAATCTTACAGAGCT		
resistant sequence	(3823)	TGGCTGTGAAAGCATAT <u>T</u> ATGGTGT <u>G</u> AAAATGCCTG <u>TT</u> GGGCAACCCTTCTC <u>C</u> CAAAATCTTACAGAGCT		

FIG 1 H

		3921		3990
consensus-sensitive	(3579)	TTWTCTTTYTGATTGTGGGGAAYYGGAGWRCCTTCCATGCMTGGGAAACTTGGWTYDTCTRAAMGTYCTC		
resistant sequence	(3893)	TTTTCTTTCTGATTGTGGGGAAC <u>TGGAGA</u> ACCTTCCATGCCTGGGAAACTTGGATCATCTAAAAGTCTC		
		3991		4060
consensus-sensitive	(3649)	CGRCTTTCGCATTTGGCRAAATTGGAGTAYATTGRAGAAGATAGCWCATCAGCTMWTTTCAGKTCTAGGC		
resistant sequence	(3963)	CGACTTTCGCATTTGGC <u>AAA</u> ATTGGAGTACATTGAAGAAGATAGCTCATCAGCTAATTTTCAGGTGTAGGC		
		4061		4130
consensus-sensitive	(3719)	CTGGACCRGAAAGTGCAGGACTATCATTATACTTCCCCTCCCTTGAACKCCTTGAGTTGAAGCRTTTGYR		
resistant sequence	(4033)	CTGGACCAGAAAGTGCAGGACTATCATTATACTTCCCCTCCCTTGAACGCCTTGAGTTGAAGCGTTTGTG		
		4131		4200
consensus-sensitive	(3789)	YAAGTTAAAAGGATGGAGGAGARGGGAAGGGTTAGGAGATGATCACCAGCCTTTTAATGAAAGCAGCAGC		
resistant sequence	(4103)	TAAGTTAAAAGGATGGAGGAGAGGGGAAGGGTTAGGAGATGATCACCAGCCTTTTAATGAAAGCAGCAGC		
		▼Retrotransposon Insertion		4270
consensus-sensitive	(3859)	AAT-----AAGTCATTGAGAATAGAAAGATGCCCATTGC		
resistant sequence	(4173)	AATACACAAGTCCAATTACAATTATGTCTTCCTCAATTGAAGTCATTGAGAATAGAAAGATGCCCATTGC		
		4271		4340
consensus-sensitive	(3893)	TGACATTTATGCCGCTGTGTCCCAAGACAGAAAAACDGCATTTAGTTGTATTTAATGAAYGACTCCGGAT		
resistant sequence	(4243)	TGACATTTATGCCGCTGTGTCCCAAGACAGAAAAAC <u>TGC</u> ATTTAGTTGTATTTAATGAACGACTCCGGAT		
		4341		4410
consensus-sensitive	(3963)	AGTGCATAC <u>TAAAG</u> GAGATGAGAATTTCTATGCTCCATTACATTCATCATCTGATCCTGAAAACCCG		
resistant sequence	(4313)	AGTGCATC <u>TAAAG</u> GAGATGAGAATTTCTATGCTCCATTACATTCATCATCATCTGATCCTGAAAACCCG		
		4411		4480
consensus-sensitive	(4033)	AGGAACACTATTCCCATTTCCCATGTTAGAGAGGTATACATAAACAATGTGGCTGGCTAAATTTCGCTGC		
resistant sequence	(4383)	AGGAACACTATTCCCATTTCCCATGTTAGAGAGGTATACATAAACAATGTGGCTGGCTAAATTTCGCTGC		

FIG 1 I

		4481		4550
consensus-sensitive	(4103)	CTATGGAGGCTTTTAGGTGTCTCACTCATATGACAATAAAAAACGAC		AGGTAGAGAGTTTGGGAGAAGT
resistant sequence	(4453)	CTATGGAGGCTTTTAGGTGTCTCACTCATATGACAATAAAAAACGAC		AGGTAGAGAGTTTGGGAGAAGT
		4551		4620
consensus-sensitive	(4173)	TGGGAGAGGTGTTTCGGAGCTRCTCATCTTCTTTGCGATCCTTGAATATCACAGGTTGCTCCAACTTAAGA		
resistant sequence	(4523)	TGGGAGAGGTGTTTCGGAGCTGCTCATCTTCTTTGCGATCCTTGAATATCACAGGTTGCTCCAACTTAAGA		
		4621		4690
consensus-sensitive	(4243)	AGTGTCTTCTGGAGGGCTGGAGCATCTCACTRCTTTGGAGATKTTAGAAATATACGACACCCATAAGCTGA		
resistant sequence	(4593)	AGTGTCTTCTGGAGGGCTGGAGCATCTCACTGCTTTGGAGATGTTAGAAATATACGACACCCATAAGCTGA		
		4691		4760
consensus-sensitive	(4313)	GTCTWTCAGAAGACCCAGAAGGTGTTGTGCCATGGAAATCCCTTCATCACTCCCTCAGCTACTTGMAAT		
resistant sequence	(4663)	GTCTATCAGAAGACCCAGAAGGTGTTGTGCCATGGAAATCCCTTCATCACTCCCTCAGCTACTTGCAAT		
		4761		4830
consensus-sensitive	(4383)	GATGAATCTCCCWCAGCTGGTCAACCTGCCTGATTGATGCAGTTCTTGGYCTCCCTCCAAACCTTTCA		
resistant sequence	(4733)	GATGAATCTCCCACAGCTGGTCAACCTGCCTGATTGATGCAGTTCTTGGCTGCCCTCCAAACCTTTCA		
		4831		4900
consensus-sensitive	(4453)	ATGGTGCAATTGCAGTAAACTGGAATCAGTGCCAGATTGGATGCCCCMGACTCACTTCYCTCAGGAAGCTTA		
resistant sequence	(4803)	ATAGTGCAATTGCAGTAAACTGGAATCAGTGCCAGATTGGATGCCCCAGACTCACTTCCTCAGGAAGCTTA		
		4901		4970
consensus-sensitive	(4523)	TGGTTTCATTCTGTTCCGCACATCTGGAGAGAAGATGYCAAATCCAACCTGGGGTGGACTGGCCTAACAT		
resistant sequence	(4873)	TGGTTTCATTCTGTTCCGCACATCTGGAGAGAAGATGCCAAATCCAACCTGGGGTGGACTGGCCTAACAT		
		4971		5037
consensus-sensitive	(4593)	TCAACACATCCCCTSCATTGATGTACCTCTAGCCCTCCTAAGTTTTTAGTGTTGCCGTATGAATAG		
resistant sequence	(4943)	TCAACACATCCCCTCCATTGATGTACCTCTAGCCCTCCTAAGTTTTTAGTGTTGCCGTATGAATAG		

FIG 2 A

	1	70
Res. sequence	(1)	MDVVG T ALSAAQSLFAALQSSELKEILSIFGYKS Q LDLQRTVSTINAVFRDAETKQELTHEAQHWLEEL
Sen. sequence	(1)	MDVVG T ALSAAQSLFAALQSSELKEILSIFGYKS Q LDLQRTVSTINAVFRDAETKQELTHEAQHWLEEL
Sen. sequence2	(1)	MDVVG T ALSAAQSLFAALQSSELKEILSIFGYKS Q LDLQRTVSTINAVFRDAETKQELTHEAQHWLEEL
Sen. sequence3	(1)	MDVVG T ALSAAQSLFAALQSSELKEILSIFGYKS Q LDLQRTVSTINAVFRDAETKQELTHEAQHWLEEL
Sen. sequence4	(1)	MDVVG T ALSAAQSLFAALQSSELKEILSIFGYKS Q LDLQRTVSTINAVFRDAETKQELTHEAQHWLEEL
Sen. sequence5	(1)	MDVVG T ALSAAQSLFAALQSSELKEILSIFGYKS Q LDLQRTVSTINAVFRDAETKQELTHEAQHWLEEL
Sen. sequence6	(1)	MDVVG T ALSAAQSLFAALQSSELKEILSIFGYKS Q LDLQRTVSTINAVFRDAETKQELTHEAQHWLEEL
Sen. sequence7	(1)	-----STINAVFRDAETKQELTHEAQHWLEEL
Sen. sequence8	(1)	MDVVG T ALSAAQSLFAALQSSELKEILSIFGYKS Q LDLQRTVSTINAVFRDAETKQELTHEAQHWLEEL
Sen. sequence9	(1)	MDVVG T ALSAAQSLFAALQSSELKEILSIFGYKS Q LDLQRTVSTINAVFRDAETKQELTHEAQHWLEEL
Sen. sequence10	(1)	MDVVG T ALSAAQSLFAALQSSELKEILSIFGYKS Q LDLQRTVSTINAVFRDAETKQELTHEAQHWLEEL
Sen. sequence11	(1)	MDVVG T ALSAAQSLFAALQSSELKEILSIFGYKS Q LDLQRTVSTINAVFRDAETKQELTHEAQHWLEEL
Sen. sequence12	(1)	MDVVG T ALSAAQSLFAALQSSELKEILSIFGYKS Q LDLQRTVSTINAVFRDAETKQELTHEAQHWLEEL
Sen. sequence13	(1)	-----STINAVFRDAETKQELTHEAQHWLEEL
Sen. sequence14	(1)	MDVVG T ALSAAQSLFAALQSSELKEILSIFGYKS Q LDLQRTVSTINAVFRDAETKQELTHEAQHWLEEL
Sen. sequence15	(1)	MDVVG T ALSAAQSLFAALQSSELKEILSIFGYKS Q LDLQRTVSTINAVFRDAETKQELTHEAQHWLEEL
Sen. sequence16	(1)	MDVVG T ALSAAQSLFAALQSSELKEILSIFGYKS Q LDLQRTVSTINAVFRDAETKQELTHEAQHWLEEL
Sen. sequence17	(1)	MDVVG T ALSAAQSLFAALQSSELKEILSIFGYKS Q LDLQRTVSTINAVFRDAETKQELTHEAQHWLEEL
Sen. sequence18	(1)	MDVVG T ALSAAQSLFAALQSSELKEILSIFGYKS Q LDLQRTVSTINAVFRDAETKQELTHEAQHWLEEL
Sen. sequence19	(1)	MDVVG T ALSAAQSLFAALQSSELKEILSIFGYKS Q LDLQRTVSTINAVFRDAETKQELTHEAQHWLEEL
Sen. sequence20	(1)	MDVVG T ALSAAQSLFAALQSSELKEILSIFGYKS Q LDLQRTVSTINAVFRDAETKQELTHEAQHWLEEL
Sen. sequence21	(1)	MDVVG S ALSAAQSLFAALQSSELKEILSIFGYKS Q LDLQRTVSTINAVFRDAETKQELTHEAQHWLEEL
Sen. sequence22	(1)	MDVVG S ALSAAQSLFAALQSSELKEILSIFGYKS Q LDLQRTVSTINAVFRDAETKQELTHEAQHWLEEL

