(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) **World Intellectual Property Organization**
International Bureau

(43) **International Publication Date**
10 May 2007 (10.05.2007)

PCT

(10) **International Publication Number**
WO 2007/051626 A2

(51) **International Patent Classification:**
C12N 15/82 (2006.01)

(21) **International Application Number:** PCT/EP2006/010535

(22) **International Filing Date:** 1 November 2006 (01.11.2006)

(25) **Filing Language:** English

(26) **Publication Language:** English

(30) **Priority Data:**
PCT/EP2005/011718 1 November 2005 (01.11.2005) EP

(71) **Applicant** *(for all designated States except US)*: **UNIVERSITEIT UTRECHT HOLDING B.V.** [NL/NL]; Yalelaan 40, NL-3584 CM Utrecht (NL).

(72) **Inventors; and**
(75) **Inventors/Applicants** *(for US only)*: **VAN DEN ACKERVEKEN, Augustinus, Franciscus, Johannes, Maria** [NL/NL]; Dijkgraafpolder 12, NL-3991 ZC Houten (NL). **VAN DAMME, Mireille, Maria, Augusta** [NL/NL]; Ververstraat 2G, NL-1011 JA Amsterdam (NL).

(74) **Agent: VAN SOMEREN, Petronella, Francisca, Hendrika, Maria**; Arnold & Siedsma, Sweelinckplein 1, NL-2517 GK The Hague (NL).

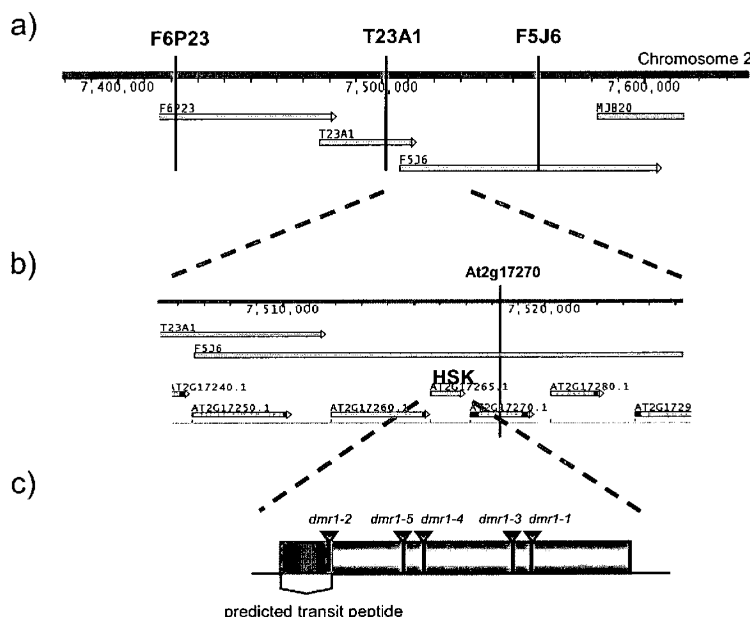
(81) **Designated States** *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) **Designated States** *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

[Continued on next page]

(54) **Title:** DISEASE RESISTANT PLANTS



(57) **Abstract:** The present invention relates to a plant, which is resistant to a pathogen of viral, bacterial, fungal or oomycete origin, wherein the plant has an increased homoserine level as compared to a plant that is not resistant to the said pathogen, in particular organisms of the phylum Oomycota. The invention further relates to a method for obtaining a plant, which is resistant to a pathogen of viral, bacterial, fungal or oomycete origin, comprising increasing the endogenous homoserine level in the plant.

WO 2007/051626 A2

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

DISEASE RESISTANT PLANTS

The present invention relates to disease resistant plants, in particular plants resistant to organisms of the phylum *Oomycota*, the oomycetes. The invention further relates to plant genes conferring disease resistance and methods of obtaining such disease resistant plants for providing protection to *Oomycota* pathogens.

Resistance of plants to pathogens has been extensively studied, for both pathogen specific and broad resistance. In many cases resistance is specified by dominant genes for resistance. Many of these race-specific or gene-for-gene resistance genes have been identified that mediate pathogen recognition by directly or indirectly interacting with avirulence gene products or other molecules from the pathogen. This recognition leads to the activation of a wide range of plant defence responses that arrest pathogen growth.

In plant breeding there is a constant struggle to identify new sources of mostly monogenic dominant resistance genes. In cultivars with newly introduced single resistance genes, protection from disease is often rapidly broken, because pathogens evolve and adapt at a high frequency and regain the ability to successfully infect the host plant. Therefore, the availability of new sources of disease resistance is highly needed.

Alternative resistance mechanisms act for example through the modulation of the defence response in plants, such as the resistance mediated by the recessive *mlo* gene in barley to the powdery mildew pathogen *Blumeria graminis f.sp. hordei*. Plants carrying mutated alleles of the wildtype *MLO* gene exhibit almost complete resistance coinciding with the abortion of attempted fungal penetration of the cell wall of

single attacked epidermal cells. The wild type *MLO* gene thus acts as a negative regulator of the pathogen response. This is described in WO9804586.

Other examples are the recessive powdery mildew resistance genes, found in a screen for loss of susceptibility to *Erysiphe cichoracearum*. Three genes have been cloned so far, named *PMR6*, which encodes a pectate lyase-like protein, *PMR4* which encodes a callose synthase, and *PMR5* which encodes a protein of unknown function. Both *mlo* and *pmr* genes appear to specifically confer resistance to powdery mildew and not to oomycetes such as downy mildews.

Broad pathogen resistance, or systemic forms of resistance such as SAR, has been obtained by two main ways. The first is by mutation of negative regulators of plant defence and cell death, such as in the *cpr*, *lsd* and *acd* mutants of *Arabidopsis*. The second is by transgenic overexpression of inducers or regulators of plant defence, such as in *NPR1* overexpressing plants.

The disadvantage of these known resistance mechanisms is that, besides pathogen resistance, these plants often show detectable additional and undesirable phenotypes, such as stunted growth or the spontaneous formation of cell death.

It is an object of the present invention to provide a form of resistance that is broad, durable and not associated with undesirable phenotypes.

In the research that led to the present invention, an *Arabidopsis thaliana* mutant screen was performed for reduced susceptibility to the downy mildew pathogen *Hyaloperonospora parasitica*. EMS-mutants were generated in the highly susceptible *Arabidopsis* line Ler *eds1-2*. Eight downy mildew resistant (*dmr*) mutants were analysed in detail, corresponding to 6 different loci. Microscopic analysis showed that in all mutants *H. parasitica* growth was severely

reduced. Resistance of *dmr*3, *dmr*4 and *dmr*5 was associated with constitutive activation of plant defence. Furthermore, *dmr*3 and *dmr*4, but not *dmr*5, were also resistant to *Pseudomonas syringae* and *Golovinomyces orontii*.

5 In contrast, enhanced activation of plant defense was not observed in the *dmr*1, *dmr*2, and *dmr*6 mutants. The results of this research have been described in Van Damme *et al.* (2005) Molecular Plant-Microbe Interactions 18(6) 583-592. This article does however not disclose the identification and 10 characterization of the *DMR* genes.

According to the present invention it was now found that *DMR1* is the gene encoding homoserine kinase (HSK). For *Arabidopsis* five different mutant *dmr1* alleles have been sequenced each leading to a different amino acid change in 15 the HSK protein. HSK is a key enzyme in the biosynthesis of the amino acids methionine, threonine and isoleucine and is therefore believed to be essential. The various *dmr1* mutants show defects in HSK causing the plants to accumulate homoserine. The five different alleles show different levels 20 of resistance that correlate to different levels of homoserine accumulation in the mutants.

The present invention thus provides a plant, which is resistant to a pathogen of viral, bacterial, fungal or oomycete origin, characterized in that the plant has an 25 altered homoserine level as compared to a plant that is not resistant to the said pathogen.

This form of resistance is in particular effective against pathogens of the phylum Oomycota, such as *Albugo, Aphanomyces, Basidiophora, Bremia, Hyaloperonospora,* 30 *Pachymetra, Paraperonospora, Perofascia, Peronophythora, Peronospora, Peronosclerospora, Phytium, Phytophthora, Plasmopara, Protobremia, Pseudoperonospora, Sclerospora, Viennotia* species.

The resistance is based on an altered level of homoserine *in planta*. More in particular, the resistance is based on an increased level of homoserine *in planta*. Such increased levels can be achieved in various ways.

5 First, homoserine can be provided by an external source. Second, the endogenous homoserine level can be increased. This can be achieved by lowering the enzymatic activity of the homoserine kinase gene which leads to a lower conversion of homoserine and thus an accumulation thereof.
10 Alternatively, the expression of the homoserine kinase enzyme can be reduced. This also leads to a lower conversion of homoserine and thus an accumulation thereof. Another way to increase the endogenous homoserine level is by increasing its biosynthesis via the aspartate pathway. Reducing the
15 expression of the homoserine kinase gene can in itself be achieved in various ways, either directly, such as by gene silencing, or indirectly by modifying the regulatory sequences thereof or by stimulating repression of the gene.

Modulating the HSK gene to lower its activity or
20 expression can be achieved at various levels. First, the endogenous gene can be directly mutated. This can be achieved by means of a mutagenic treatment. Alternatively, a modified HSK gene can be brought into the plant by means of transgenic techniques or by introgression, or the expression of HSK can
25 be reduced at the regulatory level, for example by modifying the regulatory sequences or by gene silencing.

In one embodiment of the invention, an increase (accumulation) in homoserine level in the plant is achieved by administration of homoserine to the plant. This is
30 suitably done by treating plants with L-homoserine, e.g. by spraying or infiltrating with a homoserine solution.

Treatment of a plant with exogenous homoserine is known from WO00/70016. This publication discloses how

homoserine is applied to a plant resulting in an increase in the phenol concentration in the plant. The publication does not show that plants thus treated are resistant to pathogens. In fact, WO00/70016 does not disclose nor suggest that an
5 increase in endogenous homoserine would lead to pathogen resistance.

Alternatively, endogenous homoserine is increased by modulating plant amino acid biosynthetic or metabolic pathways.

10 In one embodiment, the increased endogenous production is the result of a reduced endogenous HSK gene expression thus leading to a less efficient conversion of homoserine into phospho-homoserine and the subsequent biosynthesis of methionine and threonine. This reduced
15 expression of HSK is for example the result of a mutation in the HSK gene leading to reduced mRNA or protein stability.

In another embodiment reduced expression can be achieved by downregulation of the HSK gene expression either at the transcriptional or the translational level, e.g. by
20 gene silencing or by mutations in the regulatory sequences that affect the expression of the HSK gene. An example of a method of achieving gene silencing is by means of RNAi.

In a further embodiment the increase in endogenous homoserine level can be obtained by inducing changes in the
25 biosynthesis or metabolism of homoserine. In a particular embodiment this is achieved by mutations in the HSK coding sequence that result in a HSK protein with a reduced enzymatic activity thus leading to a lower conversion of homoserine into phospho-homoserine. Another embodiment is the
30 upregulation of genes in the aspartate pathway causing a higher production and thus accumulation of L-homoserine *in planta*.

This invention is based on research performed on resistance to *Hyaloperonospora parasitica* in *Arabidopsis* but is a general concept that can be more generally applied in plants, in particular in crop plants that are susceptible to infections with pathogens, such as Oomycota.

The invention is suitable for a large number of plant diseases caused by oomycetes such as, but not limited to, *Bremia lactucae* on lettuce, Peronospora farinosa on spinach, *Pseudoperonospora cubensis* on members of the Cucurbitaceae family, e.g. cucumber, *Peronospora destructor* on onion, *Hyaloperonospora parasitica* on members of the Brasicaceae family, e.g. cabbage, *Plasmopara viticola* on grape, *Phytophthora infestans* on tomato and potato, and *Phytophthora sojae* on soybean.

The homoserine level in these other plants can be increased with all techniques described above. However, when the modification of the HSK gene expression in a plant is to be achieved via genetic modification of the HSK gene or via the identification of mutations in the HSK gene, and the gene is not yet known it must first be identified. To generate pathogen-resistant plants, in particular crop plants, via genetic modification of the HSK gene or via the identification of mutations in the HSK gene, the orthologous HSK genes must be isolated from these plant species. Orthologs are defined as the genes or proteins from other organisms that have the same function.

Various methods are available for the identification of orthologous sequences in other plants.

A method for the identification of HSK orthologous sequences in a plant species, may for example comprise identification of homoserine kinase ESTs of the plant species in a database; designing primers for amplification of the complete homoserine kinase transcript or cDNA; performing

amplification experiments with the primers to obtain the corresponding complete transcript or cDNA; and determining the nucleotide sequence of the transcript or cDNA.

Suitable methods for amplifying the complete transcript or cDNA in situations where only part of the coding sequence is known are the advanced PCR techniques 5'RACE, 3'RACE, TAIL-PCR, RLM-RACE and vectorette PCR.

Alternatively, if no nucleotide sequences are available for the plant species of interest, primers are designed on the HSK gene of a plant species closely related to the plant of interest, based on conserved domains as determined by multiple nucleotide sequence alignment, and used to PCR amplify the orthologous sequence. Such primers are suitably degenerate primers.

Another reliable method to assess a given sequence as being a HSK ortholog is by identification of the reciprocal best hit. A candidate orthologous HSK sequence of a given plant species is identified as the best hit from DNA databases when searching with the *Arabidopsis* HSK protein or DNA sequence, or that of another plant species, using a Blast programme. The obtained candidate orthologous nucleotide sequence of the given plant species is used to search for homology to all *Arabidopsis* proteins present in the DNA databases (e.g. at NCBI or TAIR) using the BlastX search method. If the best hit and score is to the *Arabidopsis* HSK protein, the given DNA sequence can be described as being an ortholog, or orthologous sequence.

HSK is encoded by a single gene in *Arabidopsis* and rice as deduced from the complete genome sequences that are publicly available for these plant species. In most other plant species tested so far, HSK appears to be encoded by a single gene, as determined by the analysis of mRNA sequences and EST data from public DNA databases, except for potato,

tobacco and poplar for which two HSK homologs have been identified. The orthologous genes and proteins are identified in these plants by nucleotide and amino acid comparisons with the information that is present in public databases.

5 Alternatively, if no DNA sequences are available for the desired plant species, orthologous sequences are isolated by heterologous hybridization using DNA probes of the HSK gene of *Arabidopsis* or another plant or by PCR methods, making use of conserved domains in the HSK coding sequence to 10 define the primers. For many crop species, partial HSK mRNA sequences are available that can be used to design primers to subsequently PCR amplify the complete mRNA or genomic sequences for DNA sequence analysis.

In a specific embodiment the ortholog is a gene of 15 which the encoded protein shows at least 50% identity with the *Arabidopsis* HSK protein or that of other plant HSK proteins. In a more specific embodiment the homology is at least 55%, more specifically at least 60%, even more specifically at least 65%, at least 70%, at least 75%, at 20 least 80%, at least 85%, at least 90%, at least 95% or at least 99%.

Fig. 1 shows orthologous HSK sequences that have been identified in publicly available databases and obtained by PCR amplification on cDNA and subsequent sequencing.

25 After orthologous HSK sequences are identified, the complete nucleotide sequence of the regulatory and coding sequence of the gene is identified by standard molecular biological techniques. For this, genomic libraries of the plant species are screened by DNA hybridization or PCR with 30 probes or primers derived from a known homoserine kinase gene, such as the above described probes and primers, to identify the genomic clones containing the HSK gene. Alternatively, advanced PCR methods, such as RNA Ligase

Mediated RACE (RLM-RACE), can be used to directly amplify gene and cDNA sequences from genomic DNA or reverse-transcribed mRNA. DNA sequencing subsequently results in the characterization of the complete gene or coding
5 sequence.

Once the DNA sequence of the gene is known this information is used to prepare the means to modulate the expression of the homoserine kinase gene in any one of the ways described above.

10 More in particular, to achieve a reduced HSK activity the expression of the HSK gene can be down-regulated or the enzymatic activity of the HSK protein can be reduced by amino acid substitutions resulting from nucleotide changes in the HSK coding sequence.

15 In a particular embodiment of the invention, downregulation of HSK gene expression is achieved by gene-silencing using RNAi. For this, transgenic plants are generated expressing a HSK anti-sense construct, an optimized micro-RNA construct, an inverted repeat construct, or a
20 combined sense-anti-sense construct, so as to generate dsRNA corresponding to HSK that leads to gene silencing.

In an alternative embodiment, one or more regulators of the HSK gene are downregulated (in case of transcriptional activators) by RNAi.

25 In another embodiment regulators are upregulated (in case of repressor proteins) by transgenic overexpression. Overexpression is achieved in a particular embodiment by expressing repressor proteins of the HSK gene from a strong promoter, e.g. the 35S promoter that is commonly used in
30 plant biotechnology.

The downregulation of the HSK gene can also be achieved by mutagenesis of the regulatory elements in the promoter, terminator region, or potential introns. Mutations

in the HSK coding sequence in many cases lead to amino acid substitutions or premature stop codons that negatively affect the expression or activity of the encoded HSK enzyme.

These and other mutations that affect expression of HSK are induced in plants by using mutagenic chemicals such as ethyl methane sulfonate (EMS), by irradiation of plant material with gamma rays or fast neutrons, or by other means. The resulting nucleotide changes are random, but in a large collection of mutagenized plants the mutations in the HSK gene can be readily identified by using the TILLING (Targeting Induced Local Lesions IN Genomes) method (McCallum *et al.* (2000) Targeted screening for induced mutations. Nat. Biotechnol. 18, 455-457, and Henikoff *et al.* (2004) TILLING. Traditional mutagenesis meets functional genomics. Plant Physiol. 135, 630-636). The principle of this method is based on the PCR amplification of the gene of interest from genomic DNA of a large collection of mutagenized plants in the M2 generation. By DNA sequencing or by looking for point mutations using a single-strand specific nuclease, such as the CEL-I nuclease (Till et al. (2004) Mismatch cleavage by single-strand specific nucleases. Nucleic Acids Res. 32, 2632-2641) the individual plants that have a mutation in the gene of interest are identified.

By screening many plants, a large collection of mutant alleles is obtained, each giving a different effect on gene expression or enzyme activity. The gene expression or enzyme activity can be tested by analysis of HSK transcript levels (e.g. by RT-PCR), quantification of HSK protein levels with antibodies or by amino acid analysis, measuring homoserine accumulation as a result of reduced HSK activity. These methods are known to the person skilled in the art.

The skilled person can use the usual pathogen tests to see if the homoserine accumulation is sufficient to induce pathogen resistance.

Plants with the desired reduced HSK activity or expression are then back-crossed or crossed to other breeding lines to transfer only the desired new allele into the background of the crop wanted.

The invention further relates to mutated HSK genes encoding HSK proteins with a reduced enzymatic activity. In a particular embodiment, the invention relates to the *dmr1* alleles *dmr1-1, dmr1-2, dmr1-3, dmr1-4* and *dmr1-5*.

In another embodiment, the invention relates to mutated versions of the HSK genes of *Lactuca sativa*, *Vitis vinifera*, *Cucumis sativus, Spinacia oleracea* and *Solanum lycopersicum* as shown in **Figures 10-14**.

The present invention demonstrates that plants having an increased homoserine level show resistance to pathogens, in particular of oomycete origin. With this knowledge the skilled person can actively modify the HSK gene by means of mutagenesis or transgenic approaches, but also identify so far unknown natural variants in a given plant species that accumulate homoserine or that have variants of the HSK gene that lead to an increase in homoserine, and to use these natural variants according to the invention.

In the present application the terms "homoserine kinase" and "HSK" are used interchangeably.

The present invention is illustrated in the following examples that are not intended to limit the invention in any way. In the examples reference is made to the following figures.

Figure 1 shows the alignment of the amino acid sequences of the HSK proteins of *Arabidopsis thaliana* and orthologs from *Citrus sinensis*, *Populus trichocarpa* (1),

Populus trichocarpa (2), *Solanum tuberosum* (2), *Vitis vinifera*, *Lactuca sativa*, *Solanum tuberosum* (1), *Solanum lycopersicum*, *Nicotiana benthamiana*, *Ipomoea nil*, *Glycine max*, *Phaseolus vulgaris*, *Cucumis sativus*, *Spinacia oleracea*,
5 *Pinus taeda*, *Zea mays*, and *Oryza sativa* using the CLUSTAL W (1.82) multiple sequence alignment programme (EBI). Below the sequences the conserved amino acids are indicated by the dots, and the identical amino acids are indicated by the asteriks. The black triangles and corresponding text indicate
10 the amino acids that are substituted in the five *Arabidopsis dmr* mutants.

Table 2 shows the Genbank accession numbers and GenInfo identifiers of the *Arabidopsis* HSK mRNA and orthologous sequences from other plant species.

15 **Figure 2** shows the percentage of conidiophore formation by two *Hyaloperonospora parasitica* isolates, Cala2 and Waco9, on the mutants *dmr1-1, dmr1-2, dmr1-3* and *dmr1-4* and the parental line, L*er eds1-2*, 7 days post inoculation. The conidiophores formed on the parental line were set to
20 100%.

Figure 3 is a graphic overview of the three major steps in the cloning of *DMR1*. a) Initial mapping of *dmr1* resulted in positioning of the locus on the lower arm of chromosome 2 between positions 7,42 and 7,56 Mb. Three
25 insert/deletion (INDEL) markers were designed (position of the markers F6P23, T23A1 and F5J6 is indicated by the black lines). These markers were used to identify recombinants from several 100 segregating F2 and F3 plants. Primer sequences of these INDEL markers and additional markers to identify the
30 breakpoints in the collected recombinants is presented in table 3. b) One marker, At2g17270 (indicated by the grey line), showed the strongest linkage with resistance. The *dmr1* locus could be further delimited to a region containing 8

genes, at2g17250-at2g17290. The eight genes were amplified and sequenced to look for mutations in the coding sequences using the primers described in **table 4**. DNA sequence analysis of all 8 candidate genes led to the discovery of point
5 mutations in the At2g17265 gene in all 5 *dmr1* mutants. c) Each *dmr1* mutant has a point mutation at a different location in the At2g17265 gene, which encodes homoserine kinase.

Figure 4 shows a schematic drawing of the HSK coding sequence and the positions and nucleotide substitutions of
10 the 5 different *dmr1* mutations in the HSK coding sequence (the nucleotide positions, indicated by the black triangles, are relative to the ATG start codon which start on position 1). The 5'UTR and 3'UTR are shown by light grey boxes. Below the nucleotide sequence the protein sequence is shown. The
15 HSK protein contains a putative transit sequence for chloroplast targeting (dark grey part). The amino acid changes resulting from the 5 *dmr1* mutations are indicated at their amino acid (aa) position number (black triangles) in the HSK protein.

20 **Figure 5** shows the position of the homoserine kinase enzyme in the aspartate pathway for the biosynthesis of the amino acids threonine, methionine and isoleucine.

Figure 6 shows the number of conidiophores per L*er eds1-2* seedlings 5 days post inoculation with two different
25 isolates of *H. parasitica*, Waco9 and Cala2. The inoculated seedlings were infiltrated with dH2O, D-homoserine (5mM) or L-homoserine (5 mM) at 3 days post inoculation with the pathogen. Seedlings treated with L-homoserine show a complete absence of conidiophore formation and are thus resistant.

30 **Figure 7** shows the growth and development of *H. parasitica* in seedlings treated with water, D-homoserine (5 mM), or L-homoserine (5 mM) as analysed by microscopy of trypan blue stained seedlings.

a: Conidiophore formation after HS treatment on *Ler eds1-2* seedlings (10x magnification). No conidiophore formation was detected after L-homoserine infiltration, whereas control plants show abundant sporulation.

5 b: Haustorial development is affected by L-homoserine (5 mM) infiltration (40 x magnification), but not in plants treated with water or D-homoserine.

Figure 8 and **9** show the nucleotide and amino acid sequence of the homoserine kinase gene (At2g17265, NM_127281, 10 GI:18398362) and protein (At2g17265, NP_179318, GI:15227800) of *Arabidopsis thaliana*, respectively.

Figure 10 shows the nucleotide and the predicted amino acid sequence of the homoserine kinase coding sequence (CDS) and protein, respectively, of *Lactuca sativa*.

15 **Figure 11** shows the nucleotide and the predicted amino acid sequence of the homoserine kinase coding sequence (CDS) and protein, respectively, of *Vitis vinifera*.

Figure 12 shows the nucleotide and the predicted amino acid sequence of the homoserine kinase coding sequence 20 (CDS) and protein, respectively, of *Cucumis sativus*.

Figure 13 shows the nucleotide and the predicted amino acid sequence of the homoserine kinase coding sequence (CDS) and protein, respectively, of *Spinacia oleracea*.

Figure 14 shows the nucleotide and the predicted 25 amino acid sequence of the homoserine kinase coding sequence (CDS) and protein, respectively, of *Solanum lycopersicum*.

EXAMPLES

EXAMPLE 1

30 Characterization of the gene responsible for pathogen resistance in *dmr* mutants

Van Damme *et al.*, 2005, supra disclose four mutants, *dmr1-1, dmr1-2, dmr1-3 and dmr1-4* that are resistant to *H.*

parasitica. The level of resistance can be examined by counting conidiophores per seedling leaf seven day post inoculation with the *H. parasitica* Cala2 isolate (obtainable from Dr. E. Holub (Warwick HRI, Wellesbourne, UK or Dr. G. Van den Ackerveken, Department of Biology, University of Utrecht, Utrecht, NL). For the parental line, L*er eds1-2* (Parker *et al.*, 1996, Plant Cell 8:2033-2046), which is highly susceptible, the number of conidiophores is set at 100%. The reduction in conidiophore formation on the infected *dmr1* mutants compared to seedlings of the parental line is shown in **Fig. 2.**

According to the invention, the gene responsible for resistance to *H. parasitica* in the *dmr1* mutants of van Damme *et al.*, 2005, supra has been cloned by a combination of mapping and sequencing of candidate genes.

DMR1 was isolated by map-based cloning. The *dmr1* mutants were crossed to the *FN2* Col-0 mutant to generate a mapping population. The *FN2* mutant is susceptible to the *H. parasitica* isolate Cala2, due to a fast neutron mutation in the *RPP2A* gene (Sinapidou *et al.*, 2004, Plant J. 38:898-909). All 5 *dmr1* mutants carry single recessive mutations as the F1 plants were susceptible, and approximately a quarter of the F2 plants displayed *H. parasitica* resistance.

The DMR1 cloning procedure is illustrated in **Fig. 3** and described in more detail below. The map location of the *dmr1* locus was first determined by genotyping 48 resistant F2 plants to be located on the lower arm of chromosome 2. From an additional screen for new recombinants on 650 F2 plants ~90 F2 recombinant plants between two INDEL (insertion/ deletion) markers on BAC T24I12 at 7,2 Mb and BAC F5J6 at 7,56 Mb (according to the TIGR *Arabidopsis* genome release Version 5.0 of January 2004) were identified, which allowed to map the gene to a region containing a contig of 5 BACs.

The F2 plants were genotyped and the F3 generation was phenotyped in order to fine map the *dmr1* locus. The *dmr1* mutation could be mapped to a ~130 kb region (encompassing 3 overlapping BAC clones: F6P23, T23A1, and F5J6) between two INDEL markers located on BAC F6P23, at 7,42 Mb and F5J6 at 7,56 Mb (according to the TIGR *Arabidopsis* genome release Version 5.0 of January 2004). This resulted in an area of 30 putative gene candidates for the *dmr1* locus, between the *Arabidopsis* genes with the TAIR codes AT2g17060 and AT2g17380. Additionally cleaved amplified polymorphic sequences (CAPS) markers were designed based on SNPs linked to genes AT2g17190, AT2g17200, AT2g17270, At2g17300, At2g17310 and At2g17360 genes.

Analyses of 5 remaining recombinants in this region with these CAPS marker data left 8 candidate genes, At2g17230 (NM_127277, GI:30679913), At2g17240 (NM_127278, GI:30679916), At2g17250 (NM_127279, GI:22325730), At2g17260 (NM_127280, GI:30679922), At2g17265 (NM_127281, GI:18398362), At2g17270 (NM_127282, GI:30679927), At2g17280 (NM_127283, GI:42569096), At2g17290 (NM_127284, GI:30679934). Sequencing of all the 8 genes resulted in the finding of point mutations in the AT2g17265 coding gene in the five *dmr1* alleles; *dmr1-1, dmr1-2, dmr1-3*, *dmr1-4* and *dmr1-5,* clearly demonstrating that AT2g17265 is *DMR1*. **Fig. 3** shows a scheme of *dmr1* with point mutations of different alleles.

At2g17265 encodes the homoserine kinase (HSK) enzyme, so far the only *Arabidopsis* gene exhibiting this function.