FIG 2 B

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

FIG 2 C

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

FIG 2 D

	211	280
Res. sequence	(211)	DPRVRTAFPLRCWNCVSDQDQKQLDVKEILGKILSTATGKNHEGSTMDQVQTQLREQLCGKRYLLVLDDV
Sen. sequence1	(211)	DPRVRTAFPLRCWNCVSDQDQKKLDVKEILGKILSTATGKNHEGSTMDHVQTQLREQLCGKRYLLVLDDV
Sen. sequence2	(211)	DPRVRTAFPLRCWNCVSDQDQKKLDVKEILGKILSTATGKNHEGSTMDHVQTQLQEQLCGKRYLLVLDDV
Sen. sequence3	(211)	DPRVRTAFPLRCWNCVSDQDQKKLDVKEILGKILSTATGKNHEGSTMDHVQTQLQEQLCGKRYLLVLDDV
Sen. sequence4	(211)	DPRVRTAFPLRCWNCVSDQDQKKLDVKEILGKILSTATGKNHEGSTMDHVQTQLQEQLCGKRYLLVLDDV
Sen. sequence5	(211)	DPRVRTAFPLRCWNCVSDQDQKKLDVKEILGKILSTATGKNHEGSTMDHVQTQLQEQLCGKRYLLVLDDV
Sen. sequence6	(211)	DPRVRTAFPLRCWNCVSDQDQKKLDVKEILGKILSTATGKNHEGSTMDHVQTQLREQLCGKRYLLVLDDV
Sen. sequence7	(168)	DPRVRTAFPLRCWNCL*-----
Sen. sequence8	(211)	DPRVRTAFPLRCWNCVSDQDQKKLDVKEILGKILSTATGKNHEGSTMDHVQTQLQEQLCGKRYLLVLDDV
Sen. sequence9	(211)	DPRVRTAFPLRCWNCVSDQDQKKLDVKEILGKILSTATGKNHEGSTMDHVQTQLQEQLCGKRYLLVLDDV
Sen. sequence10	(211)	DPRVRTAFPLRCWNCVSDQDQKKLDVKEILGKILSTATGKNHEGSTMDHVQTQLREQLCGKRYLLVLDDV
Sen. sequence11	(211)	DPRVRTAFPLRCWNCVSDQDQKKLDVKEILGKILSTATGKNHEGSTMDHVQTQLREQLCGKRYLLVLDDV
Sen. sequence12	(211)	DPRVRTAFPLRCWNCVSDQDQKKLDVKEILGKILSTATGKNHEGSTMDHVQTQLREQLCGKRYLLVLDDV
Sen. sequence13	(168)	DPRVRTAFPLRCWNCL*-----
Sen. sequence14	(211)	DPRVRTAFPLRCWNCVSDQDQKKLDVKEILGKILSTATGKNHEGSTMDHVQTQLREQLCGKRYLLVLDDV
Sen. sequence15	(211)	DPRVRTAFPLRCWNCVSDQDQKKLDVKEILGKILSTATGKNHEGSTMDHVQTQLREQLCGKRYLLVLDDV
Sen. sequence16	(211)	DPRVRTAFPLRCWNCVSDQDQKKLDVKEILGKILSTATGKNHEGSTMDHVQTQLREQLCGKRYLLVLDDV
Sen. sequence17	(211)	DPRVRTAFPLRCWNCVSDQDQKKLDVKEILGKILSTATGKNHEGSTMDHVQTQLQEQLCGKRYLLVLDDV
Sen. sequence18	(211)	DPRVRTAFPLRCWNCVSDQDQKQLDVKEILGKILSTATGKNHEGSTMDQVQTQLRE-LCGKRYLLVLDDV
Sen. sequence19	(211)	DPRVRTAFPLRCWNCVSDQDQKKLDVKEILGKILSTATGKNHEGSTMDHVQTQLREQLCGKRYLLVLDDV
Sen. sequence20	(211)	DPRVRTAFPLRCWNCVSDQDQKKLDVKEILGKILSTATGKNHEGSTMDHVQTQLREQLCGKRYLLVLDDV
Sen. sequence21	(211)	DPRVRTAFPLRCWNCVSDQDQNQLDVKEILGKILSTATGKNHKGSTMDQVQTQLREQLCGKRYLLVLDDV
Sen. sequence22	(211)	DPRVRTAFPLRCWNCVSDQDQNQLDVKEILGKILSTATGKNHKGSTMDQVQTQLREQLCGKRYLLVLDDV

FIG 2 E

		281		350
Res. sequence	(281)	WNENPNQLRDLVEFFMGGQ	IRNWIVVTTRSHETARI	IRDGPLHKLQGLSEENSWRLFVRWTFGSVQPKFP
Sen. sequence1	(281)	WNENPNQLR ~~Y~~ LV	EFFMGGQ	RGNWIVVTTRSHETARI
Sen. sequence2	(281)	WNENPNQLRXXXX	FFMGGQ	RGNWIVVTTRSHETTRI
Sen. sequence3	(281)	WNENPNQLRDLG	KKFFMGGQ	RGNWIVVTTRSHETTRI
Sen. sequence4	(281)	WNENPNQLRDX	VEFFMGGQ	RGNWIVVTTRSHETTRI
Sen. sequence5	(281)	WNENPNQLXXXX	FFMGGQ	RGNWIVVTTRSHETTRI
Sen. sequence6	(281)	WNENPNQLRIX	VEFFMGGQ	RGNWIVVTTRSHETARI
Sen. sequence7	(184)	-----		
Sen. sequence8	(281)	WNENPNQLRXXXX	FFMGGQ	RGNWIVVTTRSHETTRI
Sen. sequence9	(281)	WNENPNQLRDX	VEFFMGGQ	RGNWIVVTTRSHETTRI
Sen. sequence10	(281)	WNENPNQLR ~~Y~~ LV	EFFMGGQ	RGNWIVVTTRSHETARI
Sen. sequence11	(281)	WNENPNQLR ~~Y~~ LV	EFFMGGQ	RGNWIVVTTRSHETARI
Sen. sequence12	(281)	WNENPNQLXXXX	FFMGGQ	RGNWIVVTTRSHETARI
Sen. sequence14	(281)	WNENPNQLRIL	VEFFMGGQ	RGNWIVVTTRSHETARI
Sen. sequence15	(281)	WNENPNQLRXXXX	FFMGGQ	RGNWIVVTTRSHETARI
Sen. sequence16	(281)	WNENPNQLRXX	VEFFMGGQ	RGNWIVVTTRSHETARI
Sen. sequence17	(281)	WNENPNQLRDL	VEFFMGGQ	RGNWIVVTTRSHETTRI
Sen. sequence18	(280)	WNENPNQLRDL	VEFFMGGQ	RGNWIVVTTRSHETARI
Sen. sequence19	(281)	WNENPNQLRIL	VEFFMXXXXXXXXXXXXXXXXXXXXXXXXXXXX	QGLSEENSWRLFVRWTFGSVQPKFP
Sen. sequence20	(281)	WNENPNQLRDL	VEFFMGGQ	RGNWIVVTTRSHETARI
Sen. sequence21	(281)	WNENPNQLRDL	VEFFMXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX	
Sen. sequence22	(281)	WNENPNQLRXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX		

FIG 2 F

		351		420
Res. sequence	(351)	NDFVMIARDIVDKCARNPLAIRVVGSL	CGQDKSKWLSFHEICLANIRKSHNDIMPILNLSYHHLEPPIR	
Sen. sequence1	(351)	NDFVMIARDIVDKCARNPLAIRVVGSL	CGQDKSKWLSFHEICLANIRKSHNDIMPILNLSYHHLEPPIR	
Sen. sequence2	(351)	NDFVMIARDIVDKCARNPLAIRVVGSL	CGQDKSKWLSFHEICLANIRKSHNDIMLILNLSYHHLEPPIR	
Sen. sequence3	(351)	NDFVMIARDIVDKCARNPLAIRVVGSL	CGQDKSKWLSFHEICLANIRKSHNDIMLILNLSYHHLEPPIR	
Sen. sequence4	(351)	NDFVMIARDIVDKCARNPLAIRVVGSL	CGQDKSKWLSFHEICLANIRKSHNDIMLILNLSYHHLEPPIR	
Sen. sequence5	(351)	NDFVMIARDIVDKCARNPLAIRVVGSL	CGQDKSKWLSFHEICLANIRKSHNDIMLILNLSYHHLEPPIR	
Sen. sequence6	(351)	NDFVMIARDIVDKCARNPLAIRVVGSL	CGQDKSKWLSFHEICLANIRKSHNDIMPILNLSYHHLEPPIR	
Sen. sequence8	(351)	NDFVMIARDIVDKCARNPLAIRVVGSL	CGQDKSKWLSFHEICLANIRKSHNDIMLILNLSYHHLEPPIR	
Sen. sequence9	(351)	NDFVMIARDIVDKCARNPLAIRVVGSL	CGQDKSKWLSFHEICLANIRKSHNDIMLILNLSYHHLEPPIR	
Sen. sequence10	(351)	NDFVMIARDIVDKCARNPLAIRVVGSL	CGQDKSKWLSFHEICLANIRKSHNDIMPILNLSYHHLEPPIR	
Sen. sequence11	(351)	NDFVMIARDIVDKCARNPLAIRVVGSL	CGQDKSKWLSFHEICLANIRKSHNDIMPILNLSYHHLEPPIR	
Sen. sequence12	(351)	NDFVMIARDIVDKCARNPLAIRVVGSL	CGQDKSKWLSFHEICLANIRKSHNDIMPILNLSYHHLEPPIR	
Sen. sequence14	(351)	NDFVMIARDIVDKCARNPLAIRVVGSL	CGQDKSKWLSFHEICLANIRKSHNDIMPILNLSYHHLEPPIR	
Sen. sequence15	(351)	NDFVMIARDIVDKCARNPLAIRVVGSL	CGQDKSKWLSFHEICLANIRKSHNDIMPILNLSYHHLEPPIR	
Sen. sequence16	(351)	NDFVMIARDIVDKCARNPLAIRVVGSL	CGQDKSKWLSFHEICLANIRKSHNDIMPILNLSYHHLEPPIR	
Sen. sequence17	(351)	NDFVMIARDIVDKCARNPLAIRVVGSL	CGQDKSKWLSFHEICLANIRKSHNDIMLILNLSYHHLEPPIR	
Sen. sequence18	(350)	NDFVMIARDIVDKCARNPLAIRVVGSL	CGQDKSKWLSFHEICLANIRKSHNDIMPILNLSYHHLEPPIR	
Sen. sequence19	(351)	NDFVMIARDIVDKCARNPLAIRVVGSL	CGQDKSKWLSFHEICLANIRKSHNDIMPILNLSYHHLEPPIR	
Sen. sequence20	(351)	NDFVMIARDIVDKCARNPLAIRVVGSL	CGQDKSKWLSFHEICLANIRKSHNDIMPILNLSYHHLEPPIR	
Sen. sequence21	(333)	XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX		
Sen. sequence22	(351)	XXXXXXXXXXXXDKCARNPLAIRVVGSL	CGQDKSKWLSFHEICLANIRKSHNDIMPILNLSYHHLEPPIR	

FIG 2 G

	421	490
Res. sequence	(421)	RCFSYCA ~~Y~~ FPKDFLIGK ~~T~~ TLINLWMAQGYIVPLDKDQSIDDASEEYISILL ~~R~~ RCFFEN ~~T~~ G ~~A~~ EKDGV ~~I~~ KI ~~H~~
Sen. sequence1	(421)	RCFSYCAMFPKDFLIGKKTLINLWMAQGYIVPLDKDQSIDDASEEYISILLQRCFFENIGTEKD ~~Y~~ VIKI ~~H~~
Sen. sequence2	(421)	RCFSYCAMFPKDFLIGKKTLINLWMAQGYIVPLDKDQSIDDASEEYISILLQRCFFENIGTEKD ~~Y~~ VIKI ~~H~~
Sen. sequence3	(421)	RCFSYCAMFPKDFLIGKKTLINLWMAQGYIVPLDKDQSIDDASEEYISILLQRCFFENIGTEKD ~~Y~~ VIKI ~~H~~
Sen. sequence4	(421)	RCFSYCAMFPKDFLIGKKTLINLWMAQGYIVPLDKDQSIDDASEEYISILLQRCFFENIGTEKD ~~Y~~ VIKI ~~H~~
Sen. sequence5	(421)	RCFSYCAMFPKDFLIGKKTLINLWMAQGYIVPLDKDQSIDDASEEYISILLQRCFFENIGTEKD ~~Y~~ VIKI ~~H~~
Sen. sequence6	(421)	RCFSYCAMFPKDFLIGKKTLINLWMAQGYIVPLDKDQSIDDASEEYISILLQRCFFENIGTEKD ~~Y~~ VIKI ~~H~~
Sen. sequence8	(421)	RCFSYCAMFPKDFLIGKKTLINLWMAQGYIVPLDKDQSIDDASEEYISILLQRCFFENIGTEKD ~~Y~~ VIKI ~~H~~
Sen. sequence9	(421)	RCFSYCAMFPKDFLIGKKTLINLWMAQGYIVPLDKDQSIDDASEEYISILLQRCFFENIGTEKD ~~Y~~ VIKI ~~H~~
Sen. sequence10	(421)	RCFSYCAMFPKDFLIGKKTLINLWMAQGYIVPLDKDQSIDDASEEYISILLQRCFFENIGTEKD ~~Y~~ VIKI ~~H~~
Sen. sequence11	(421)	RCFSYCAMFPKDFLIGKKTLINLWMAQGYIVPLDKDQSIDDASEEYISILLQRCFFENIGTEKD ~~Y~~ VIKI ~~H~~
Sen. sequence12	(421)	RCFSYCAMFPKDFLIGKKTLINLWMAQGYIVPLDKDQSIDDASEEYISILLQRCFFENIGTEKD ~~Y~~ VIKI ~~H~~
Sen. sequence14	(421)	RCFSYCAMFPKDFLIGKKTLINLWMAQGYIVPLDKDQSIDDASEEYISILLQRCFFENIGTEKD ~~Y~~ VIKI ~~H~~
Sen. sequence15	(421)	RCFSYCAMFPKDFLIGKKTLINLWMAQGYIVPLDKDQSIDDASEEYISILLQRCFFENIGTEKD ~~Y~~ VIKI ~~H~~
Sen. sequence16	(421)	RCFSYCAMFPKDFLIGKKTLINLWMAQGYIVPLDKDQSIDDASEEYISILLQRCFFENIGTEKD ~~Y~~ VIKI ~~H~~
Sen. sequence17	(421)	RCFSYCAMFPKDFLIGKKTLINLWMAQGYIVPLDKDQSIDDASEEYISILLQRCFFENIGTEKD ~~Y~~ VIKI ~~H~~
Sen. sequence18	(420)	RCFSYCVVFPKDFLIGKKTLINLWMAQGYIVPLDKDQSIDDASEEYISILLQRCFFENIGTEKDD ~~V~~ IKI ~~H~~
Sen. sequence19	(421)	RCFSYCAMFPKDFLIGKKTLINLWMAQGYIVPLDKDQSIDDASEEYISILLQRCFFENIGTEKD ~~Y~~ VIKI ~~H~~
Sen. sequence20	(421)	RCFSYCAMFPKDFLIGKKTLINLWMAQGYIVPLDKDQSIDDASEEYISILLQRCFFENIGTEKD ~~Y~~ VIKI ~~H~~
Sen. sequence21	(378)	XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
Sen. sequence22	(421)	RCFSYCAVFPKDFLIGKKTLINLWMAQGYIVPLDKDQSIDDASEEYISILXQRCFFENIGTEKDD ~~V~~ IKI ~~H~~

FIG 2 H

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

FIG 2 I

		561	630
Res. sequence	(561)	ALVPNCLCLRALDLAWSKIKSLPDSIGGLLHLRYLDLSYNEDLEVLPNSIAKLYNLQTLQLKGCKRLEGL	
Sen. sequence1	(561)	ALVPNCLCLRALDLAWSKIKSVDPDSIGGLLHLRYLDLSYNEDLEVLPNSIAKLYNLQTLQLKGCKRLEGL	
Sen. sequence2	(561)	ALVPNCLCLRALDLAWSKIKSVDPDSIGGLLHLRYLDLSYNEDLEVLPNSIAKLYNLQTLQLKGCKRLEGL	
Sen. sequence3	(561)	ALVPNCLCLRALDLAWSKIKSVDPDSIGGLLHLRYLDLSYNEDLEVLPNSIAKLYNLQTLQLKGCKRLEGL	
Sen. sequence4	(561)	ALVPNCLCLRALDLAWSKIKSVDPDSIGGLLHLRYLDLSYNEDLEVLPNSIAKLYNLQTLQLKGCKRLEGL	
Sen. sequence5	(561)	ALVPNCLCLRALDLAWSKIKSVDPDSIGGLLHLRYLDLSYNEDLEVLPNSIAKLYNLQTLQLKGCKRLEGL	
Sen. sequence6	(561)	ALVPNCLCLRALDLAWSKIKSVDPDSIGGLLHLRYLDLSYNEDLEVLPNSIAKLYNLQTLQLKGCKRLEGL	
Sen. sequence8	(561)	ALVPNCLCLRALDLAWSKIKSVDPDSIGGLLHLRYLDLSYNEDLEVLPNSIAKLYNLQTLQLKGCKRLEGL	
Sen. sequence9	(561)	ALVPNCLCLRALDLAWSKIKSVDPDSIGGLLHLRYLDLSYNEDLEVLPNSIAKLYNLQTLQLKGCKRLEGL	
Sen. sequence10	(561)	ALVPNCLCLRALDLAWSKIKSVDPDSIGGLLHLRYLDLSYNEDLEVLPNSIAKLYNLQTLQLKGCKRLEGL	
Sen. sequence11	(561)	ALVPNCLCLRALDLAWSKIKSVDPDSIGGLLHLRYLDLSYNEDLEVLPNSIAKLYNLQTLQLKGCKRLEGL	
Sen. sequence12	(561)	ALVPNCLCLRALDLAWSKIKSVDPDSIGGLLHLRYLDLSYNEDLEVLPNSIAKLYNLQTLQLKGCKRLEGL	
Sen. sequence14	(561)	ALVPNCLCLRALDLAWSKIKSVDPDSIGGLLHLRYLDLSYNEDLEVLPNSIAKLYNLQTLQLKGCKRLEGL	
Sen. sequence15	(561)	ALVPNCLCLRALDLAWSKIKSVDPDSIGGLLHLRYLDLSYNEDLEVLPNSIAKLYNLQTLQLKGCKRLEGL	
Sen. sequence16	(561)	ALVPNCLCLRALDLAWSKIKSVDPDSIGGLLHLRYLDLSYNEDLEVLPNSIAKLYNLQTLQLKGCKRLEGL	
Sen. sequence17	(561)	ALVPNCLCLRALDLAWSKIKSVDPDSIGGLLHLRYLDLSYNEDLEVLPNSIAKLYNLQTLQLKGCKRLEGL	
Sen. sequence18	(560)	ALVPNCLCLRALDLAWSKIKSVDPDSIGGLLHLRYLDLSYNEDLEVLPNSIAKLYNLQTLQLKGCKRLEGL	
Sen. sequence19	(561)	ALVPNCLCLRALDLAWSKIKSVDPDSIGGLLHLRYLDLSYNEDLEVLPNSIAKLYNLQTLQLKGCKRLEGL	
Sen. sequence20	(561)	ALVPNCLCLRALDLAWSKIKSVDPDSIGGLLHLRYLDLSYNEDLEVLPNSIAKLYNLQTLQLKGCKRLEGL	
Sen. sequence21	(504)	ALVPNCLCLRALDLAWSKIKSLPDSIGGLLHLRYLDLSYNEDLEVLPNSIAKLYNLQTLQLKGCKRLEGL	
Sen. sequence22	(533)	XXXXNCLCLRALDLAWSKIKSLPDSIGGLLHLRYLDLSYNEDLEVLPNSIAKLYNLQTLQLKGCKRLEGL	