In *Arabidopsis*, HSK is encoded by a single gene, At2g17265 (Lee & Leustek, 1999, Arch. Biochem. Biophys. 372: 135-142). HSK is the fourth enzyme in the aspartate pathway required for the biosynthesis of the amino acids methionine, threonine and isoleucine. HSK catalyzes the phosphorylation of homoserine to homoserine phosphate (**Fig. 5**).

EXAMPLE 2

Amino acid analysis

Homoserine phosphate is an intermediate in the production of methionine, isoleucine and threonine in *Arabidopsis*. Since homoserine kinase has a key role in the production of amino acids, free amino acid levels were determined in the parental line L*er eds1-2* and the four different *dmr1* mutants. For this amino acids from total leaves were extracted with 80% methanol, followed by a second extraction with 20% methanol. The combined extracts were dried and dissolved in water. After addition of the internal standard, S-amino-ethyl-cysteine (SAEC) amino acids were detected by automated ion-exchange chromatography with post column ninhydrin derivatization on a JOEL AminoTac JLC-500/V (Tokyo, Japan).

Amino acid analysis of four different *dmr1* mutants and the parental line, L*er eds1-2* showed an accumulation of homoserine in the *dmr1* mutants, whereas this intermediate amino acid was not detectable in the parental line L*er eds1-2*. There was no reduction in the level of methionine, isoleucine and threonine in the *dmr1* mutants (**Table 1**).

Table 1

Concentration (in pmol/mg fresh weight) of homoserine, methionine, threonine and isoleucine in above-ground parts of 2-week old seedlings of the parental line L*er eds1-2* and the mutants *dmr1-1, dmr1-2, dmr1-3* and *dmr1-4*.

	Homoserine	Methionine	Isoleucine	Threonine
dmr1-1	964	29	12	264
dmr1-2	7128	14	29	368
dmr1-3	466	11	16	212
dmr1-4	6597	11	32	597
L*er eds 1-2*	0	7	10	185

Due to the reduced activity of the HSK in the *dmr1* mutants, homoserine accumulates. This effect could be further enhanced by a stronger influx of aspartate into the pathway leading to an even higher level of homoserine. The high concentration of the substrate homoserine would still allow sufficient phosphorylation by the mutated HSK so that the levels of methionine, isoleucine and threonine are not reduced in the *dmr1* mutants and the parental line, L*er* *eds1-2* (**Table 1**).

EXAMPLE 3

Pathogen resistance is achieved by application of L-homoserine

To test if the effect is specific for homoserine the stereo-isomer D-homoserine was tested. Whole seedlings were infiltrated with water, 5 mM D-homoserine and 5 mM L-homoserine. Only treatment with the natural amino acid L-homoserine resulted in resistance to *H. parasitica*. Seedlings treated with water or D-homoserine did not show a large reduction in pathogen growth and were susceptible to *H. parasitica*. The infiltration was applied to two *Arabidopsis* accessions, L*er* *eds1-2* and Ws *eds1-1*, susceptible to Cala2 and Waco9, respectively. Conidiophore formation was determined as an indicator for *H. parsitica* susceptibility. Conidiophores were counted 5 days post inoculation with *H. parasitica* and 2 days post infiltration with water, D-homoserine or L-homoserine. **(Fig. 6)**. L-homoserine infiltration clearly results in reduction of conidiophore formation and *H. parasitica* resistance. This was further confirmed by studying pathogen growth *in planta* by trypan blue staining of *Arabidopsis* seedlings. Plants were inoculated with isolate Cala2. Two days later the plants were treated by infiltration with water, 5 mM D-homoserine, and 5 mM L-homoserine. Symptoms were scored at 5 days post

inoculation and clearly showed that only the L-homoserine-infiltrated seedlings showed a strongly reduced pathogen growth and no conidiophore formation **(Fig. 7)**.

Microscopic analysis showed that only in L-homoserine treated leaves the haustoria, feeding structures that are made by *H. parasitica* during the infection process, are disturbed. Again it is shown that increased levels of homoserine *in planta* lead to pathogen resistance.

EXAMPLE 4

Identification of HSK orthologs in crops

1. *Screening of libraries on the basis of sequence homology*

The nucleotide and amino acid sequences of the homoserine kinase gene and protein of *Arabidopsis thaliana* are shown in **Figs. 8** and **9**.

Public libraries of nucleotide and amino acid sequences were compared with the sequences of **Fig. 8** and **9**. This comparison resulted in identification of the complete HSK coding sequences and predicted amino acid sequences in *Citrus sinensis*, *Populus trichocarpa* (1), *Populus trichocarpa* (2), *Solanum tuberosum* (2), *Solanum tuberosum* (1), *Nicotiana benthamiana*, *Ipomoea nil*, *Glycine max*, *Phaseolus vulgaris*, *Pinus taeda*, *Zea mays*, and *Oryza sativa*. The sequence information of the orthologous proteins thus identified is given in **Fig. 1**. For many other plant species orthologous DNA fragments could be identified by BlastX as reciprocal best hits to the *Arabidopsis* or other plant HSK protein sequences.

2. *Identification of orthologs by means of heterologous hybridisation*

The HSK DNA sequence of *Arabidopsis thaliana* as shown in **Fig. 8** is used as a probe to search for homologous sequences by hybridization to DNA on any plant species using

standard molecular biological methods. Using this method orthologous genes are detected by southern hybridization on restriction enzyme-digested DNA or by hybridization to genomic or cDNA libraries. These techniques are well known to the person skilled in the art. As an alternative probe the HSK DNA sequence of any other more closely related plant species can be used as a probe.

3. *Identification of orthologs by means of PCR*

For many crop species, partial HSK mRNA or gene sequences are available that are used to design primers to subsequently PCR amplify the complete cDNA or genomic sequence. When 5' and 3' sequences are available the missing internal sequence is PCR amplified by a HSK specific 5' forward primer and 3' reverse primer. In cases where only 5', internal or 3' sequences are available, both forward and reverse primers are designed. In combination with available plasmid polylinker primers, inserts are amplified from genomic and cDNA libraries of the plant species of interest. In a similar way, missing 5' or 3' sequences are amplified by advanced PCR techniques; 5'RACE, 3' RACE, TAIL-PCR, RLM-RACE or vectorette PCR.

As an example the sequencing of the *Lactuca sativa* (lettuce) HSK cDNA is provided. From the Genbank EST database at NCBI several *Lactuca* HSK ESTs were identified using the tblastn tool starting with the *Arabidopsis* HSK amino acid sequence. Clustering and alignment of the ESTs resulted in a consensus sequence for a 5' HSK fragment and one for a 3' HSK fragment. To obtain the complete lettuce HSK cDNA the RLM-RACE kit (Ambion) was used on mRNA from lettuce seedlings. The 5' mRNA sequence was obtained by using a primer that was designed in the 3' HSK consensus sequence derived from ESTs (R1S1a: GCCTTCTTCACAGCATCCATTCC) and the 5'

RACE primers from the kit. The 3' cDNA sequence was obtained by using two primers designed on the 5'RACE fragment (Let3RACEOut: CCGTTGCGGTTAATGAGATT, and Let3RACEInn: TCGTGTTGGTGAATCCTGAA) and the 3' RACE primers from the kit.
5 Based on the assembled sequence new primers were designed to amplify the complete HSK coding from cDNA to provide the nucleotide sequence and derived protein sequence as presented in **Figure 10**. A similar approach was a used for *Solanum lycopersicum* (**Figure 14**) and *Vitis vinifera* (**Figure 11**).

10 The complete HSK coding sequences from more than 10 different plants species have been identified from genomic and EST databases. From the alignment of the DNA sequences, conserved regions in the coding sequence were selected for the design of degenerate oligonucleotide primers (for the
15 degenerate nucleotides the abbreviations are according to the IUB nucleotide symbols that are standard codes used by all companies synthesizing oligonucleotides; G = Guanine, A = Adenine, T = Thymine, C = Cytosine, R = A or G, Y = C or T, M = A or C, K = G or T, S = C or G, W = A or T, B = C or G or
20 T, D = G or A or T, H = A or C or T, V = A or C or G, N = A or C or G or T).

The procedure for obtaining internal HSK cDNA sequences of a given plant species is as follows:

1. mRNA is isolated using standard methods,
25 2. cDNA is synthesized using an oligo dT primer and standard methods,

3. using degenerate forward and reverse oligonucleotides a PCR reaction is carried out,

4. PCR fragments are separated by standard agarose
30 gel electrophoresis and fragments of the expected size are isolated from the gel,

5. isolated PCR fragments are cloned in a plasmid vector using standard methods,

6. plasmids with correct insert sizes, as determined by PCR, are analyzed by DNA sequencing,

7. Sequence analysis using blastX reveals which fragments contain the correct internal HSK sequences,

5 8. The internal DNA sequence can then be used to design gene- and species- specific primers for 5' and 3' RACE to obtain the complete HSK coding sequence by RLM-RACE (as described above).

As an example the sequencing of the *Cucumis sativus*

10 (cucumber) HSK cDNA is provided. For cucumber two primer combinations were successful in amplifying a stretch of internal coding sequence from cDNA; combination 1: primer F1Kom (GAYTTCYTHGGMTGYGCCGT) and M1RC (GCRGCGATKCCRGCRCAGTT), and combination 2: primer M1Kom (AACTGYGCYGGMATCGCYGC) and

15 R1Kom (CCATDCCVGGAATCAANGGVGC). After cloning and sequencing of the amplified fragments cucumber HSK-specific primers were designed for 5' RACE (Cuc5RACEOut : AGAGGATTTTTACTAAGTTTATTCGTG and Cuc5RACEInn: AGACATAATCTCCCAAGCCATCA) and 3' RACE (Cuc3RACEOut:

20 TGATGGCTTGGGAGATTATGTCT and Cuc3RACEInn: CACGAATAAACTTAGTAAAAATCCTCT). Finally the complete cucumber HSK cDNA sequence was amplified and sequenced (**Figure 12**). A similar approach was a used for spinach, *Spinacia oleracea* (**Figure 13**).

25 Orthologs identified as described in this example can be modified using well-known techniques to induce mutations that reduce the HSK expression or activity. Alternatively, the genetic information of the orthologs can be used to design vehicles for gene silencing. All these sequences are

30 then used to transform the corresponding crop plants to obtain plants that are resistant to *Oomycota*.

EXAMPLE 5

Reduction of homoserine kinase expression in *Arabidopsis* by means of RNAi

The production of HSK silenced lines has been achieved in *Arabidopsis* by RNAi. A construct containing two ~750bp fragments of the HSK exon in opposite directions was successfully transformed into the *Arabidopsis* Col-0 accession. The transformants were analysed for resistance to *H. parasitica*, isolate Waco9. Several transgenic lines were obtained that confer resistance to *H. parasitica*. Analysis of HSK expression and homoserine accumulation confirm that in the transformed lines the HSK gene is silenced, resulting in resistance to *H. parasitica*.

EXAMPLE 6

Mutation of seeds

Seeds of the plant species of interest are treated with a mutagen in order to introduce random point mutations in the genome. Mutated plants are grown to produce seeds and the next generation is screened for increased accumulation of homoserine. This is achieved by measuring levels of the amino acid homoserine, by monitoring the level of HSK gene expression, or by searching for missense mutations in the HSK gene by the TILLING method, by DNA sequencing, or by any other method to identify nucleotide changes.

The selected plants are homozygous or are made homozygous by selfing or inter-crossing. The selected homozygous plants with increased homoserine levels are tested for increased resistance to the pathogen of interest to confirm the increased disease resistance.

EXAMPLE 7

Transfer of a mutated allele into the background of a desired crop

Introgression of the desired mutant allele into a
5 crop is achieved by crossing and genotypic screening of the mutant allele. This is a standard procedure in current-day marker assistant breeding of crops.

TABLES

Table 2

GI numbers (GenInfo identifier) and Genbank accession number for Expressed Sequence Tags (ESTs) and mRNA sequences of the Arabidopsis HSK mRNA and orthologous sequences from other plant species.

A GI number (genInfo identifier, sometimes written in lower case, "gi") is a unique integer which identifies a particular sequence. The GI number is a series of digits that are assigned consecutively to each sequence record processed by NCBI. The GI number will thus change every time the sequence changes. The NCBI assigns GI numbers to all sequences processed into Entrez, including nucleotide sequences from DDBJ/EMBL/GenBank, protein sequences from SWISS-PROT, PIR and many others. The GI number thus provides a unique sequence identifier which is independent of the database source that specifies an exact sequence. If a sequence in GenBank is modified, even by a single base pair, a new GI number is assigned to the updated sequence. The accession number stays the same. The GI number is always stable and retrievable. Thus, the reference to GI numbers in the table provides a clear and unambiguous identification of the corresponding sequence.