FIG 2 J

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

FIG 2 K

	701	770
Res. sequence	(701)	LEIQIRWPENTTDAVKEDVKREGLYLNHKEHLNHIVVDFRCEE ~~GG~~ GRMDDEEARRLMEELRPHPYLENLA
Sen. sequence1	(701)	LEIQIRWPENTTDAVKEDVKREGLYLNHKEHLNHIVVDFRCEE ~~GG~~ GRMDDEEARRLMEELRPHPYLENLA
Sen. sequence2	(701)	LEIQIRWPENTTDAVKEDVKREGLYLNHKEHLNHIVVD-----
Sen. sequence3	(701)	LEIQIRWPENTTDAVKEDVKREGLYLNHKEHLNHIVVD-----
Sen. sequence4	(701)	LEIQIRWPENTTDAVKEDVKREGLYLNHKEHLNHIVVD-----
Sen. sequence5	(701)	LEIQIRWPENTTDAVKEDVKREGLYLNH-----
Sen. sequence6	(701)	LEIQIRWPENTTDAVKEDVKREGLYLN-----
Sen. sequence8	(701)	LEIQIRWPENTTDAVKEDVKREGLYLNH-----
Sen. sequence9	(701)	LEIQIRWPENTTDAVKEDVKREGLYLNH-----
Sen. sequence10	(701)	LEIQIRWPENTTDAVKEDVKREGLYLNHKEHLNHIVVDFRCEE ~~GG~~ GRMDDEEARRLMEELRPHPYLENLA
Sen. sequence11	(701)	LEIQIRWPENTTDAVKEDVKREGLYLNHKEHLNHIVVD-----
Sen. sequence12	(701)	LEIQIRWPENTTDAVKEDVKREGLYLNHKEHLNHIVVD-----
Sen. sequence14	(701)	LEIQIRWPENTTDAVKEDVKREGLYLNHKEHLNHIVVD-----
Sen. sequence15	(701)	LEIQIRWPENTTDAVKEDVKREGLYLNHKEHLNHIVVD-----
Sen. sequence16	(701)	LEIQIRWPENTTDAVKEDVKREGLYLNHKEHLNHIVVD-----
Sen. sequence17	(701)	LEIQIRWPENTTDAVKEDVKREGLYLNHKEHLNHIVVD-----
Sen. sequence18	(700)	LEIQIRWPENTTDAVKEDVKREGLYLNHKEHLNHIVVD-----
Sen. sequence19	(701)	LEIQIRWPENTTDAVKEDVKREGLYLNHKEHLNHIVVD-----
Sen. sequence20	(701)	LEIQIRWPENTTDAVKEDVKREGLYLNHKEHLNHIVVD-----
Sen. sequence21	(644)	LEIQIRWPENTTDAVKEDVKREGLYLNH-----
Sen. sequence22	(669)	LEIQIRWPENTTDAVKEDVKREGLYLNHKEHLNHIVVD-----

FIG 2 L

	771		840
Res. sequence	(771)	VKAYYG <u>V</u> KMPGWATLLPNLT <u>E</u> FLSDCGE <u>L</u> ENLP <u>C</u> LGNLDHLKVLRLSHLAKLEYIEEDSSSAN <u>F</u> RC <u>R</u> PG	
Sen. sequence1	(771)	VKAYYGAKMPGWATLLPNLT <u>E</u> LYLSDCGE <u>S</u> ECLPCMGNLDCLKVLRLSHLAKLEYIEEDSTSAN <u>F</u> S <u>F</u> RP <u>G</u>	
Sen. sequence10	(771)	VKAYYGAKMPGWATLLPNLT <u>E</u> LYLSDCGE <u>S</u> ECLPCMGNLDCLKVLRLSHLAKLEYIEEDSTSAN <u>F</u> S <u>F</u> RP <u>G</u>	
	841		910
Res. sequence	(841)	PESAGLSLYFPSLE <u>R</u> LE <u>L</u> K <u>R</u> L <u>C</u> KLKGWRR <u>G</u> EGLGDDHQPFNESSS> <u>Ret. Ins</u> <NTQVQLQLCLPQLKSLRIERCPLLT	
Sen. sequence1	(841)	PESAGLSLYFPSLE <u>L</u> LE <u>L</u> ELKRLHKLKGWRR <u>R</u> EGLGDDHQPFNESSS> <u>Ret. Ins</u> <-----	
Sen. sequence10	(841)	PESAGLSLYFPSLE <u>L</u> LE <u>L</u> ELKRLHKLKGWRR <u>R</u> EGLGDDHQPFNESSS> <u>Ret. Ins</u> <-----	
	911		980
Res. sequence	(911)	FMPLCPKTEKLHLVVFNERLRIVHAKRDENFYAPLHSSSSDPENPRNTIPIPMFREVYINNVAVLNSLPM	
	981		1050
Res. sequence	(981)	EAFRCLTHMTIKNDEVESLGEVGEVFRSCSSSLRSLNITGCSNLRSVSGGLEHLTALEMLEIYDTHKLSL	
	1051		1120
Res. sequence	(1051)	SEDPEGVVPWKSLSLHLSYLQLMNLPQLVNLPSMQFLAALRTLIVHCTKLQSVDPDWMPLRLSLRKLMV	
	1121		1163
Res. sequence	(1121)	SFCSAHLERRCQNPTGVDWPNIQHIPSIDVTSSLPKFLVLPYE	

FIG 3 A

	1	70
Res. sequence	(1) MERVVYRWNTAETAGHLWPTNP	LILIPYSALQSSSELKEILSIFGYKSQ
Sen. sequence1	(1) MERVVYRWN---TAGHLWPTNQ	LILIPY-ALQSSSELKEILSIFGYKSQ
Sen. sequence2	(1) MERVVYRWNTAETAGHLWPTNP	LILIPYSALQSSSELKEILSIFGYKSQ
Sen. sequence3	(1) MERVVYRWNTAETAGHLWPTNP	LILIPYSALQSSSELKEILSIFGYKSQ
Sen. sequence4	(1) MERVVYRWNTAETAGHLWPTNP	LILIPYSALQSSSELKEILSIFGYKSQ
Sen. sequence5	(1) MERVVYRWNTAETAGHLWPTNP	LILIPYSALQSSSELKEILSIFGYKSQ
Sen. sequence6	(1) MERVVYRWN---TAGHLWPTNQ	LILIPYSALQSSSELKEILSIFGYKSQ
Sen. sequence7	(1) -----	STINAVFRDAETKQELTHEAQHWLEEL
Sen. sequence8	(1) MERVVYRWNTAETAGHLWPTNP	LILIPYSALQSSSELKEILSIFGYKSQ
Sen. sequence9	(1) MERVVYRWNTAETAGHLWPTNP	LILIPYSALQSSSELKEILSIFGYKSQ
Sen. sequence10	(1) MERVVYRWN---TAGHLWPTNQ	LILIPY-ALQSSSELKEILSIFGYKSQ
Sen. sequence11	(1) MERVVYRWN---TAGHLWPTNQ	LILIPYSALQSSSELKEILSIFGYKSQ
Sen. sequence12	(1) MERVVYRWN---TAGHLWPTNQ	LILIPYSALQSSSELKEILSIFGYKSQ
Sen. sequence13	(1) -----	STINAVFRDAETKQELTHEAQHWLEEL
Sen. sequence14	(1) MERVVYRWN---TAGHLWPTNQ	LILIPYSALQSSSELKEILSIFGYKSQ
Sen. sequence15	(1) MERVVYRWN---TAGHLWPTNQ	LILIPYSALQSSSELKEILSIFGYKSQ
Sen. sequence16	(1) MERVVYRWN---TAGHLWPTNQ	LILIPYSALQSSSELKEILSIFGYKSQ
Sen. sequence17	(1) MERVVYRWNTAETAGHLWPTNP	LILIPYSALQSSSELKEILSIFGYKSQ
Sen. sequence18	(1) MERVVYRWNTAETAGHLWPTNP	LILIPYSALQSSSELKEILSIFGYKSQ
Sen. sequence19	(1) MERVVYRWN---TAGHLWPTNQ	LILIPYSALQSSSELKEILSIFGYKSQ
Sen. sequence20	(1) MERVVYRWN---TAGHLWPTNQ	LILIPYSALQSSSELKEILSIFGYKSQ
Sen. sequence21	(1) -----	ALQSSSELKEILSIFGYKSQ
Sen. sequence22	(1) -----	ALQSSSELKEILSIFGYKSQ