Species	Common name	Detail	GI number	Genbank
Arabidopsis thaliana	Thale cress	mRNA	39104571	AK117871
Citrus sinensis	Sweet Orange	ESTs	55935768	CV886642
			28618675	CB293218
			55935770	CV886643
			28619455	CB293998
Glycine max	Soybean	ESTs	10846810	BF069552
			17401269	BM178051
			8283472	BE021031
			16348965	BI974560
			7285286	AW597773
			58024665	CX711406
			58017647	CX704389
			20449357	BQ253481
			16105339	BI893079
			37996979	CF808568
			37996460	CF808049
			6072786	AW102173
			26057235	CA800149
			6455775	AW186458
			6072724	AW102111
			9203587	BE329811
Ipomoea nil	Japanese morning glory	ESTs	74407098	CJ761918
			74402449	CJ757269
			74402115	CJ756935
			74388670	CJ743490
Nicotiana benthamiana	Tobacco	ESTs	39880685	CK295868
			39859026	CK284950
			39864851	CK287885
			39864855	CK287887
			39859024	CK284949
			39864853	CK287886
			39880683	CK295867
			39864849	CK287884
Oryza sativa	Rice	mRNA	50916171	XM_468550
			32970537	AK060519
Phaseolus vulgaris	Common Bean	ESTs	62708660	CV535256
			62710636	CV537232
			62708052	CV534648
			62709395	CV535991
			62710761	CV537357
			62708535	CV535131
			62708534	CV535130
			62711318	CV537914
			62707924	CV534520
			62710733	CV537329
			62709601	CV536197
			62709064	CV535660
			62708834	CV535430
Pinus taeda	Loblolly Pine	ESTs	70780626	DR690274

			67490638	DR092267
			48933532	CO162991
			34354980	CF396563
			67706241	DR117931
			17243465	BM158115
			34349136	CF390719
			66981484	DR057917
			48932595	CO162054
			66689208	DR011702
			48933450	CO162909
			34350236	CF391819
			67706323	DR118013
			48932678	CO162137
			66981399	DR057832
			34354850	CF396433
Populus trichocarpa 1	Poplar	Genome v1.0, LG_IX, 149339-148242		
		Expression confirmed by ESTs		
Populus trichocarpa 2	Poplar	Genome v1.0, scaffold_66, 1415935-1417032		
		Expression confirmed by ESTs		
Solanum tuberosum 1	Potato	ESTs	66838966	DR037071
			61238361	DN588007
			39804315	CK251362
			39801776	CK250065
			9250052	BE340521
			39832341	CK275363
			21917848	BQ116921
			9249876	BE340345
			39815050	CK258070
			39804985	CK251702
			39804987	CK251703
			39825384	CK268406
			39832342	CK275364
			66838967	DR037072
			9250394	BE340863
			39804317	CK251363
			39825385	CK268407
			21375072	BQ516203
Solanum tuberosum 2	Potato	ESTs	39813353	CK256373
			39793361	CK246131
			39793359	CK246130
			39813352	CK256372
Zea Mays	Maize	ESTs	76021237	DT948407
			76913306	DV165065
			71446162	DR827212
			71449720	DR830770
			78117576	DV535963
			91048486	EB158904
			71439095	DR820145
			76936546	DV174774
			76012246	DT939416

		78085419	DV513812
		71766843	DR964780
		76924795	DV170131
		71449067	DR830117
		91875652	EB405609
		71450175	DR831225
		78103551	DV521979
		78090555	DV518929
		78104654	DV523072
		76926251	DV170768
		78111568	DV529965
		71773353	DR971257
		71425952	DR807002
		93282458	EB674722
		78074199	DV502633
		76293328	DV032896
		78075462	DV503896
		91054750	EB165168
		86469295	DY235665
		74243218	DT651132
		74242899	DT650813
		101384764	EB814428
		91054750	EB165168
		71440426	DR821476
		78121780	DV540164
		78103550	DV521978
		86469294	DY235664
		91877777	EB407734
		67014441	CO443190
		76924794	DV170130
		76021236	DT948406
		71446161	DR827211
		78110960	DV529358
		78074736	DV503170
		71428043	DR809093
		86469052	DY235422
		71440425	DR821475
		78121779	DV540163
		78104653	DV523071
		37400920	CF637820
		78074198	DV502632
		71449719	DR830769
Solanum lycopersicum	Tomato	58213736	BP877213
		7333245	AW621598
		4386685	AI482761
		Unigene SGN-U223239 from Sol Genomics Network	
		Sequence described in this patent application	
Lactuca sativa	Lettuce	Sequence described in this patent application	
Vitis vinifera	Grape vine	Sequence described in this patent application	
Spinacia oleracea	Spinach	Sequence described in this patent application	
Cucumis sativus	Cucumber	Sequence described in this patent application	

Table 3

Primer sequences of insertion/deletion (INDEL, size difference indicated in brackets) markers and cleaved amplified polymorphics sequences (CAP, polymorphic restriction site indicated in brackets) used in the mapping of the *DMR1* locus.

Primer name: BAC and/or TAIR At code	**Forward primer sequence**	**Reverse primer sequence**	**TYPE (size/enzyme)**	**GI number of TAIR At code**
T24I12 (At2g16670)	AATCCAAATTTCTT GCGAGAACACA	AAACGAAGAGTGAC AATGGTTGGAG	INDEL (33)	18398180
F5J6 (AT2g17370-80)	CCGTCAGATCAGTC CTCATCTTGTT	CAGAAGCTGATGAT CGTGGAAAGTA	INDEL (30)	23506018, 30679966
F6P23 (AT2g17060)	CGGTTTCATGTCGA GGAAGATCATA	AAGAAGAGAACTGC GTCAACCTTCC	INDEL (37)	22325728
T23A1 (AT2g17220-30)	TCCTTCCATGTCCG AAACCA	AACAAATTTGCTTC CAGCCTTT	INDEL (26)	42570808, 30679913
AT2g17190	GAATAGAGGTTGAT GGAAATCAAGA	CTCTTGTATGTTTT ACTGGGCTGAT	CAP (MseI)	30679898
AT2g17200	CCTCTCCACCCATT TCTAATTTCG	CGATCCATTTCGTC AAGCAATCTAC	CAP (MboII)	30679902
AT2g17270	GATGCAGCTAAATT ATCAGTGTGAA	ACGAAAATATCAAA AAGCTCCTTC	CAP (NlaIII)	30679927
AT2g17300-05	AGGTAGGATGGTAT TATGTTTGAACT	GCATGTTTTCTCTA AGCGATAGAAG	CAP (EcoRI)	30679937, 22325732
AT2g17310	ATGGGTAACGAAAG AGAGGATTAGT	CACATGTATAAGGT CTTCCCATAGA	CAP (MseI)	42569097
AT2g17360	CCAACAAGTATCCT CTTTTGTTGTT	CCACATCAAACTTA ATGAACTCCAC	CAP (MaeIII)	30679959

Table 4

Primer sequences used for amplifying and sequencing of eight candidate *DMR1* genes for which the TAIR and GI codes are indicated

Primer name	Primer sequence	TAIR codes	GI codes
MvD17230_F	TTCCCGAAGTGTACATTAAAAGCTC	At2g17230	30679913
MvD17230_R	TATGTCATCCCCAAGAGAAGAAGAC	At2g17230	30679913
MvD17240-F	CAATAAAAGCCTTTAAAAGCCCACT	At2g17240	30679916
MvD17240_R	TAGCTTCTGAAACTGTGGCATTACA	At2g17240	30679916
MvD17250_1F	CATGATTTGAGGGGTATATCCAAAA	At2g17250	22325730
MvD17250_1R	GGAGGTGGGATTTGAGATAAAACTT	At2g17250	22325730
MvD17250_2F	TAGCCTAGAACTCTCTGTTCGCAAG	At2g17250	22325730
MvD17250_2R	CATTATTTTGCGTAGTTGTGAGTGG	At2g17250	22325730
MvD17250_3F	CGAAGAAATCCTACAATCAACCATC	At2g17250	22325730
MvD17250_3R	TCTCACAATTCCCATCTCTTACTCC	At2g17250	22325730
MvD17260_1F	TTACTCATTTGGGTGAACAGAACAA	At2g17260	30679922
MvD17260_1R	ATCATCCCTAATCTCTCTGCTTCCT	At2g17260	30679922
MvD17260_2F	GATTAAGATACGGGGAATGGAGTCT	At2g17260	30679922
MvD17260_2R	ATGCAGACAAATAAGATGGCTCTTG	At2g17260	30679922
MvD17260_3F	GTTGTTGCTCCTGTCACAAGACTTA	At2g17260	30679922
MvD17260_3R	GAACAAAGACGAAGCCTTTAAACAA	At2g17260	30679922
MvD17265_F	GAGGACTGCATCTAGAAGACCCATA	At2g17265	18398362
MvD17265_R	TGGGCTCTCAACTATAAAGTTTGCT	At2g17265	18398362
MvD17270_F1	TAACGGTAAAGCAACGAATCTATCC	At2g17270	30679927
MvD17270_R1	TCAAACTGATAACGAGAGACGTTGA	At2g17270	30679927
MvD17270_F2	TTGCGTTCGTTTTTGAGTCTTTTAT	At2g17270	30679927
MvD17270_R2	AAACCAGACTCATTCCTTTGACATC	At2g17270	30679927
MvD17280_F1	TTTAGGATCTCTGCCTTTTCTCAAC	At2g17280	42569096
MvD17280_R1	GAGAAATCAATAGCGGGAAAGAGAG	At2g17280	42569096
MvD17280_F2	GCTTAAATAGTCCTCCTTTCCTTGC	At2g17280	42569096
MvD17280_R2	TCTGCTGGTTCTCATGTTGATAGAG	At2g17280	42569096
MvD17290_F1	CTCTCCTTCATCATTTCACAAATCC	At2g17290	30679934
MvD17290_R1	TTCCTCTCGCTGTAATGACCTCTAT	At2g17290	30679934
MvD17290_F2	TGCCACAGGTGTTGACTATGC	At2g17290	30679934
MvD17290_R2	TGCTCTTAAACCCGCAATCTC	At2g17290	30679934
MvD17290_F3	GAAGATGGCTTTAAAGGTCAGTTTGT	At2g17290	30679934
MvD17290_R3	AGCAACAACAACTAAAAGGTGGAAG	At2g17290	30679934

CLAIMS

1. Plant, which is resistant to a pathogen of viral, bacterial, fungal or oomycete origin, characterized in that the plant has an increased homoserine level as compared to a plant that is not resistant to the said pathogen.

5 2. Plant as claimed in claim 1, wherein the pathogen is an organism of the phylum Oomycota.

3. Plant as claimed in claim 2, wherein the pathogen is a *Albugo, Aphanomyces, Basidiophora, Bremia, Hyaloperonospora, Pachymetra, Paraperonospora, Perofascia,*
10 *Peronophythora, Peronospora, Peronosclerospora, Phytium, Phytophthora, Plasmopara, Protobremia, Pseudoperonospora, Sclerospora, Viennotia* species.

4. Plant as claimed in claim 2 or 3, wherein the plant and the pathogen are selected from *Bremia lactucae* on
15 lettuce, Peronospora farinosa on spinach, *Pseudoperonospora cubensis* on members of the Cucurbitaceae family, e.g. cucumber, *Peronospora destructor* on onion, *Hyaloperonospora parasitica* on members of the Brasicaceae family, e.g. cabbage, *Plasmopara viticola* on grape, *Phytophthora infestans*
20 on tomato and potato, and *Phytophthora sojae* on soybean.

5. Plant as claimed in any one of the claims 1-4, having a mutation in its homoserine kinase gene that affects the homoserine kinase activity of the encoded enzyme.

6. Plant as claimed in claim 5, wherein the mutation
25 in the homoserine kinase gene leads to an amino acid substitution in the encoded protein.

7. Plant as claimed in any one of the claims 1-4, having a mutation in the regulatory sequences of its homoserine kinase gene that affects the expression of the
30 encoded homoserine kinase.

8. Plant as claimed in any one of the claims 1-4, having a in its genome a gene-silencing construct on the

basis of the homoserine kinase gene that affects the expression of the encoded homoserine kinase.

9. Plant as claimed in any one of the claims 1-4, having upregulated genes in the aspartate pathway leading to
5 an increase in the endogenous homoserine level in the plant.

10. Plant as claimed in any one of the claims 5-8, wherein the gene is an orthologous gene of the *Arabidopsis* gene At2g17265 (NM_127281, GI:18398362) as shown in **Fig. 8.**

11. Plant as claimed in any one of the claims 5-8,
10 wherein the gene is a homoserine kinase gene as identified in the list of **Table 2.**

12. Plant as claimed in any one of the claims 5-8, wherein the gene is the homoserine kinase gene of *Lactuca sativa* having the nucleotide sequence and amino acid sequence
15 as shown in **Fig. 10.**

13. Plant as claimed in any one of the claims 5-8, wherein the gene is the homoserine kinase gene of *Vitis vinifera* having the nucleotide sequence and amino acid sequence as shown in **Fig. 11.**

20 14. Plant as claimed in any one of the claims 5-8, wherein the gene is the homoserine kinase gene of *Cucumis sativus* having the nucleotide sequence and amino acid sequence as shown in **Fig. 12.**

15. Plant as claimed in any one of the claims 5-8,
25 wherein the gene is the homoserine kinase gene of *Spinacia oleracea* having the nucleotide sequence and amino acid sequence as shown in **Fig. 13.**

16. Plant as claimed in any one of the claims 5-8, wherein the gene is the homoserine kinase gene of *Solanum*
30 *lycopersicum* having the nucleotide sequence and amino acid sequence as shown in **Fig. 14.**

17. Method for obtaining a plant, which is resistant to a pathogen of viral, bacterial, fungal or oomycete origin,

comprising increasing the endogenous homoserine level in the plant.

18. Method as claimed in claim 17, wherein increasing the endogenous homoserine level in the plant is achieved by external administration of homoserine to the plant.

19. Method as claimed in claim 18, wherein homoserine is administered to the plant by treatment of seedlings by spraying or infiltration with homoserine.

20. Method as claimed in claim 17, wherein increasing the endogenous homoserine level in the plant is achieved by mutation of the homoserine kinase gene of the plant.

21. Method as claimed in claim 20, wherein the mutation results in one or more amino acid changes that lead to a lower homoserine kinase activity.

22. Method as claimed in claim 20 or 21, wherein the mutation is effected by mutagenic treatment of the plant, in particular with a mutagen or radiation.

23. Method as claimed in claim 17, wherein increasing the endogenous homoserine level in the plant is achieved by reducing the expression of the homoserine kinase gene of the plant.

24. Method as claimed in claim 23, wherein reducing the expression of the homoserine kinase gene of the plant is achieved by gene silencing or RNAi.

25. Method as claimed in claim 23, wherein reducing the expression of the homoserine kinase gene of the plant is achieved by mutagenesis of regulatory elements in the promoter region, terminator region or an intron.

26. Method as claimed in claim 23, wherein reducing the expression of the homoserine kinase gene of the plant is achieved by overexpressing repressor proteins of the homoserine kinase gene.

27. Method as claimed in claim 26, wherein the homoserine kinase gene is expressed from the 35S promoter.

28. Method as claimed in claim 23, wherein reducing the expression of the homoserine kinase gene of the plant is achieved by silencing or mutation of plant genes encoding homoserine kinase activating or regulatory proteins.

5 29. Method as claimed in claim 17, wherein increasing the endogenous homoserine level in the plant is achieved by inducing alterations in the aspartate pathway.

30. Method as claimed in claim 21 to 29, wherein the homoserine kinase gene to be mutated is an orthologous gene

10 of the *Arabidopsis* gene At2g17265 (NM_127281, GI:18398362) as shown in **Fig. 8**.

31. Method as claimed in claim 30, wherein the orthologous gene is the homoserine kinase gene of *Lactuca sativa* having the nucleotide sequence and amino acid sequence

15 as shown in **Fig. 10**.

32. Method as claimed in claim 30, wherein the orthologous gene is the homoserine kinase gene of *Vitis vinifera* having the nucleotide sequence and amino acid sequence as shown in **Fig. 11**.

20 33. Method as claimed in claim 30, wherein the orthologous gene is the homoserine kinase gene of *Cucumis sativus* having the nucleotide sequence and amino acid sequence as shown in **Fig. 12**.

34. Method as claimed in claim 30, wherein the

25 orthologous gene is the homoserine kinase gene of *Spinacia oleracea* having the nucleotide sequence and amino acid sequence as shown in **Fig. 13**.