FIG 3 B

	71	140
Res. sequence	(71)	KDAVFEADDLFDEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYRMSRGVKKIKKKLDAIAYNHQF
Sen. sequence1	(71)	KDAVFEADDLFDEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYRMSRGVKKIKKKLDAIAYNHQF
Sen. sequence2	(71)	KDAVFEADDLFDEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYRMSRGVKKIKKKLDAIAYNHQF
Sen. sequence3	(71)	KDAVFEADDLFDEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYRMSRGVKKIKKKLDAIAYNHQF
Sen. sequence4	(71)	KDAVFEADDLFDEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYRMSRGVKKIKKKLDAIAYNHQF
Sen. sequence5	(71)	KDAVFEADDLFDEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYRMSRGVKKIKKKLDAIAYNHQF
Sen. sequence6	(71)	KDAVFEADDLFDEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYRMSRGVKKIKKKLDAIAYNHQF
Sen. sequence7	(28)	KDAVFEADDLFDEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYRMSRGVKKIKKKLDAIAYNHQF
Sen. sequence8	(71)	KDAVFEADDLFDEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYRMSRGVKKIKKKLDAIAYNHQF
Sen. sequence9	(71)	KDAVFEADDLFDEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYRMSRGVKKIKKKLDAIAYNHQF
Sen. sequence10	(71)	KDAVFEADDLFDEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYRMSRGVKKIKKKLDAIAYNHQF
Sen. sequence11	(71)	KDAVFEADDLFDEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYRMSRGVKKIKKKLDAIAYNHQF
Sen. sequence12	(71)	KDAVFEADDLFDEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYRMSRGVKKIKKKLDAIAYNHQF
Sen. sequence13	(28)	KDAVFEADDLFDEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYRMSRGVKKIKKKLDAIAYNHQF
Sen. sequence14	(71)	KDAVFEADDLFDEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYRMSRGVKKIKKKLDAIAYNHQF
Sen. sequence15	(71)	KDAVFEADDLFDEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYRMSRGVKKIKKKLDAIAYNHQF
Sen. sequence16	(71)	KDAVFEADDLFDEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYRMSRGVKKIKKKLDAIAYNHQF
Sen. sequence17	(71)	KDAVFEADDLFDEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYRMSRGVKKIKKKLDAIAYNHQF
Sen. sequence18	(71)	KDAVFEADDLFDEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYRMSRGLKKIKKKLDAIAYNHQF
Sen. sequence19	(71)	KDAVFEADDLFDEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYRMSRGVKKIKKKLDAIAYNHQF
Sen. sequence20	(71)	KDAVFEADDLFDEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYRMSRGVKKIKKKLDAIAYNHQF
Sen. sequence21	(71)	KDAVFEADDLFDEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYKMSQG _Q VKKIKKKLDV _I AYNHQF
Sen. sequence22	(71)	KDAVFEADDLFDEFVTLAEQKQLVEAGGSLSKKMRQFFSDSNPLGIAYKMSQG _Q VKKIKKKLDV _I AYNHQF

FIG 3 C

	141	210
Res. sequence	(141)	SFKIDLEPIKERRLETGSVVNAGDIIIGREDDLEKIVGLLLDSNIQRDVSFLTIVGMGGLGKTALAQLVYN
Sen. sequence1	(141)	SFKIDLEPIKERRLETGSVVNAGDIIIGREDDLEKIVGLLLDSNIQRDVSFLTIVGMGGLGKTALAQLVYN
Sen. sequence2	(141)	SFKIDLEPIKERRLETGSVVNAGDIIIGREDDLEKIVGLLLDSNIQRDVSFLTIVGMGGLGKTALAQLVYN
Sen. sequence3	(141)	SFKIDLEPIKERRLETGSVVNAGDIIIGREDDLEKIVGLLLDSNIQRDVSFLTIVGMGGLGKTALAQLVYN
Sen. sequence4	(141)	SFKIDLEPIKERRLETGSVVNAGDIIIGREDDLEKIVGLLLDSNIQRDVSFLTIVGMGGLGKTALAQLVYN
Sen. sequence5	(141)	SFKIDLEPIKERRLETGSVVNAGDIIIGREDDLEKIVGLLLDSNIQRDVSFLTIVGMGGLGKTALAQLVYN
Sen. sequence6	(141)	SFKIDLEPIKERRLETGSVVNAGDIIIGREDDLEKIVGLLLDSNIQRDVSFLTIVGMGGLGKTALAQLVYN
Sen. sequence7	(98)	SFKIDLEPIKERRLETGSVVNAGDIIIGREDDLEKIVGLLLDSNIQRDVSFLTIVGMGGLGKTALAQLVYN
Sen. sequence8	(141)	SFKIDLEPIKERRLETGSVVNAGDIIIGREDDLEKIVGLLLDSNIQRDVSFLTIVGMGGLGKTALAQLVYN
Sen. sequence9	(141)	SFKIDLEPIKERRLETGSVVNAGDIIIGREDDLEKIVGLLLDSNIQRDVSFLTIVGMGGLGKTALAQLVYN
Sen. sequence10	(141)	SFKIDLEPIKERRLETGSVVNAGDIIIGREDDLEKIVGLLLDSNIQRDVSFLTIVGMGGLGKTALAQLVYN
Sen. sequence11	(141)	SFKIDLEPIKERRLETGSVVNAGDIIIGREDDLEKIVGLLLDSNIQRDVSFLTIVGMGGLGKTALAQLVYN
Sen. sequence12	(141)	SFKIDLEPIKERRLETGSVVNAGDIIIGREDDLEKIVGLLLDSNIQRDVSFLTIVGMGGLGKTALAQLVYN
Sen. sequence13	(98)	SFKIDLEPIKERRLETGSVVNAGDIIIGREDDLEKIVGLLLDSNIQRDVSFLTIVGMGGLGKTALAQLVYN
Sen. sequence14	(141)	SFKIDLEPIKERRLETGSVVNAGDIIIGREDDLEKIVGLLLDSNIQRDVSFLTIVGMGGLGKTALAQLVYN
Sen. sequence15	(141)	SFKIDLEPIKERRLETGSVVNAGDIIIGREDDLEKIVGLLLDSNIQRDVSFLTIVGMGGLGKTALAQLVYN
Sen. sequence16	(141)	SFKIDLEPIKERRLETGSVVNAGDIIIGREDDLEKIVGLLLDSNIQRDVSFLTIVGMGGLGKTALAQLVYN
Sen. sequence17	(141)	SFKIDLEPIKERRLETGSVVNAGDIIIGREDDLEKIVGLLLDSNIQRDVSFLTIVGMGGLGKTALAQLVYN
Sen. sequence18	(141)	SFKIDLEPIKERRLETGSVVNAGDIIIGREDDLEKIVGLLLDSNIQRDMSFLTIVGMGGLGKTALAQLVYN
Sen. sequence19	(141)	SFKIDLEPIKERRLETGSVVNAGDIIIGREDDLEKIVGLLLDSNIQRDV _S FLTIVGMGGLGKTALAQLVYN
Sen. sequence20	(141)	SFKIDLEPIKERRLETGSVVNAGDIIIGREDDLEKIVGLLLDSNIQRDVSFLTIVGMGGLGKTALAQLVYN
Sen. sequence21	(141)	SFKIDLEPIKERRLETGSVVNAGDIIIGREDDLEKIVGL <u>F</u> LD _S NIQRDVSFLTIVGMGGLGKTALAQLVYN
Sen. sequence22	(141)	SFKIDLEPIKERRLETGSVVNAGDIIIGREDDLEKIVGL <u>F</u> LD _S NIQRDVSFLTIVGMGGLGKTALAQLVYN

FIG 3 D

	211	280
Res. sequence	(211)	DPRVRTAFPLRCWNCVSDQDQKQLDVKEILGKILSTATGKNHEGSTMDQVQTQLREQLCGKRYLLVLDDV
Sen. sequence1	(211)	DPRVRTAFPLRCWNCVSDQDQKKLDVKEILGKILSTATGKNHEGSTMDHVQTQLREQLCGKRYLLVLDDV
Sen. sequence2	(211)	DPRVRTAFPLRCWNCVSDQDQKKLDVKEILGKILSTATGKNHEGSTMDHVQTQLQEQLCGKRYLLVLDDV
Sen. sequence3	(211)	DPRVRTAFPLRCWNCVSDQDQKKLDVKEILGKILSTATGKNHEGSTMDHVQTQLQEQLCGKRYLLVLDDV
Sen. sequence4	(211)	DPRVRTAFPLRCWNCVSDQDQKKLDVKEILGKILSTATGKNHEGSTMDHVQTQLQEQLCGKRYLLVLDDV
Sen. sequence5	(211)	DPRVRTAFPLRCWNCVSDQDQKKLDVKEILGKILSTATGKNHEGSTMDHVQTQLQEQLCGKRYLLVLDDV
Sen. sequence6	(211)	DPRVRTAFPLRCWNCVSDQDQKKLDVKEILGKILSTATGKNHEGSTMDHVQTQLREQLCGKRYLLVLDDV
Sen. sequence7	(168)	DPRVRTAFPLRCWNCL*-----
Sen. sequence8	(211)	DPRVRTAFPLRCWNCVSDQDQKKLDVKEILGKILSTATGKNHEGSTMDHVQTQLQEQLCGKRYLLVLDDV
Sen. sequence9	(211)	DPRVRTAFPLRCWNCVSDQDQKKLDVKEILGKILSTATGKNHEGSTMDHVQTQLQEQLCGKRYLLVLDDV
Sen. sequence10	(211)	DPRVRTAFPLRCWNCVSDQDQKKLDVKEILGKILSTATGKNHEGSTMDHVQTQLREQLCGKRYLLVLDDV
Sen. sequence11	(211)	DPRVRTAFPLRCWNCVSDQDQKKLDVKEILGKILSTATGKNHEGSTMDHVQTQLREQLCGKRYLLVLDDV
Sen. sequence12	(211)	DPRVRTAFPLRCWNCVSDQDQKKLDVKEILGKILSTATGKNHEGSTMDHVQTQLREQLCGKRYLLVLDDV
Sen. sequence13	(168)	DPRVRTAFPLRCWNCL*-----
Sen. sequence14	(211)	DPRVRTAFPLRCWNCVSDQDQKKLDVKEILGKILSTATGKNHEGSTMDHVQTQLREQLCGKRYLLVLDDV
Sen. sequence15	(211)	DPRVRTAFPLRCWNCVSDQDQKKLDVKEILGKILSTATGKNHEGSTMDHVQTQLREQLCGKRYLLVLDDV
Sen. sequence16	(211)	DPRVRTAFPLRCWNCVSDQDQKKLDVKEILGKILSTATGKNHEGSTMDHVQTQLREQLCGKRYLLVLDDV
Sen. sequence17	(211)	DPRVRTAFPLRCWNCVSDQDQKKLDVKEILGKILSTATGKNHEGSTMDHVQTQLQEQLCGKRYLLVLDDV
Sen. sequence18	(211)	DPRVRTAFPLRCWNCVSDQDQKQLDVKEILGKILSTATGKNHEGSTMDQVQTQLRE-LCGKRYLLVLDDV
Sen. sequence19	(211)	DPRVRTAFPLRCWNCVSDQDQKKLDVKEILGKILSTATGKNHEGSTMDHVQTQLREQLCGKRYLLVLDDV
Sen. sequence20	(211)	DPRVRTAFPLRCWNCVSDQDQKKLDVKEILGKILSTATGKNHEGSTMDHVQTQLREQLCGKRYLLVLDDV
Sen. sequence21	(211)	DPRVRTAFPLRCWNCVSDQDQNQLDVKEILGKILSTATGKNHKGSTMDQVQTQLREQLCGKRYLLVLDDV
Sen. sequence22	(211)	DPRVRTAFPLRCWNCVSDQDQNQLDVKEILGKILSTATGKNHKGSTMDQVQTQLREQLCGKRYLLVLDDV

FIG 3 E

		281		350
Res. sequence	(281)	WNENPNQLRDLVEFFMGGQ	IRNWIVVTTRSHETARI	IRDGPLHKLQGLSEENSWRLFVRWTFGSVQPKFP
Sen. sequence1	(281)	WNENPNQLR ~~Y~~ LV	EFFMGGQ	RGNWIVVTTRSHETARI
Sen. sequence2	(281)	WNENPNQLRXXXX	FFMGGQ	RGNWIVVTTRSHETTRI
Sen. sequence3	(281)	WNENPNQLRDLG	KKFFMGGQ	RGNWIVVTTRSHETTRI
Sen. sequence4	(281)	WNENPNQLRDX	VEFFMGGQ	RGNWIVVTTRSHETTRI
Sen. sequence5	(281)	WNENPNQLXXXX	FFMGGQ	RGNWIVVTTRSHETTRI
Sen. sequence6	(281)	WNENPNQLRIX	VEFFMGGQ	RGNWIVVTTRSHETARI
Sen. sequence7	(184)	-----		
Sen. sequence8	(281)	WNENPNQLRXXXX	FFMGGQ	RGNWIVVTTRSHETTRI
Sen. sequence9	(281)	WNENPNQLRDX	VEFFMGGQ	RGNWIVVTTRSHETTRI
Sen. sequence10	(281)	WNENPNQLR ~~Y~~ LV	EFFMGGQ	RGNWIVVTTRSHETARI
Sen. sequence11	(281)	WNENPNQLR ~~Y~~ LV	EFFMGGQ	RGNWIVVTTRSHETARI
Sen. sequence12	(281)	WNENPNQLXXXX	FFMGGQ	RGNWIVVTTRSHETARI
Sen. sequence14	(281)	WNENPNQLRIL	VEFFMGGQ	RGNWIVVTTRSHETARI
Sen. sequence15	(281)	WNENPNQLRXXXX	FFMGGQ	RGNWIVVTTRSHETARI
Sen. sequence16	(281)	WNENPNQLRXX	VEFFMGGQ	RGNWIVVTTRSHETARI
Sen. sequence17	(281)	WNENPNQLRDL	VEFFMGGQ	RGNWIVVTTRSHETTRI
Sen. sequence18	(280)	WNENPNQLRDL	VEFFMGGQ	RGNWIVVTTRSHETARI
Sen. sequence19	(281)	WNENPNQLRIL	VEFFMXXXXXXXXXXXXXXXXXXXXXXXXXXXX	QGLSEENSWRLFVRWTFGSVQPKFP
Sen. sequence20	(281)	WNENPNQLRDL	VEFFMGGQ	RGNWIVVTTRSHETARI
Sen. sequence21	(281)	WNENPNQLRDL	VEFFMXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX	
Sen. sequence22	(281)	WNENPNQLRXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX		

FIG 3 F

		351		420
Res. sequence	(351)	NDFVMIARDIVDKCARNPLAIRVVGSL	CGQDKSKWLSFHEICLANIRKSHNDIMPILNLSYHHLEPPIR	
Sen. sequence1	(351)	NDFVMIARDIVDKCARNPLAIRVVGSL	CGQDKSKWLSFHEICLANIRKSHNDIMPILNLSYHHLEPPIR	
Sen. sequence2	(351)	NDFVMIARDIVDKCARNPLAIRVVGSL	CGQDKSKWLSFHEICLANIRKSHNDIMLILNLSYHHLEPPIR	
Sen. sequence3	(351)	NDFVMIARDIVDKCARNPLAIRVVGSL	CGQDKSKWLSFHEICLANIRKSHNDIMLILNLSYHHLEPPIR	
Sen. sequence4	(351)	NDFVMIARDIVDKCARNPLAIRVVGSL	CGQDKSKWLSFHEICLANIRKSHNDIMLILNLSYHHLEPPIR	
Sen. sequence5	(351)	NDFVMIARDIVDKCARNPLAIRVVGSL	CGQDKSKWLSFHEICLANIRKSHNDIMLILNLSYHHLEPPIR	
Sen. sequence6	(351)	NDFVMIARDIVDKCARNPLAIRVVGSL	CGQDKSKWLSFHEICLANIRKSHNDIMPILNLSYHHLEPPIR	
Sen. sequence8	(351)	NDFVMIARDIVDKCARNPLAIRVVGSL	CGQDKSKWLSFHEICLANIRKSHNDIMLILNLSYHHLEPPIR	
Sen. sequence9	(351)	NDFVMIARDIVDKCARNPLAIRVVGSL	CGQDKSKWLSFHEICLANIRKSHNDIMLILNLSYHHLEPPIR	
Sen. sequence10	(351)	NDFVMIARDIVDKCARNPLAIRVVGSL	CGQDKSKWLSFHEICLANIRKSHNDIMPILNLSYHHLEPPIR	
Sen. sequence11	(351)	NDFVMIARDIVDKCARNPLAIRVVGSL	CGQDKSKWLSFHEICLANIRKSHNDIMPILNLSYHHLEPPIR	
Sen. sequence12	(351)	NDFVMIARDIVDKCARNPLAIRVVGSL	CGQDKSKWLSFHEICLANIRKSHNDIMPILNLSYHHLEPPIR	
Sen. sequence14	(351)	NDFVMIARDIVDKCARNPLAIRVVGSL	CGQDKSKWLSFHEICLANIRKSHNDIMPILNLSYHHLEPPIR	
Sen. sequence15	(351)	NDFVMIARDIVDKCARNPLAIRVVGSL	CGQDKSKWLSFHEICLANIRKSHNDIMPILNLSYHHLEPPIR	
Sen. sequence16	(351)	NDFVMIARDIVDKCARNPLAIRVVGSL	CGQDKSKWLSFHEICLANIRKSHNDIMPILNLSYHHLEPPIR	
Sen. sequence17	(351)	NDFVMIARDIVDKCARNPLAIRVVGSL	CGQDKSKWLSFHEICLANIRKSHNDIMLILNLSYHHLEPPIR	
Sen. sequence18	(350)	NDFVMIARDIVDKCARNPLAIRVVGSL	CGQDKSKWLSFHEICLANIRKSHNDIMPILNLSYHHLEPPIR	
Sen. sequence19	(351)	NDFVMIARDIVDKCARNPLAIRVVGSL	CGQDKSKWLSFHEICLANIRKSHNDIMPILNLSYHHLEPPIR	
Sen. sequence20	(351)	NDFVMIARDIVDKCARNPLAIRVVGSL	CGQDKSKWLSFHEICLANIRKSHNDIMPILNLSYHHLEPPIR	
Sen. sequence21	(333)	XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX		
Sen. sequence22	(351)	XXXXXXXXXXXXDKCARNPLAIRVVGSL	CGQDKSKWLSFHEICLANIRKSHNDIMPILNLSYHHLEPPIR	