35. Method as claimed in claim 30, wherein the orthologous gene is the homoserine kinase gene of *Solanum*

30 *lycopersicum* having the nucleotide sequence and amino acid sequence as shown in **Fig. 14**.

36. Mutated plant HSK gene encoding a homoserine kinase having a reduced enzymatic activity.

37. Mutated plant HSK gene as claimed in claim 36, selected from the group consisting of the *dmr1* alleles *dmr1-1*, *dmr1-2*, *dmr1-3*, *dmr1-4* and *dmr1-5*.

Fig. 1

```
Arabidopsis_thaliana        ---MASLCFQSPS---KPISYFQPKS-----NPSPPLFAKVSVFRCRASVQTLVA----- 44
Citrus_sinensis             ----MAICFSSAV---KPANHFTVFF-----NPAP----KKPIFKCSCSLPTVTT----- 39
Populus_trichocarpa_1       ---MAICCFPSPL---KPMTPATPLT-----NLKPK---RPDILRCNFSLPTITT----- 41
Populus_trichocarpa_2       ---MAIC-FLSPL---KPITPTTSLTT----NLNPK---KPNILRCNFSLPIITT----- 41
Solanum_tuberosum_2         ----MAVLCQSP---LNLKLITSSSSS-----SRNRTANP----SFRLNLSAHSR----- 39
Vitis_vinifera              ----MAICFHSP---SKPTCISPSSNH-----YRPNLHAR----SFRCNFSKTLT----- 39
Lactuca_sativa              ----MAIRHYQPPFASTSSSISSTDLF-----KPPKLYLSS---SVRCNISVASK----- 43
Solanum_tuberosum_1         ----MAITYQSP---MKLNFITSNGFS-----NPPSLYPINTHFSFGFNLSSVSSKTQT- 47
Solanum_lycopersicum        ----MAITFQSP---MKLSFITSNGFS-----NPPSLYPINTHFSFGFNLSSVSSKTQT- 47
Nicotiana_benthamiana       ---MAAICYQSP---VKLNFTTSNAFSNPIPNNPPPLYPIKTRFSSGFNLSAVPSKTQTT 54
Ipomoea_nil                 ------------------ASISSTRHP-----NPPLCLPALNISRCGPLFSAVTSSTLA- 36
Glycine_max                 ----MATSTCFLC--PSTASLKGRARF------RIR---IRCSSS----VSVNIR----- 36
Phaseolus_vulgaris          ----MATAMSFLC--PSPATFKGTEMP------IAR---FRCCSSNTNSVSLNTR----- 40
Cucumis_sativus             ----MAMLSYQPP--LKSLTIPPVSLS------NPKPVLFRCSLSLPSRTAVTS------ 42
Spinacia_oleracea           ----MAICAQSP---FKPVNLSPHSPS----PTHKSPFICKLSLSSHSTHSPLTT----- 44
Pinus_taeda                 MESVFAQTKNHCFYLEPDLGLINSCFGLSRFRTKFSRGHLPHVFNVRCNAQQVSLKP--- 57
Zea_mays                    ----MAPAATSTAS--APSSFHSTGRHR----ARVGARPSLVSLRVRAANPNVT------ 46
Oryza_sativa                ---MAAAAAAAAAP--SPAPCFPSTRH-----TLPGLVSVRVSRRVKVA--VAI------ 44

                                 K (dmr1-2)
                                 ▼
Arabidopsis_thaliana        ----VEPEPVFVSVKTFAPATVANLGPGFDFLGCAVDG----LGDHVTLRVDPSVRAGEV 96
Citrus_sinensis             ----TEPEPVFTSVKTFAPATVANLGPCFDFLGCAVDG----LGDYVSLKVDPSVHPGEV 91
Populus_trichocarpa_1       ----TEPEPVFTSVRSFAPATVANLGPGFDFLGCAVDG----LGDFVSLRVDPSVHPGEL 93
Populus_trichocarpa_2       ----TEPEPVFTSVRSFAPATVANLGPGFDFLGCAVDG----LGDFVSLRVDPSVHPGEL 93
Solanum_tuberosum_2         ----SEPSPVFTSVKSFAPATVANLGPGFDFLGCAVDG----IGDFVTLRLDPNVHPGEV 91
Vitis_vinifera              ----ADPQPVFTSVKSFAPATVANLGPGFDFLGAAVDG----IGDFVSLRVDPDVRPGEI 91
Lactuca_sativa              --LEPEPHPVFTSVKSFAPATVANLGPGFDFLGCAIDG----IGDYVTLTVDPQVQPGRL 97
Solanum_tuberosum_1         HITIPEPEPVFTSVKSFAPATVANLGPGFDFLGCAVDG----IGDFVTLRVDPNVKAGEV 103
Solanum_lycopersicum        HITIPEPEPVFTSVKSFAPATVANLGPGFDFLGCAVDG----VGDFVTLRVDPNVKAGEV 103
Nicotiana_benthamiana       HITIPEPEPVFASVKSFAPATVANLGPGFDFLGCAVDG----IGDFITLRVDSKVKPGEV 110
Ipomoea_nil                 ---VSDPEPVYASVKSFAPATVANLGPGFDFLGCAVDG----IGDFVTVRVDPDVPPGQV 89
Glycine_max                 ----REPEPVTTLVKAFAPATVANLGPGFDFLGCAVDG----LGDIVSVKVDPQVHPGEI 88
Phaseolus_vulgaris          ----TEPQPVTTFVKAFAPATVANLGPGFDFLGCAVDG----IGDIVSVRVDPEVRPGEI 92
Cucumis_sativus             ----VEPQPVFSSVKAFAPATVANLGPGFDFLGCAVDG----LGDYVSLSVDSNVHPGEV 94
Spinacia_oleracea           -----EPTPLLTSVTTFAPATVANLGPGFDFLGCAVDG----LGDFVSLSVDPSVHPGQL 95
Pinus_taeda                 -VIQFEATPILQSVKAFAPATIANLGPGFDFLGCAVEG----LGDHVTVEVNEDVEPGKI 112
Zea_mays                    ----ADPAPAFQSVTTFAPATVANLGPGFDFLGCAVADASLSLGDTVTATLDPSLPPATV 100
Oryza_sativa                ----ADPAPAFNSVTAFAPATVANLGPGFDFLGCAVADASLSLGDTVTATLDPSLPPGTV 98
                                 :  *    * :*****:***** *****.*: .    :** ::  :: .: .. :

Arabidopsis_thaliana        SISEITGTTT-----KLSTNPLRNCAGIAAIATMKMLGIRSVGLSLDLHKGLPLGSGLGS 151
Citrus_sinensis             SISEVIGPS------KLSKNPLWNCAGIAAISAMKMLGVRSVGLSLSLEKGLPLGSGLGS 145
Populus_trichocarpa_1       SISDISGPK------KLSKNPLYNCAGIAAIATMKMLNIRSVGLSLSLEKGLPLGSGLGS 147
Populus_trichocarpa_2       SISDISGTK------NLSKNPLNNCAGIAAIATMKMLNIRSVGLSLSLEKGLPLGSGLGS 147
Solanum_tuberosum_2         SISDISGAG-----KKLRRNPRWNCAGIAAISVMKMLNIRSVGLTLSLHKGLPLGSCLGS 146
Vitis_vinifera              SIVDIDGVGNS--AKKLSKNPLWNCAGIAAISVMKMLGVRSVGLSLSLEKGLPLGSGLGS 149
Lactuca_sativa              SIAEINGVDKS--SKRLSRNPLWNCAGIAAISVMKMLKIRSVGLSLSINTCLPLRGGLGS 155
Solanum_tuberosum_1         SISDISGAG-----NRLSKDPLSNCAGIAAISVMKMLNIQSVGLSISLEKGLPLGSGLGS 158
Solanum_lycopersicum        SISDISGAG-----NRLSKDPLSNCAGIAAISVMKMLNIQSVGLSISLEKGLPLGSGLGS 158
Nicotiana_benthamiana       SISDISGAG-----GKLSKDPLSNCAGIAAISVMKMLNIQSVGLSISLEKGLPLGSGLGS 165
Ipomoea_nil                 SISEISGAG-----NKLSKNPLWNCAGIAAIAVMKMLRIQSVGLSLSLEKGLPLGSGLGS 144
Glycine_max                 CISDISGHAP----NKLSKNPLWNCAGIAAIEVMKMLSIRSVGLSLSLEKGLPLGSGLGS 144
Phaseolus_vulgaris          RISDITGHAP----NKLSTNPLWNCAGIAAIEVMKMLAIRSVGLSLSLQKGLPLGSGLGS 148
Cucumis_sativus             AISDITGNNT----NKLSKNPLYNCAGIAAIEVMKMLGIRSVGLSLSLEKGLPLGSGLGS 150
Spinacia_oleracea           SISSISGDAS----SKLSKDPLLNCAGISALAAMKLLNIRSVGLSLSLQKGLPLGSGLGS 151
Pinus_taeda                 VISFIDGDN-----NRLSLNPMKNCAGIAAKATMELLGVRSVGLSLGLHKGLPLGSGLGS 167
Zea_mays                    SIASVTSPSRPNLAERLSRDPLRNCAGVAAIAALRALGVRSHAVSIHLTKGLPLGSGLGS 160
Oryza_sativa                AIASVTSPSRPTLADRLSRDPLRNCAGVAAIAALRALDVKSHAVSIHLTKGLPLGSGLGS 158
                             *  : .        .*  :*  ****::*  .:. * ::* .::: : . *** . ***


                                                        D (dmr1-5)            R (dmr1-4)
                                                        ▼                     ▼
Arabidopsis_thaliana        SAASAAAAAVAVNEIFGRKLGSDQLVLAGLESEAKVSGYHADNIAPAIMGGFVLIRNYEP 211
Citrus_sinensis             SAASAAAAAVAVNEMFGNKLLPDELVLAGLESEAKVSGYHADNIAPAIMGGFVLIRSYEP 205
Populus_trichocarpa_1       SAASAAAAAVAVNELFGRKLEVKDLVLAGLESEAKVSGYHADNIAPAIMGGFVLIRSYDP 207
Populus_trichocarpa_2       SAASAAAAAVAVNEMFGRKLEVKDLVLAGLESEAKVSGYHADNIAPAIMGGFVLIRSYDP 207
Solanum_tuberosum_2         SAASAAAAAVAVNELFGRPLTLTDLVLAGLDSESKVSGYHADNVAPAIMGGFVLIRSYHP 206
Vitis_vinifera              SAASAAAAAVAVNEIFGRKLGVDDLVLAGLDSEAKVSGYHANNVAPALMGGFVLIRSYDP 209
Lactuca_sativa              SAASAAAAAVAVNEIFGGKLHDSDLILAGLEAEAKLSGYHADNIAPAIMGGFVLIRSYDP 215
Solanum_tuberosum_1         SAASAAAAAVAVNEIFGRKLSVDDLVLAGLESETKVSGYHADNIAPSIMGGFVLVRSYDP 218
Solanum_lycopersicum        SAASAAAAAVAVNEIFGRKLSVDDLVLAGLESETKVSGYHADNIAPSIMGGFVLIRSYDP 218
Nicotiana_benthamiana       SAASAAAAAVAVNELFGGKLSVSDLVLAGLESETKVSGYHADNIAPAIMGGFVLIRSYDP 225
Ipomoea_nil                 SAASAAAAAVAVNELFGSRLSVSDLVFAGLESESKVSGYHADNVAPSIMGGFVLIRSYDP 204
Glycine_max                 SAASAAAAAVAVNELFGKKLSVEELVLASLKSEEKVSGYHADNVAPSIMGGFVLIGSYSP 204
Phaseolus_vulgaris          SAASAAAAAVAVNEMFGKRLSVEDLVVASLKSEEKVSGYHADNVAPAIMGGFVLIQSYEP 208
Cucumis_sativus             SAASAAAAAIAVNGLFGGKLGVEELVLAGLKSEEKVSGYHADNVAPAIMGGFILIRNYEP 210
```

```
(continued)

Spinacia_oleracea         SAASAAAAAVAVNSLFGSPLSPLDLVHAGLESESKVSGYHADNIAPAIMGGFILIRSYEP 211
Pinus_taeda               SAASAAAAAVAVNGLFGNKLTKSDLVLAGLESEAAVSGYHADNVAPSLMGGFVLVRSYSP 227
Zea_mays                  SAASAAAAAKAVDALFGSRLGRDDLVLAGLESEKAVSGFHADNIAPAILGGFVLVRSYDP 220
Oryza_sativa              SAASAAAAAKAVDALFGSLLHQDDLVLAGLESEKAVSGFHADNIAPAILGGFVLVRSYDP 218
                          ********* **: :**  *    :*: *.*.:*  :**:**:*:**:::***:*: .* *

                                               I (dmr1-3)                V (dmr1-1)
                                               ▼                         ▼
Arabidopsis_thaliana      LDLKPLRFPSDKDLFFVLVSPDFEAPTKKMRAALPTEIPMVHHVWNSSQAAALVAAVLEG 271
Citrus_sinensis           LDLMRLNFPEKKQLLFVLVTPEFEAPTKKMRAALPAEVGMPHHIWNCSQAGALVAAVLNG 265
Populus_trichocarpa_1     LELMSLQFPVEKDLIFVLVSPDFEAPTKKMRAALPAEIGMPHHVWNCSQAGAFVASVLQG 267
Populus_trichocarpa_2     LELMSLQFPVEKDLIFVLVSPDFEAPTKKMRAALPAEIGMSHHVWNCSQAGALVASVLQG 267
Solanum_tuberosum_2       LELIQLNFPHEKDLFFVLANPEFEAPTKKMRAALPQEITMSHHIWNCSQAGALVASVLLG 266
Vitis_vinifera            LELIPLTFPSDKELFFVLVNPEFEAPTKKMRAALPSEIGMSDHVWNCSQAAALVASILQG 269
Lactuca_sativa            LELISLKFPPEKNLFFVLVNPEFQAQTKKMRAVLPTEITMSDHVWNCSQAAALVAGVLQG 275
Solanum_tuberosum_1       LELISLKFPFEKDLFFVLVNPEFEAPTKKMRAVLPSEVTMSHHIWNCSQAGALVAAILQG 278
Solanum_lycopersicum      LELIPLKFPFEKDLFFVLVNPEFEAPTKKMRAVLPSEVTMSHHIWNCSQAGALVAAILQG 278
Nicotiana_benthamiana     LELIELKFPLEKDLFFVLVNPEFEAPTKKMRAALPNEVTMSHHIWNSSQAGALVAAILQG 285
Ipomoea_nil               LELIQLKFPQEKSLFFVLVNPEFEAPTKKMRAALPAEITMSSHVWNCSQAGALVASVLQG 264
Glycine_max               LELMPLKFPAEKELYFVLVTPEIEAPTKKMRAALPTEIGMPHHVWNCSQAGALVASVLQG 264
Phaseolus_vulgaris        LRLIELKFPAEKELYFVLVSPEFEAPTKKMRAALPGEIAMAHHVWNCSQAGALVAAVLQG 268
Cucumis_sativus           LELIRLKFPVEKELFFVLVSPEFEAPTKKMRAALPAEVGMPHHVWNSSQAGALVAAVLQG 270
Spinacia_oleracea         LDLMKLEFPETNDLYFVLVSPEFEAPTKKMRAALPKEIGMPHHIWNSSQAAALVAAVLMG 271
Pinus_taeda               LDLIHLPFPSEKELFFVLVTPAFEAPTKEMRAVLPKNITMKDHIQNCSQAAALVAAILQG 287
Zea_mays                  FHLVPLSFPPALRLHFVLVTPDFEAPTSKMRAALPRQVDVQQHVRNSSQAAALVAAVLQG 280
Oryza_sativa              FHLIPLSSPPALRLHFVLVTPDFEAPTSKMRAALPKQVAVHQHVRNSSQAAALVAAVLQG 278
                          : *  *  *    * ***..* ::* *.:** .** :: :   *: *.***.*:**.:* *

Arabidopsis_thaliana      DAVMLGKALSSDKIVEPTRAPLIPGMEAVKKAALEAGAFGCTISGAGPTAVAVIDSEEKG 331
Citrus_sinensis           DPVGLGKALSSDKIVEPNRAPLIPGMEAVKKVAVEAGAYGCTISGAGPTAVAVVDNEEKG 325
Populus_trichocarpa_1     DLVGLGKALSSDKIVEPKRAPLIPGMEGVKKAALEAGAFGCTISGAGPTAVAVVDSEERG 327
Populus_trichocarpa_2     DLVGLGKALSSDKIVEPKRAPLIPGMVGVKKAALEAGAFGCTISGAGPTAVAVVGSEDRG 327
Solanum_tuberosum_2       DVSGFGKALSSDKIVEPRRTPLIPGMEGVKKAAMEAGAFGCTIRGAGPTVVAVTDNEETG 326
Vitis_vinifera            DLRGLGKALSSDRIVEPRRAPLIPGMEGVKKAALEAGAFGCTISGAGPTAVAITDDEEKG 329
Lactuca_sativa            DLVGFGKALSSDRIVEPRRAPLLPGMEDVKKAAMEAGAYGCTISGSGPTVVAVTDDEDRG 335
Solanum_tuberosum_1       DSRGLGKALSSDKIVEPRRGPLIPGMEGVKKAALKAGAFGCTISGAGPTLVAVTDDEERG 338
Solanum_lycopersicum      DSRGLGKALSSDKIVEPRRGPLIPGMEGVKKAALKAGAFGCTISGAGPTLVAVTDDEERG 338
Nicotiana_benthamiana     DSRGLGKALSSDKIVEPKRGPLIPGMEGVKKAALEAGAFGCTISGAGPTLVAVTDGEERG 345
Ipomoea_nil               DLPGLGKALSSDKIVEPRRAPLIPGMEAVKKAAIQAGAFGCTISGAGPTAVAVTDDEEKG 324
Glycine_max               DVVGLGKALSSDKIVEPRRAPLIPGMEAVKKRAAIQAGAFGCTISGAGPTAVAVIDDEQTG 324
Phaseolus_vulgaris        DVVGLGKALSSDKIVEPRRAPLIPGMEAVKKAALQAGAFGCTISGAGPTAVAVIDDELAG 328
Cucumis_sativus           DTMGLGKALSSDKIVEPRRSPLIPGMDGVKKAAIAAGAFGCTISGAGPTAVAVIDNEEKG 330
Spinacia_oleracea         DVEGIGKAMSSDKVVEPRRAPLIPGMMAVKKAAIEGGAFGCTISGAGPTAVAVTDREEKG 331
Pinus_taeda               DPCLLGAALSSDSIVEPKRGPFIPGMMAVKKAALETGAFGCTISGAGPTAVAITDTAEKG 347
Zea_mays                  DAGLIGSAMSSDGIVEPTRAPLIPGMAAVKAAALQAGALGCTISGAGPTVVAVIQGEERG 340
Oryza_sativa              DATLIGSAMSSDGIVEPTRAPLIPGMAAVKAAALEAGALGCTISGAGPTAVAVIDGEEKG 338
                          *    :* *:*** :*** * *::***  ** .*:  ** **** *:*** **:       *

Arabidopsis_thaliana      QVIGEKMVEAFWKVGHLKSVASVKKLDNVGARLVNSVSR---- 370
Citrus_sinensis           KVIGEKMVEAFWKEGNLKAVSMVKRLDRVGARLVGSVRAPR-- 366
Populus_trichocarpa_1     VEIGERMVETFWKEGKLKAVASVKRLDRVGARLVGSVPR---- 366
Populus_trichocarpa_2     MEVGERMVEAFWKEGNLKAVAMVKRLDRVGARLVGSVPR---- 366
Solanum_tuberosum_2       REIGQRMVEVFLEHGKLKALAMVKKLDRIGARLVSSQPI---- 365
Vitis_vinifera            REIGERMVEAFLEEGKLKAVAMVKQLDRVGARLMSSILR---- 368
Lactuca_sativa            REIGEKMVEAFVEKGKLKALAMVKKLDRVGARVISRISSQ--- 375
Solanum_tuberosum_1       REIGERMVDAFMKEGNLKALAMVKKLDRVGARLVSSNS----- 376
Solanum_lycopersicum      REIGERMVEAFMKEGNLKALAMVKKLDRVGARLVSSNS----- 376
Nicotiana_benthamiana     REIGERMVEAFMKEGKLKALAMVKQLDRVGARLVSSNPR---- 384
Ipomoea_nil               MEIGKRMVEAFIQEGNLKALAMVKRLDRVGARLVSKNGSICN- 366
Glycine_max               HLIAKHMIDAFLHVGNLKASANVKQLDRLGARRIPN------- 360
Phaseolus_vulgaris        NAIAEHMIHAFLHHGNLKASAKVLQLDRLGARRILD------- 364
Cucumis_sativus           KEIGERMVMAFLKEGNLKATASVKRLDRVGARLIGSTPLDRVL 373
Spinacia_oleracea         REIGERMVEAFWKEGGLKAAAVIQKLDRVGARVVSSVPR---- 370
Pinus_taeda               KAIAVAMVDMFQKKGQLEATASVQKLDRIGARVV--------- 381
Zea_mays                  EEVARKMVDAFWSAGKLKATATVAQLDTLGARVIATSSLN--- 380
Oryza_sativa              EEVGRRMVEAFANAGNLKATATVAQLDRVGARVISTSTLE--- 378
                            :.   *:  *    * *:: : : :** :*** :
```

Fig. 2

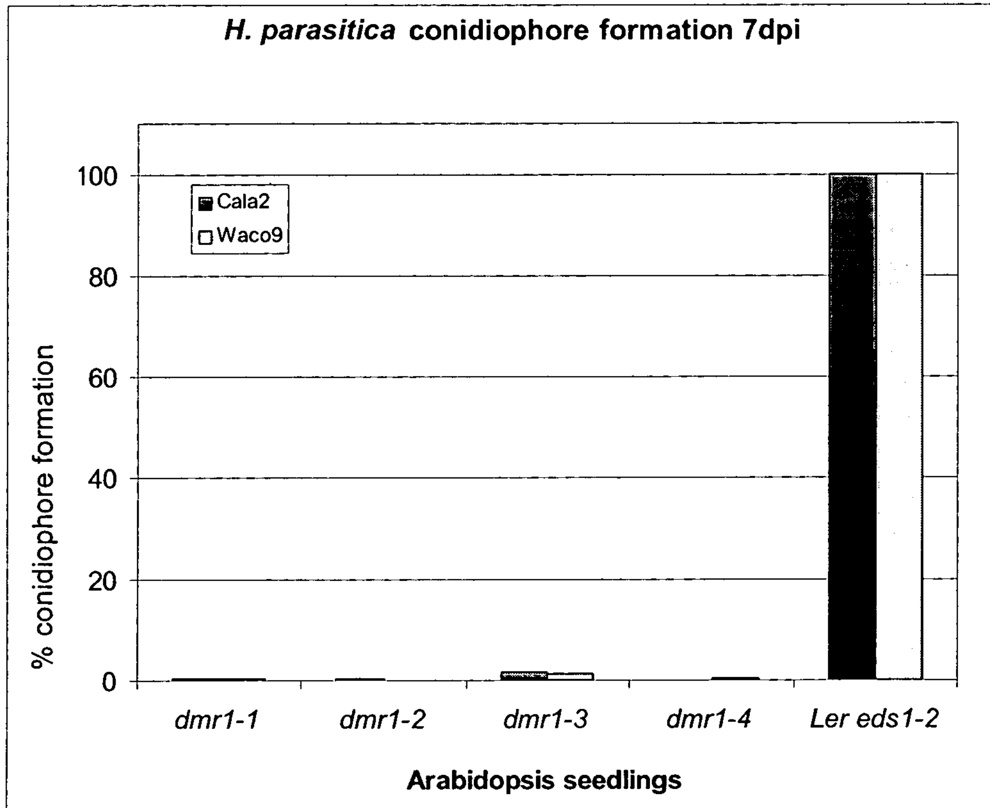


Fig. 3

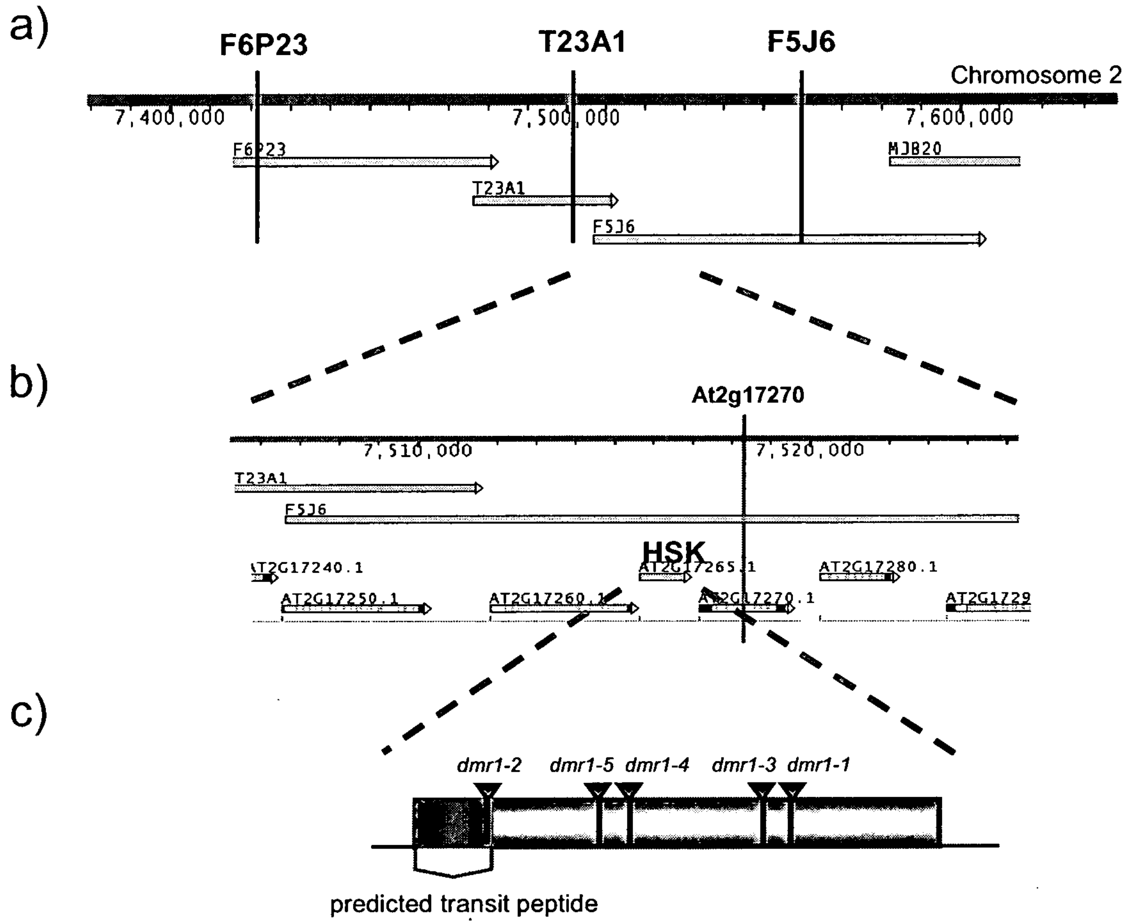


Fig. 4

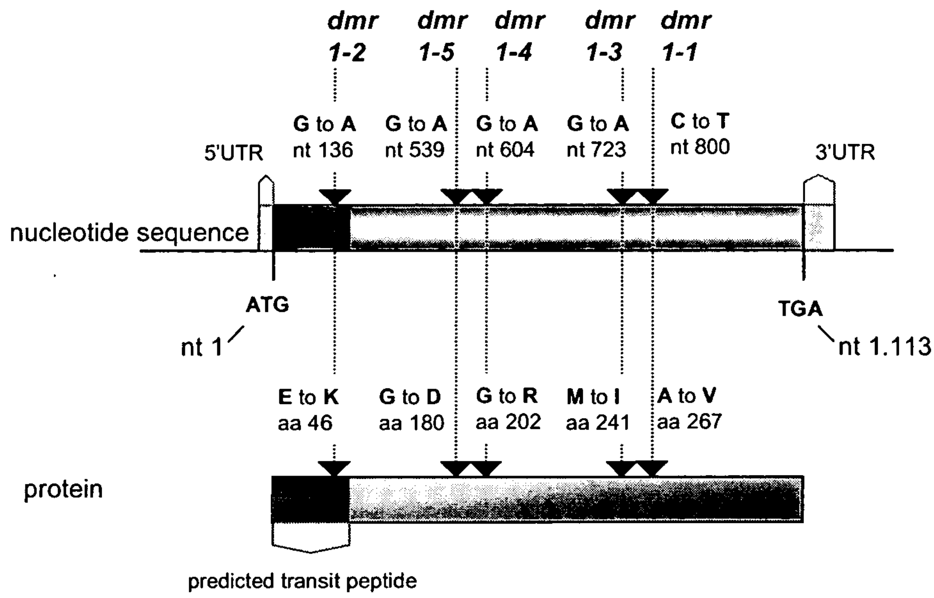


Fig. 5

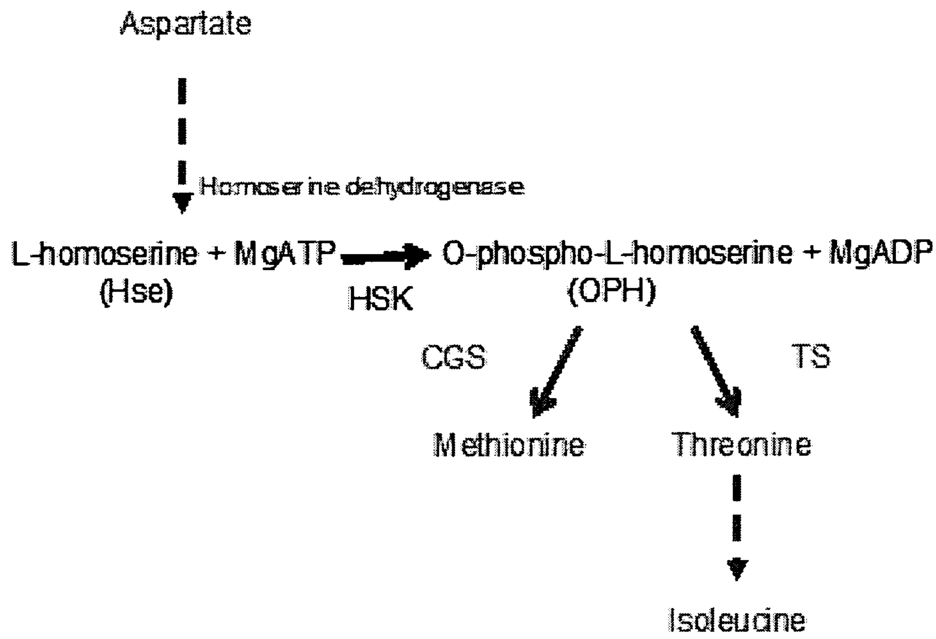


Fig. 6

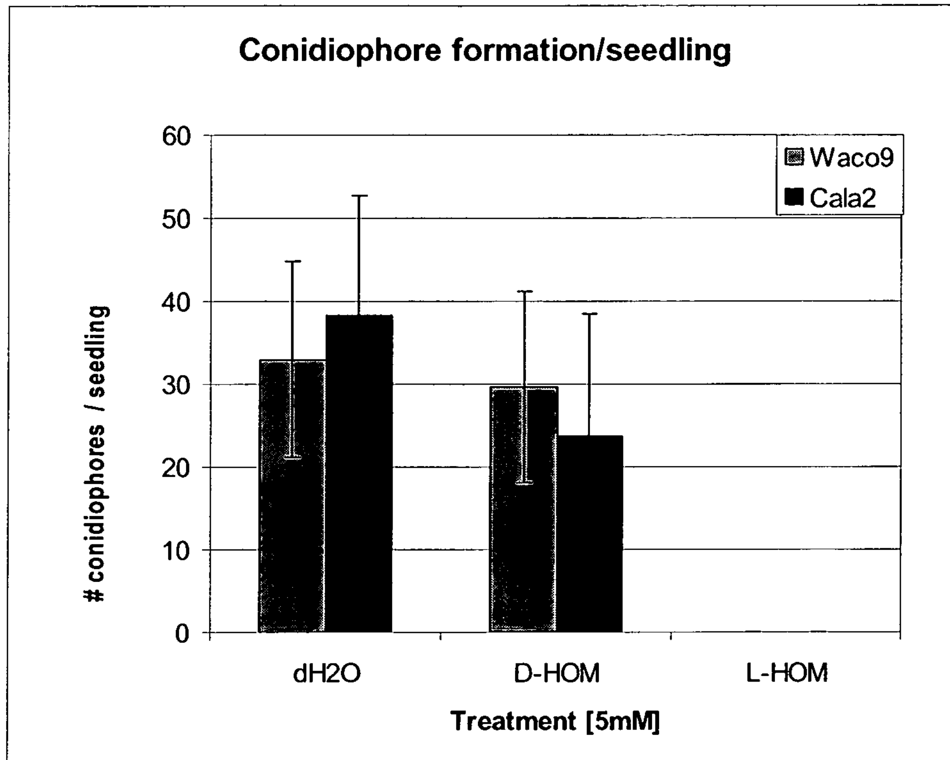


Fig. 7

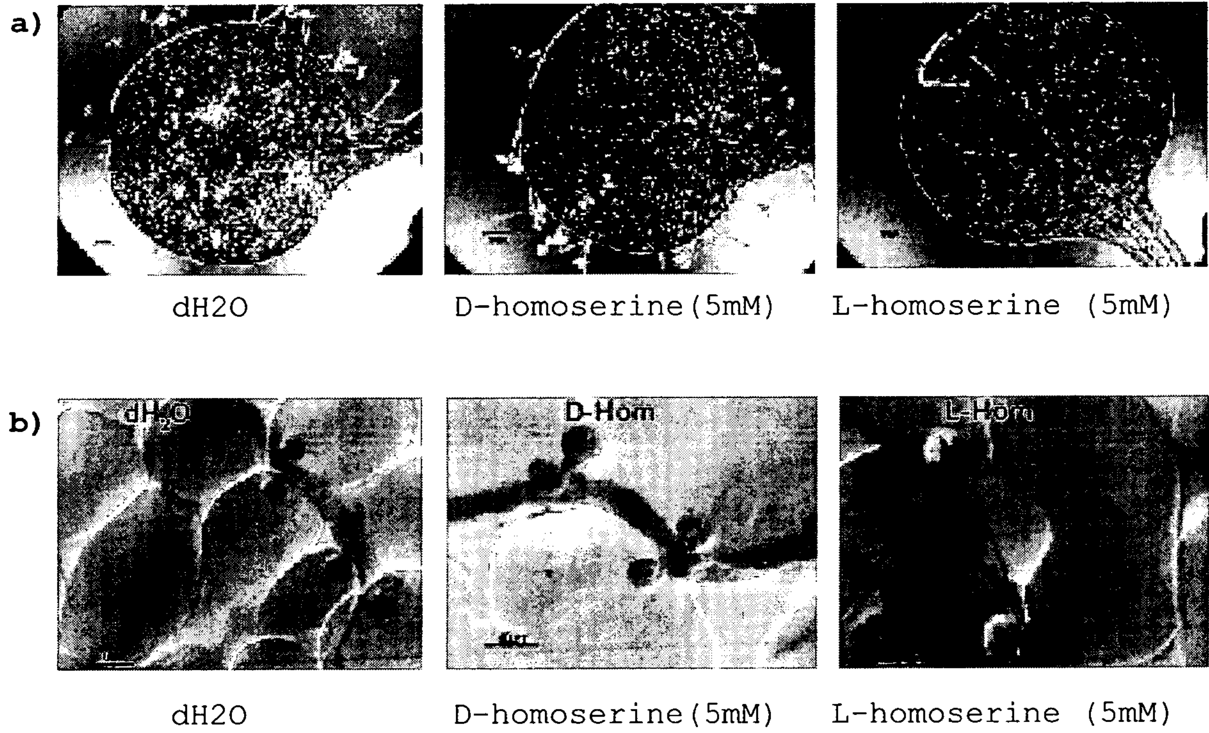


Fig. 8

```
   1 CTCATTACTT GTTCATCAAT GGCAAGTCTT TGTTTCCAAT CTCCTTCCAA
  51 ACCCATTTCC TATTTCCAAC CCAAATCCAA TCCATCGCCG CCGTTATTCG
 101 CCAAAGTCTC CGTCTTTCGA TGCAGAGCTT CCGTACAAAC CCTCGTCGCC
 151 GTTGAGCCGG AGCCAGTTTT CGTCTCCGTC AAGACTTTTG CGCCAGCCAC
 201 CGTCGCTAAT TTAGGACCAG GGTTTGATTT CTTAGGATGC GCCGTCGATG
 251 GTCTCGGAGA CCATGTGACT CTCCGTGTAG ATCCCTCTGT ACGAGCCGGT
 301 GAGGTCTCAA TCTCGGAGAT CACCGGAACG ACAACAAAAC TCAGCACAAA
 351 TCCTCTCCGG AACTGCGCCG GAATCGCTGC TATTGCTACA ATGAAGATGT
 401 TAGGGATCAG ATCGGTTGGT TTATCATTAG ATTTGCATAA AGGTCTTCCT
 451 TTAGGTAGCG GTTTAGGTTC TAGTGCAGCT AGCGCCGCCG CAGCTGCTGT
 501 GGCGGTTAAT GAGATCTTTG GTCGGAAATT AGGGAGTGAT CAATTGGTAT
 551 TAGCCGGTTT AGAATCGGAA GCGAAAGTCT CCGGTTATCA CGCTGATAAT
 601 ATCGCACCAG CGATCATGGG TGGATTCGTT TTGATTCGAA ACTACGAACC
 651 ACTTGATTTG AAACCATTGA AGTTCCCATC TGATAAAGAT CTCTTCTTTG
 701 TTCTAGTAAG CCCTGAGTTT GAAGCTCCAA CTAAGAAAAT GAGAGCTGCA
 751 TTGCCTACAG AGATTCCAAT GGTTCATCAT GTTTGGAACA GTAGCCAAGC
 801 AGCTGCTTTA GTCGCTGCTG TGTTAGAAGG TGACGCAGTG ATGCTTGGGA
 851 AGGCATTGTC GTCGGATAAG ATTGTGGAGC CAACTAGAGC GCCTTTGATT
 901 CCGGGGATGG AAGCTGTGAA GAAGGCAGCT TTGGAAGCTG GAGCGTTTGG
 951 ATGTACAATT AGCGGAGCTG GACCAACAGC AGTTGCGGTG ATTGATTCGG
1001 AGGAGAAGGG TCAAGTGATT GGAGAGAAGA TGGTGGAAGC GTTTTGGAAA
1051 GTTGGTCATT TGAAATCTGT TGCTTCTGTG AAGAAGCTTG ATAAGGTTGG
1101 TGCTAGGCTT GTCAACAGCG TCTCCAGATG ATCTTTTAAT GATGTTTGAT