FIG 3 G

	421		490
Res. sequence	(421)	RCFSYCA ~~Y~~ FPKDFLIGK ~~T~~ TLINLWMAQGYIVPLDKDQSIDDASEEYISILL ~~R~~ RCFFEN ~~T~~ G ~~A~~ EKDGV ~~I~~ KI ~~H~~	
Sen. sequence1	(421)	RCFSYCAMFPKDFLIGKKT ~~L~~ INLWMAQGYIVPLDKDQSIDDASEEYISILLQRCFFENIGTEKD ~~Y~~ VIK ~~I~~ H	
Sen. sequence2	(421)	RCFSYCAMFPKDFLIGKKT ~~L~~ INLWMAQGYIVPLDKDQSIDDASEEYISILLQRCFFENIGTEKD ~~Y~~ VIK ~~I~~ H	
Sen. sequence3	(421)	RCFSYCAMFPKDFLIGKKT ~~L~~ INLWMAQGYIVPLDKDQSIDDASEEYISILLQRCFFENIGTEKD ~~Y~~ VIK ~~I~~ H	
Sen. sequence4	(421)	RCFSYCAMFPKDFLIGKKT ~~L~~ INLWMAQGYIVPLDKDQSIDDASEEYISILLQRCFFENIGTEKD ~~Y~~ VIK ~~I~~ H	
Sen. sequence5	(421)	RCFSYCAMFPKDFLIGKKT ~~L~~ INLWMAQGYIVPLDKDQSIDDASEEYISILLQRCFFENIGTEKD ~~Y~~ VIK ~~I~~ H	
Sen. sequence6	(421)	RCFSYCAMFPKDFLIGKKT ~~L~~ INLWMAQGYIVPLDKDQSIDDASEEYISILLQRCFFENIGTEKD ~~Y~~ VIK ~~I~~ H	
Sen. sequence8	(421)	RCFSYCAMFPKDFLIGKKT ~~L~~ INLWMAQGYIVPLDKDQSIDDASEEYISILLQRCFFENIGTEKD ~~Y~~ VIK ~~I~~ H	
Sen. sequence9	(421)	RCFSYCAMFPKDFLIGKKT ~~L~~ INLWMAQGYIVPLDKDQSIDDASEEYISILLQRCFFENIGTEKD ~~Y~~ VIK ~~I~~ H	
Sen. sequence10	(421)	RCFSYCAMFPKDFLIGKKT ~~L~~ INLWMAQGYIVPLDKDQSIDDASEEYISILLQRCFFENIGTEKD ~~Y~~ VIK ~~I~~ H	
Sen. sequence11	(421)	RCFSYCAMFPKDFLIGKKT ~~L~~ INLWMAQGYIVPLDKDQSIDDASEEYISILLQRCFFENIGTEKD ~~Y~~ VIK ~~I~~ H	
Sen. sequence12	(421)	RCFSYCAMFPKDFLIGKKT ~~L~~ INLWMAQGYIVPLDKDQSIDDASEEYISILLQRCFFENIGTEKD ~~Y~~ VIK ~~I~~ H	
Sen. sequence14	(421)	RCFSYCAMFPKDFLIGKKT ~~L~~ INLWMAQGYIVPLDKDQSIDDASEEYISILLQRCFFENIGTEKD ~~Y~~ VIK ~~I~~ H	
Sen. sequence15	(421)	RCFSYCAMFPKDFLIGKKT ~~L~~ INLWMAQGYIVPLDKDQSIDDASEEYISILLQRCFFENIGTEKD ~~Y~~ VIK ~~I~~ H	
Sen. sequence16	(421)	RCFSYCAMFPKDFLIGKKT ~~L~~ INLWMAQGYIVPLDKDQSIDDASEEYISILLQRCFFENIGTEKD ~~Y~~ VIK ~~I~~ H	
Sen. sequence17	(421)	RCFSYCAMFPKDFLIGKKT ~~L~~ INLWMAQGYIVPLDKDQSIDDASEEYISILLQRCFFENIGTEKD ~~Y~~ VIK ~~I~~ H	
Sen. sequence18	(420)	RCFSYCVVFPKDFLIGKKT ~~L~~ INLWMAQGYIVPLDKDQSIDDASEEYISILLQRCFFENIGTEKDD ~~V~~ IK ~~I~~ H	
Sen. sequence19	(421)	RCFSYCAMFPKDFLIGKKT ~~L~~ INLWMAQGYIVPLDKDQSIDDASEEYISILLQRCFFENIGTEKD ~~Y~~ VIK ~~I~~ H	
Sen. sequence20	(421)	RCFSYCAMFPKDFLIGKKT ~~L~~ INLWMAQGYIVPLDKDQSIDDASEEYISILLQRCFFENIGTEKD ~~Y~~ VIK ~~I~~ H	
Sen. sequence21	(378)	XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX	
Sen. sequence22	(421)	RCFSYCAVFPKDFLIGKKT ~~L~~ INLWMAQGYIVPLDKDQSIDDASEEYISILXQRCFFENIGTEKDD ~~V~~ IK ~~I~~ H	

FIG 3 I

		561	630
Res. sequence	(561)	ALVPNCLCLRALDLAWSKIKSLPDSIGGLLHLRYLDLSYNEDLEVLPNSIAKLYNLQTLQLKGCKRLEGL	
Sen. sequence1	(561)	ALVPNCLCLRALDLAWSKIKSVDPDSIGGLLHLRYLDLSYNEDLEVLPNSIAKLYNLQTLQLKGCKRLEGL	
Sen. sequence2	(561)	ALVPNCLCLRALDLAWSKIKSVDPDSIGGLLHLRYLDLSYNEDLEVLPNSIAKLYNLQTLQLKGCKRLEGL	
Sen. sequence3	(561)	ALVPNCLCLRALDLAWSKIKSVDPDSIGGLLHLRYLDLSYNEDLEVLPNSIAKLYNLQTLQLKGCKRLEGL	
Sen. sequence4	(561)	ALVPNCLCLRALDLAWSKIKSVDPDSIGGLLHLRYLDLSYNEDLEVLPNSIAKLYNLQTLQLKGCKRLEGL	
Sen. sequence5	(561)	ALVPNCLCLRALDLAWSKIKSVDPDSIGGLLHLRYLDLSYNEDLEVLPNSIAKLYNLQTLQLKGCKRLEGL	
Sen. sequence6	(561)	ALVPNCLCLRALDLAWSKIKSVDPDSIGGLLHLRYLDLSYNEDLEVLPNSIAKLYNLQTLQLKGCKRLEGL	
Sen. sequence8	(561)	ALVPNCLCLRALDLAWSKIKSVDPDSIGGLLHLRYLDLSYNEDLEVLPNSIAKLYNLQTLQLKGCKRLEGL	
Sen. sequence9	(561)	ALVPNCLCLRALDLAWSKIKSVDPDSIGGLLHLRYLDLSYNEDLEVLPNSIAKLYNLQTLQLKGCKRLEGL	
Sen. sequence10	(561)	ALVPNCLCLRALDLAWSKIKSVDPDSIGGLLHLRYLDLSYNEDLEVLPNSIAKLYNLQTLQLKGCKRLEGL	
Sen. sequence11	(561)	ALVPNCLCLRALDLAWSKIKSVDPDSIGGLLHLRYLDLSYNEDLEVLPNSIAKLYNLQTLQLKGCKRLEGL	
Sen. sequence12	(561)	ALVPNCLCLRALDLAWSKIKSVDPDSIGGLLHLRYLDLSYNEDLEVLPNSIAKLYNLQTLQLKGCKRLEGL	
Sen. sequence14	(561)	ALVPNCLCLRALDLAWSKIKSVDPDSIGGLLHLRYLDLSYNEDLEVLPNSIAKLYNLQTLQLKGCKRLEGL	
Sen. sequence15	(561)	ALVPNCLCLRALDLAWSKIKSVDPDSIGGLLHLRYLDLSYNEDLEVLPNSIAKLYNLQTLQLKGCKRLEGL	
Sen. sequence16	(561)	ALVPNCLCLRALDLAWSKIKSVDPDSIGGLLHLRYLDLSYNEDLEVLPNSIAKLYNLQTLQLKGCKRLEGL	
Sen. sequence17	(561)	ALVPNCLCLRALDLAWSKIKSVDPDSIGGLLHLRYLDLSYNEDLEVLPNSIAKLYNLQTLQLKGCKRLEGL	
Sen. sequence18	(560)	ALVPNCLCLRALDLAWSKIKSVDPDSIGGLLHLRYLDLSYNEDLEVLPNSIAKLYNLQTLQLKGCKRLEGL	
Sen. sequence19	(561)	ALVPNCLCLRALDLAWSKIKSVDPDSIGGLLHLRYLDLSYNEDLEVLPNSIAKLYNLQTLQLKGCKRLEGL	
Sen. sequence20	(561)	ALVPNCLCLRALDLAWSKIKSVDPDSIGGLLHLRYLDLSYNEDLEVLPNSIAKLYNLQTLQLKGCKRLEGL	
Sen. sequence21	(504)	ALVPNCLCLRALDLAWSKIKSLPDSIGGLLHLRYLDLSYNEDLEVLPNSIAKLYNLQTLQLKGCKRLEGL	
Sen. sequence22	(533)	XXXXNCLCLRALDLAWSKIKSLPDSIGGLLHLRYLDLSYNEDLEVLPNSIAKLYNLQTLQLKGCKRLEGL	

FIG 3 J

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

FIG 3 K

	701	770
Res. sequence	(701)	LEIQIRWPENTTDAVKEDVKREGLYLNHKEHLNHIVVDFRCEE ~~GG~~ GRMDDEEARRLMEELRPHPYLENLA
Sen. sequence1	(701)	LEIQIRWPENTTDAVKEDVKREGLYLNHKEHLNHIVVDFRCEE ~~GG~~ GRMDDEEARRLMEELRPHPYLENLA
Sen. sequence2	(701)	LEIQIRWPENTTDAVKEDVKREGLYLNHKEHLNHIVVD-----
Sen. sequence3	(701)	LEIQIRWPENTTDAVKEDVKREGLYLNHKEHLNHIVVD-----
Sen. sequence4	(701)	LEIQIRWPENTTDAVKEDVKREGLYLNHKEHLNHIVVD-----
Sen. sequence5	(701)	LEIQIRWPENTTDAVKEDVKREGLYLNH-----
Sen. sequence6	(701)	LEIQIRWPENTTDAVKEDVKREGLYLN-----
Sen. sequence8	(701)	LEIQIRWPENTTDAVKEDVKREGLYLNH-----
Sen. sequence9	(701)	LEIQIRWPENTTDAVKEDVKREGLYLNH-----
Sen. sequence10	(701)	LEIQIRWPENTTDAVKEDVKREGLYLNHKEHLNHIVVDFRCEE ~~GG~~ GRMDDEEARRLMEELRPHPYLENLA
Sen. sequence11	(701)	LEIQIRWPENTTDAVKEDVKREGLYLNHKEHLNHIVVD-----
Sen. sequence12	(701)	LEIQIRWPENTTDAVKEDVKREGLYLNHKEHLNHIVVD-----
Sen. sequence14	(701)	LEIQIRWPENTTDAVKEDVKREGLYLNHKEHLNHIVVD-----
Sen. sequence15	(701)	LEIQIRWPENTTDAVKEDVKREGLYLNHKEHLNHIVVD-----
Sen. sequence16	(701)	LEIQIRWPENTTDAVKEDVKREGLYLNHKEHLNHIVVD-----
Sen. sequence17	(701)	LEIQIRWPENTTDAVKEDVKREGLYLNHKEHLNHIVVD-----
Sen. sequence18	(700)	LEIQIRWPENTTDVVKEDVKREGLYLNHKEHLNHIVVD-----
Sen. sequence19	(701)	LEIQIRWPENTTDAVKEDVKREGLYLNHKEHLNHIVVD-----
Sen. sequence20	(701)	LEIQIRWPENTTDAVKEDVKREGLYLNHKEHLNHIVVD-----
Sen. sequence21	(644)	LEIQIRWPENTTDAVKEDVKREGLYLNH-----
Sen. sequence22	(669)	LEIQIRWPENTTDAVKEDVKREGLYLNHKEHLNHIVVD-----