1151 TATGCTAAGA TTGGAACAAA TCTTCCTTTG TACTGTAATT TCTAGATGAT
1201 AATAAAGTTG TTTGTTTTCT ACACT
```

Fig. 9

```
  1 MASLCFQSPS KPISYFQPKS NPSPPLFAKV SVFRCRASVQ TLVAVEPEPV
 51 FVSVKTFAPA TVANLGPGFD FLGCAVDGLG DHVTLRVDPS VRAGEVSISE
101 ITGTTTKLST NPLRNCAGIA AIATMKMLGI RSVGLSLDLH KGLPLGSGLG
151 SSAASAAAAA VAVNEIFGRK LGSDQLVLAG LESEAKVSGY HADNIAPAIM
201 GGFVLIRNYE PLDLKPLRFP SDKDLFFVLV SPDFEAPTKK MRAALPTEIP
251 MVHHVWNSSQ AAALVAAVLE GDAVMLGKAL SSDKIVEPTR APLIPGMEAV
301 KKAALEAGAF GCTISGAGPT AVAVIDSEEK GQVIGEKMVE AFWKVGHLKS
351 VASVKKLDNV GARLVNSVSR
```

Fig. 10

>*Lactuca sativa* HSK CDS
ATGGCAATTCGCCATTATCAACCTCCATTCGCCTCCACTTCTTCTTCTATCTCTAGTACA
GATTTATTCAAACCCCCTAAACTTTATCTTTCATCGTCTGTCCGGTGCAACATCTCCGTC
GCTTCCAAACTGGAACCCGAACCTCATCCAGTTTTCACCTCCGTTAAGTCATTCGCCCCC
GCCACCGTAGCCAACCTCGGGCCTGGTTTCGACTTCCTCGGCTGCGCAATCGACGGCATC
GGAGATTACGTTACCCTCACAGTCGACCCCCAAGTCCAACCCGGCAGATTATCAATTGCA
GAAATCAACGGCGTTGACAAGTCTTCCAAGAGGCTCAGCAGAAACCCTCTATGGAATTGC
GCCGGAATTGCTGCAATCTCCGTCATGAAGATGCTCAAGATCCGATCCGTTGGTCTCTCT
TTATCCATCAATACATGTCTCCCCCTTCGAGGCGGCCTAGGCTCCAGCGCCGCTAGCGCT
GCCGCCGCCGCCGTTGCGGTTAATGAGATTTTCGGAGGGAAGTTACATGATTCCGATTTG
ATACTCGCGGGGCTCGAAGCTGAAGCGAAGTTATCCGGTTATCACGCCGATAACATTGCT
CCGGCGATCATGGGCGGGTTTGTGTTGATCAGAAGCTACGATCCATTAGAGTTGATCTCC
TTGAAGTTTCCACCGGAAAAGAATCTGTTTTTCGTGTTGGTGAATCCTGAATTCCAAGCA
CAAACGAAGAAGATGAGGGCGGTTCTACCGACGGAGATAACAATGTCGGATCATGTATGG
AATTGTAGTCAGGCGGCGGCGTTGGTGGCAGGCGTATTGCAGGGGGATTTGGTGGGGTTT
GGGAAGGCATTGTCATCGGATAGAATAGTGGAGCCACGGCGGGCGCCATTGCTTCCGGGA
ATGGAAGATGTGAAGAAGGCAGCAATGGAAGCAGGGGCATATGGGTGTACGATAAGTGGG
TCAGGGCCGACGGTGGTGGCGGTGACGGATGATGAAGATAGAGGGAGGGAGATCGGGGAG
AAGATGGTGGAAGCTTTTGTAGAGAAGGGAAAGTTGAAAGCTTTGGCTATGGTGAAGAAA
CTGGACAGAGTTGGTGCTAGAGTTATCAGTCGTATCTCCAGCCAATGA