FIG 3 L

	771		840
Res. sequence	(771)	VKAYYG <u>V</u> KMPGWATLLPNLT <u>E</u> FLSDCGE <u>L</u> ENLP <u>C</u> LGNLD <u>H</u> LKVLRLSHLAKLEYIEEDSS <u>S</u> AN <u>F</u> RC <u>R</u> PG	
Sen. sequence1	(771)	VKAYYGAKMPGWATLLPNLT <u>E</u> LYLSDCGE <u>S</u> ECLPCMG <u>N</u> LDCLKVLRLSHLAKLEYIEEDSTSAN <u>F</u> S <u>F</u> RP <u>G</u>	
Sen. sequence10	(771)	VKAYYGAKMPGWATLLPNLT <u>E</u> LYLSDCGE <u>S</u> ECLPCMG <u>N</u> LDCLKVLRLSHLAKLEYIEEDSTSAN <u>F</u> S <u>F</u> RP <u>G</u>	
	841		910
Res. sequence	(841)	PESAGLSLYFPSLE <u>R</u> LE <u>L</u> K <u>R</u> L <u>C</u> KLKGWRR <u>G</u> EGLGDDHQPFNESSS> <u>Ret. Ins</u> <NTQVQLQLCLPQLKSLRIERCPLLT	
Sen. sequence1	(841)	PESAGLSLYFPSLE <u>L</u> LE <u>L</u> ELKRLHKLKGWRR <u>R</u> EGLGDDHQPFNESSS> <u>Ret. Ins</u> <-----	
Sen. sequence10	(841)	PESAGLSLYFPSLE <u>L</u> LE <u>L</u> ELKRLHKLKGWRR <u>R</u> EGLGDDHQPFNESSS> <u>Ret. Ins</u> <-----	
	911		980
Res. sequence	(911)	FMPLCPKTEKLHLVVFNERLRIVHAKRDENFYAPLHSSSSDPENPRNTIPIPMFREVYINNVAVLNSLPM	
	981		1050
Res. sequence	(981)	EAFRCLTHMTIKNDEVESLGEVGEVFRSCSSSLRSLNITGCSNLRSVSGGLEHLTALEMLEIYDTHKLSL	
	1051		1120
Res. sequence	(1051)	SEDPEGVVPWKSLSLHSLSYLQLMNLPQLVNLPDSMQFLAALRTLIVHCTKLQSVPDWMPRLTSLRKLMV	
	1121		1163
Res. sequence	(1121)	SFCSAHLERRCQNPTGVDWPNIQHIPSIDVTSSLPKFLVLPYE	

FIG 4

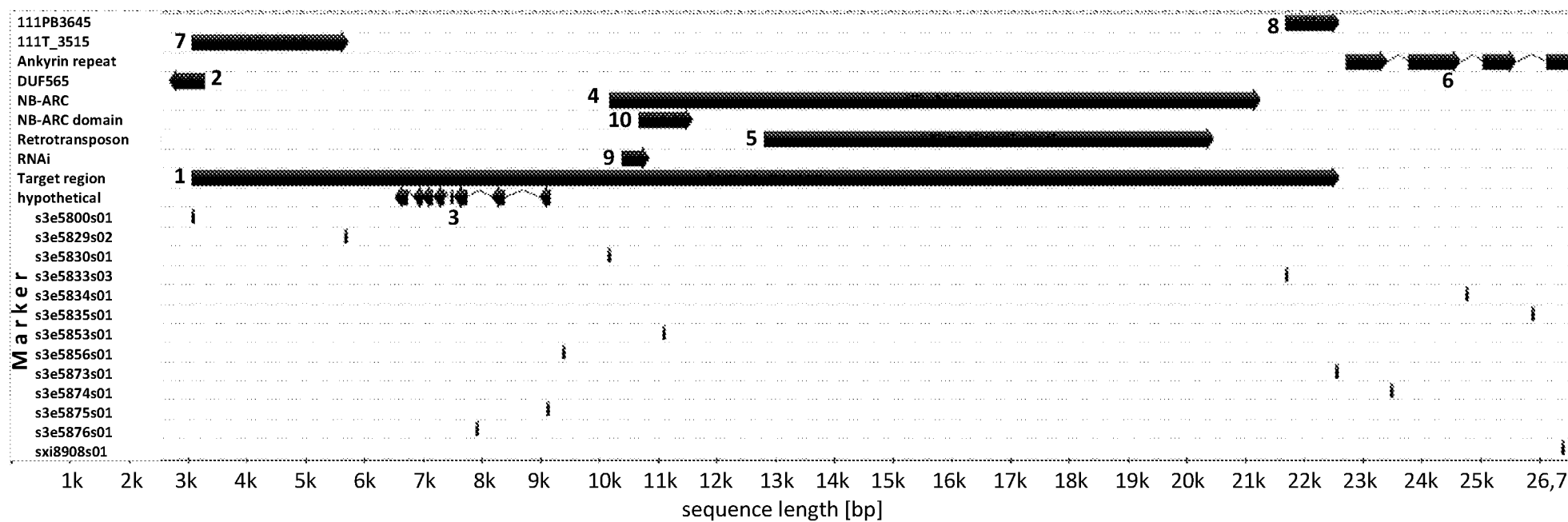
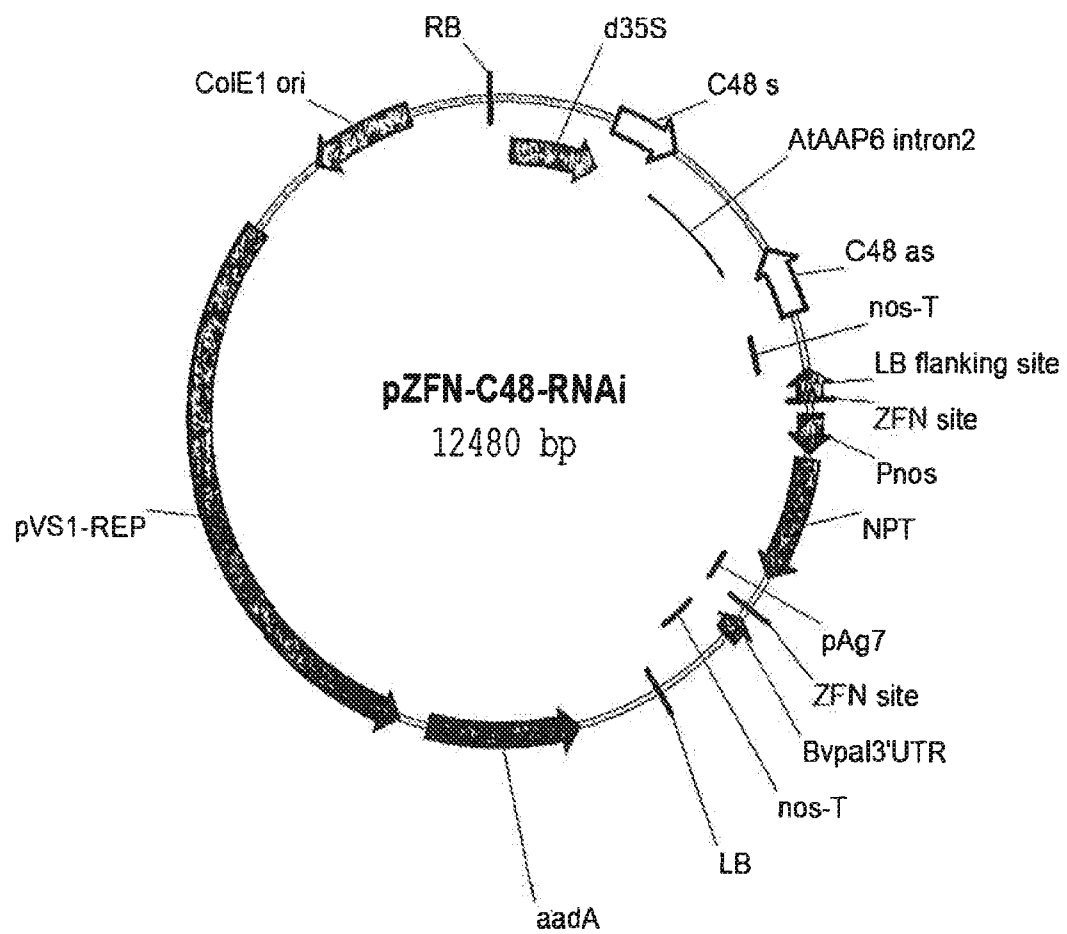


FIG 5

Recombinant Line	Population	Target Region																			Plants	ELISA values				Splicing				Assertion		
		Physical Position (bp)																				ELISA values				Splicing						
		Marker																				ELISA values				Splicing						
		s3e5299xxx	s3e5299s02	s3m0090s01	s3mp0090s1	s3e5800s01	s3e5829s02	s3e5876s01	s3e5875s01	s3e5856s01	s3e5830s01	s3e5853s01	s3e5833s03	s3e5873s01	s3e5874s01	s3e5834s01	s3e5835s01	sxi8908s01	s3e5332xxx	Number of descendants	Number RR	Number Rs	Number ss	Average all	Average RR	Average Rs	Average ss	Population splices	DIFF_RR-ss	T-Test	Wilcoxon	Position of the gene
ZR11007_03075	111T_3515	h	h	H	h	H	B	b	b	b	b	b	b	b	b	b	b	b	B	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	A	A	A	A	A	A	A	A	A	A	A	A	A	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	H	H	H	H	H	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	B	B	B	B	B	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	-	-	A	A	A	A	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	B	B	B	B	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	B	B	B	B	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	B	B	B	169	50	76	43	1.736	1.157	1.786	2.466	yes	-1.31	0.00	0.00	left of s3e5834s01

FIG 6



SEKVENSLISTE

Sekvenslisten er udeladt af skriftet og kan hentes fra det Europæiske Patent Register.

The Sequence Listing was omitted from the document and can be downloaded from the European Patent Register.