>*Lactuca sativa* HSK protein
MAIRHYQPPFASTSSSISSTDLFKPPKLYLSSSVRCNISVASKLEPEPHPV
FTSVKSFAPATVANLGPGFDFLGCAIDGIGDYVTLTVDPQVQPGRLSIAEINGVDKSSKR
LSRNPLWNCAGIAAISVMKMLKIRSVGLSLSINTCLPLRGGLGSSAASAAAAAVAVNEIF
GGKLHDSDLILAGLEAEAKLSGYHADNIAPAIMGGFVLIRSYDPLELISLKFPPEKNLFF
VLVNPEFQAQTKKMRAVLPTEITMSDHVWNCSQAAALVAGVLQGDLVGFGKALSSDRIVE
PRRAPLLPGMEDVKKAAMEAGAYGCTISGSGPTVVAVTDDEDRGREIGEKMVEAFVEKGK
LKALAMVKKLDRVGARVISRISSQ*

Fig. 11

>*Vitis vinifera* HSK CDS
ATGGCGATTTGCTTCCACTCCCCCTCAAAACCCACTTGCATTTCTCCCTCATCAAACCATTACAGACCCAATCTT
CATGCTCGGTCCTTCAGATGCAACTTCTCTAAAACATTAACTGCTGATCCTCAACCAGTTTTCACCTCTGTGAAG
TCCTTCGCACCCGCAACCGTTGCTAACCTCGGTCCCGGTTTCGATTTCCTCGGTGCTGCTGTTGATGGTATAGGC
GATTTCGTCTCCCTTCGCGTGGATCCTGATGTTCGGCCCGGGGAGATTTCGATTGTCGATATCGATGGTGTTGGG
AATAGCGCCAAGAAGCTCAGTAAAAATCCCCTCTGGAACTGCGCCGGCATTGCCGCTATCTCCGTCATGAAAATG
CTCGGAGTCCGATCGGTGGGGCTGTCCCTTTCCCTCGAGAAGGGGTTGCCATTGGGAAGTGGACTTGGGTCGAGC
GCTGCCAGTGCCGCCGCGGCTGCTGTGGCGGTGAATGAGATTTTTGGGCGGAAATTGGGAGTTGATGACCTTGTC
CTTGCTGGGCTTGACTCGGAAGCTAAAGTTTCGGGTTATCACGCGAACAATGTGGCGCCGGCTCTTATGGGAGGA
TTCGTGTTGATTCGGAGTTATGATCCTTTGGAGTTGATTCCTTTGACGTTTCCGAGCGACAAGGAGTTGTTTTTT
GTGTTGGTGAATCCGGAATTTGAAGCTCCCACCAAGAAAATGCGGGCGGCATTGCCGTCGGAGATCGGGATGTCT
GATCACGTGTGGAATTGTAGCCAGGCCGCTGCATTGGTAGCCTCGATTTTGCAAGGAGATTTGAGGGGGTTGGGC
AAGGCATTGTCCTCCGACAGAATTGTGGAGCCAAGGAGGGCACCCCTTGATCCCTGGGATGGAAGGAGTGAAAAAG
GCTGCTCTTGAGGCTGGTGCATTTGGCTGTACAATTAGTGGAGCAGGGCCGACTGCAGTTGCAATTACAGATGAC
GAAGAGAAGGGAAGGGAGATTGGAGAACGGATGGTAGAAGCTTTCTTGGAGGAAGGGAAGTTGAAGGCTGTAGCA
ATGGTGAAGCAACTCGATAGGGTTGGTGCTAGGCTTATGAGTAGCATCCTCAGATGA

>*Vitis vinifera* HSK protein
MAICFHSPSKPTCISPSSNHYRPNLHARSFRCNFS
KTLTADPQPVFTSVKSFAPATVANLGPGFDFLGAAVDGIGDFVSLRVDPDVRPGEISIVD
IDGVGNSAKKLSKNPLWNCAGIAAISVMKMLGVRSVGLSLSLEKGLPLGSGLGSSAASAA
AAAVAVNEIFGRKLGVDDLVLAGLDSEAKVSGYHANNVAPALMGGFVLIRSYDPLELIPL
TFPSDKELFFVLVNPEFEAPTKKMRAALPSEIGMSDHVWNCSQAAALVASILQGDLRGLG
KALSSDRIVEPRRAPLIPGMEGVKKAALEAGAFGCTISGAGPTAVAITDDEEKGREIGER
MVEAFLEEGKLKAVAMVKQLDRVGARLMSSILR*

Fig. 12

>*Cucumis sativus* HSK CDS
ATGGCTATGCTCTCCTATCAACCGCCATTGAAGTCGTTGACCATTCCTCCAGTTTCTTTATCTAACCCTAAACCT
GTTCTCTTCAGGTGCAGTTTGTCTCTTCCATCTAGAACCGCCGTCACTTCCGTCGAACCTCAACCCGTTTTCTCT
TCCGTCAAGGCGTTTGCTCCTGCAACCGTCGCTAATTTAGGTCCTGGGTTTGATTTCCTTGGCTGCGCTGTTGAT
GGCTTGGGAGATTATGTCTCTCTTAGTGTTGATTCCAATGTTCATCCAGGTGAAGTTGCGATTTCTGATATTACA
GGGAATAACACGAATAAACTTAGTAAAAATCCTCTCTATAATTGTGCTGGTATTGCTGCTATTGAAGTTATGAAA
ATGCTAGGGATCCGATCTGTTGGTCTTTCTCTTTCGCTTGAGAAAGGTTTGCCGTTAGGGAGTGGATTGGGATCT
AGTGCTGCGAGTGCAGCTGCTGCGGCGATTGCTGTTAATGGATTGTTCGGTGGGAAATTAGGAGTAGAGGAATTG
GTTCTCGCGGGGTTGAAATCGGAAGAGAAGGTTTCTGGGTACCATGCGGATAATGTCGCACCGGCTATCATGGGG
GGTTTCATTCTGATTCGAAATTACGAACCCTTGGAATTGATTCGTTTGAAATTCCCCGTCGAGAAGGAGCTGTTC
TTCGTGTTGGTCAGCCCGGAATTCGAAGCACCGACGAAGAAAATGCGGGCTGCGTTACCTGCTGAAGTTGGGATG
CCACACCATGTGTGGAATTCCAGCCAAGCCGGGGCGTTGGTGGCTGCGGTGCTGCAGGGTGACACGATGGGATTG
GGGAAAGCATTGTCATCAGACAAAATTGTGGAACCAAGGCGTTCGCCGTTGATTCCAGGTATGGATGGTGTTAAG
AAGGCAGCCATTGCTGCTGGGGCATTTGGGTGCACGATAAGCGGAGCAGGGCCAACAGCGGTGGCGGTGATCGAT
AACGAAGAGAAGGGGAAGGAGATTGGTGAGAGGATGGTTATGGCATTTCTGAAGGAAGGAAATTTGAAAGCTACG
GCATCTGTAAAGAGACTAGATCGAGTTGGTGCAAGGCTTATTGGATCAACTCCTTTAGATAGAGTTTTATGA

>*Cucumis sativus* HSK protein
MAMLSYQPPLKSLTIPPVSLSNPKPVLFRCSLSLPSRTAVTSVEPQPVFSSVKAFAPA
TVANLGPGFDFLGCAVDGLGDYVSLSVDSNVHPGEVAISDITGNNTNKLSKNPLYNCAGI
AAIEVMKMLGIRSVGLSLSLEKGLPLGSGLGSSAASAAAAAIAVNGLFGGKLGVEELVLA
GLKSEEKVSGYHADNVAPAIMGGFILIRNYEPLELIRLKFPVEKELFFVLVSPEFEAPTK
KMRAALPAEVGMPHHVWNSSQAGALVAAVLQGDTMGLGKALSSDKIVEPRRSPLIPGMDG
VKKAAIAAGAFGCTISGAGPTAVAVIDNEEKGKEIGERMVMAFLKEGNLKATASVKRLDR
VGARLIGSTPLDRVL*

Fig. 13

>*Spinacia oleracea* HSK CDS
ATGGCAATCTGCGCACAATCTCCATTCAAACCCGTCAATCTATCACCTCACTCCCCTTCTCCCACCCACAAATCC
CCATTCATCTGTAAACTTTCTCTCTCCTCCCACTCAACCCACTCACCTCTCACCACTGAACCAACACCACTCCTC
ACCTCCGTCACCACCTTCGCCCCCGCTACCGTCGCCAACCTCGGCCCAGGGTTCGACTTCCTCGGTTGCGCTGTC
GATGGCCTCGGTGACTTCGTTTCTCTTTCCGTTGACCCCTCCGTTCATCCCGGTCAACTCTCCATCTCCTCCATT
TCCGGCGACGCTTCTTCCAAACTCTCCAAAGATCCCCTTCTTAACTGCGCCGGTATCTCTGCCCTAGCCGCCATG
AAGCTCCTTAACATTCGCTCCGTCGGCCTTTCTCTATCTCTCCAAAAAGGGCTCCCACTTGGCTCCGGTCTCGGA
TCTTCAGCAGCTTCCGCTGCTGCTGCCGCTGTTGCTGTGAACTCCCTATTTGGCTCCCCTCTCTCTCCACTCGAC
CTCGTACACGCTGGACTTGAGTCAGAATCTAAAGTTTCCGGTTACCACGCTGACAACATTGCACCGGCGATAATG
GGTGGTTTTATCTTAATCAGGAGTTATGAGCCATTGGATTTGATGAAATTGGAGTTCCCTGAGACTAATGATTTG
TATTTCGTATTGGTTAGTCCGGAATTTGAAGCCCCAACGAAGAAGATGAGGGCGGCATTGCCGAAGGAGATCGGG
ATGCCGCACCACATATGGAATTCTAGCCAAGCGGCAGCATTGGTGGCGGCAGTTTTGATGGGTGACGTAGAAGGG
ATAGGAAAGGCAATGTCTTCCGATAAAGTGGTGGAGCCAAGGCGGGCACCATTGATTCCCGGGATGATGGCGGTG
AAGAAGGCGGCTATTGAAGGGGGAGCGTTCGGGTGTACAATTAGCGGGGCAGGGCCTACGGCTGTGGCAGTAACG
GATAGGGAGGAGAAGGGAAGAGAGATCGGAGAGAGAATGGTGGAAGCGTTTTGGAAGGAAGGAGGGTTAAAGGCT
GCCGCTGTGATTCAAAAGCTAGATAGAGTTGGTGCTAGAGTTGTTAGCAGTGTTCCCAGATGA

>*Spinacia oleracea* HSK protein
MAICAQSPFKPVNLSPHSPSPTHKSPFICKLSLSSHSTH
SPLTTEPTPLLTSVTTFAPATVANLGPGFDFLGCAVDGLGDFVSLSVDPSVHPGQLSISS
ISGDASSKLSKDPLLNCAGISALAAMKLLNIRSVGLSLSLQKGLPLGSGLGSSAASAAAA
AVAVNSLFGSPLSPLDLVHAGLESESKVSGYHADNIAPAIMGGFILIRSYEPLDLMKLEF
PETNDLYFVLVSPEFEAPTKKMRAALPKEIGMPHHIWNSSQAAALVAAVLMGDVEGIGKA
MSSDKVVEPRRAPLIPGMMAVKKAAIEGGAFGCTISGAGPTAVAVTDREEKGREIGERMV
EAFWKEGGLKAAAVIQKLDRVGARVVSSVPR*

Fig. 14

>*Solanum lycopersicum* HSK CDS
ATGGCTATAACCTTTCAATCTCCCATGAAACTCAGCTTCATCACTTCTAATGGCTTCTCAAATCCTCC
TTCTCTTTATCCCATCAATACCCATTTCTCATTTGGATTCAATCTCTCATCTGTCTCCTCCAAAACCC
AAACCCATATCACCATACCCGAACCCGAACCCGTATTCACCTCCGTCAAGTCGTTTGCTCCGGCCACT
GTTGCTAATCTAGGTCCGGGTTTTGATTTCCTCGGATGCGCCGTTGATGGAGTCGGAGATTTTGTCAC
TCTTCGGGTTGACCCAAATGTTAAAGCTGGGGAGGTTTCGATTTCTGATATCTCCGGTGCTGGAAATA
GGCTTAGTAAAGACCCTTTATCGAACTGTGCTGGAATAGCTGCTATTTCTGTTATGAAGATGTTGAAT
ATACAGTCTGTTGGTTTATCGATTTCGCTTGAAAAAGGGTTGCCGTTGGGTAGTGGACTTGGGTCTAG
TGCTGCTAGTGCTGCGGCGGCGGCGGTGGCTGTGAATGAGATTTTTGGACGGAAGTTGAGTGTTGATG
ATCTTGTGCTTGCTGGGTTGGAATCGGAAACGAAGGTTTCGGGTTATCATGCTGATAATATAGCACCT
TCGATTATGGGTGGTTTTGTGTTGATAAGAAGTTATGATCCGTTGGAATTGATCCCATTGAAGTTTCC
ATTTGAAAAAGATTTGTTTTTTGTGCTTGTGAATCCCGAATTCGAAGCTCCAACGAAGAAGATGAGGG
CGGTATTGCCATCGGAGGTGACAATGTCGCATCATATATGGAATTGTAGTCAGGCTGGGGCGTTGGTG
GCTGCGATATTGCAGGGGGATTCGAGGGGTTTAGGGAAGGCGTTGTCGTCTGATAAGATTGTGGAGCC
GAGGAGAGGGCCGTTGATTCCTGGGATGGAGGGAGTGAAGAAGGCGGCGTTGAAGGCTGGGGCATTTG
GTTGCACGATAAGCGGAGCTGGACCTACTTTGGTCGCGGTGACGGATGATGAAGAGAGAGGGAGGGAG
ATTGGGGAGAGAATGGTGGAGGCGTTTATGAAGGAAGGGAACTTGAAGGCTTTGGCTATGGTGAAGAA
GCTTGATCGAGTTGGTGCCCGCCTTGTTAGTAGCAATTCATGA

> *Solanum lycopersicum* HSK protein
MAITFQSPMKLSFITSNGFSNPPSLYPINTHFSFGFNLSSVSSKTQTHITIPEPEPVFTS
VKSFAPATVANLGPGFDFLGCAVDGVGDFVTLRVDPNVKAGEVSISDISGAGNRLSKDPL
SNCAGIAAISVMKMLNIQSVGLSISLEKGLPLGSGLGSSAASAAAAAVAVNEIFGRKLSV
DDLVLAGLESETKVSGYHADNIAPSIMGGFVLIRSYDPLELIPLKFPFEKDLFFVLVNPE
FEAPTKKMRAVLPSEVTMSHHIWNCSQAGALVAAILQGDSRGLGKALSSDKIVEPRRGPL
IPGMEGVKKAALKAGAFGCTISGAGPTLVAVTDDEERGREIGERMVEAFMKEGNLKALAM
VKKLDRVGARLVSSNS*