



(86) Date de dépôt PCT/PCT Filing Date: 2010/06/17
(87) Date publication PCT/PCT Publication Date: 2010/12/23
(45) Date de délivrance/Issue Date: 2022/05/10
(85) Entrée phase nationale/National Entry: 2011/12/15
(86) N° demande PCT/PCT Application No.: NL 2010/050373
(87) N° publication PCT/PCT Publication No.: 2010/147467
(30) Priorité/Priority: 2009/06/17 (EP09163015.2)

(51) Cl.Int./Int.Cl. *A01H 1/04* (2006.01)
(72) Inventeurs/Inventors:
DE HAAN, ANITA AFKE, NL;
VAN LUIJK, MARLEEN, NL;
ROZIER, BRAM, NL
(73) Propriétaire/Owner:
MONSANTO INVEST N.V., NL
(74) Agent: NORTON ROSE FULBRIGHT CANADA
LLP/S.E.N.C.R.L., S.R.L.

(54) Titre : NOUVEAUX PLANTS DE TOMATES AYANT UN INDICE SYMPODIAL MOYEN DE 2 ET PRODUCTION DE FRUITS DE COULEUR ROUGE
(54) Title: NOVEL TOMATO PLANTS HAVING AN AVERAGE SYMPODIAL INDEX OF 2 AND PRODUCING RED-COLORED FRUIT

(57) Abrégé/Abstract:

The present invention provides a method for the production of *S. lycopersicum* plants having an average sympodial index of 2 and producing red-colored fruits comprising crossing a plant of *S. lycopersicum* capable of producing red-colored fruits, with a plant of a *Solanum* spp. having an average sympodial index of 2, collecting the seeds resulting from the cross, regenerating the seeds into plants, providing one or more backcross generations, selfing the backcross plants, growing the selfed seed into plants, and identifying and selecting plants having an average sympodial index of 2 and producing red-colored fruits.



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

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
23 December 2010 (23.12.2010)

PCT

(10) International Publication Number
WO 2010/147467 A3

(51) International Patent Classification:

A01H 1/04 (2006.01) *A01H 5/08* (2006.01)
A01H 5/00 (2006.01) *C12Q 1/68* (2006.01)

(21) International Application Number:

PCT/NL2010/050373

(22) International Filing Date:

17 June 2010 (17.06.2010)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

09163015.2 17 June 2009 (17.06.2009) EP

(71) **Applicant** (for all designated States except US): **MON-SANTO INVEST N.V.** [NL/NL]; Leeuwenhoekweg 52, NL-2661 CZ Bergschenhoek (NL).

(72) **Inventors; and**

(75) **Inventors/Applicants** (for US only): **DE HAAN, Anita Afke** [NL/NL]; Karel Doormanlaan 10, NL-2665 AT Bleiswijk (NL). **VAN LUIJK, Marleen** [NL/NL]; Klapwijksezoom 277, NL-2652 BP Berkel en Rodenrijs (NL). **ROZIER, Bram** [NL/NL]; Fazantplein 118, NL-3334 SL Zwijndrecht (NL).

(74) **Agent:** **HATZMANN, M.J.**; Vereenigde, Johan de Wittlaan 7, NL-2517 JR Den Haag (NL).

(81) **Designated States** (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM,

AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, 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, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, 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, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

- with international search report (Art. 21(3))
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))

(88) **Date of publication of the international search report:**

3 February 2011

(54) **Title:** TOMATO PLANTS RESULTING FROM THE INTROGRESSION OF A TRAIT FROM SOLANUM PENNELLII INTO S. LYCOPERSICUM AND HAVING AN INCREASED YIELD

(57) **Abstract:** The present invention provides a method for the production of *S. lycopersicum* plants having an average sympodial index of 2 and producing red-colored fruits comprising crossing a plant of *S. lycopersicum* capable of producing red-colored fruits, with a plant of a *Solanum spp.* having an average sympodial index of 2, collecting the seeds resulting from the cross, regenerating the seeds into plants, providing one or more backcross generations, selfing the backcross plants, growing the selfed seed into plants, and identifying and selecting plants having an average sympodial index of 2 and producing red-colored fruits.



WO 2010/147467 A3

Novel Tomato Plants Having An Average Sympodial Index Of 2 And Producing Red-Colored Fruit

FIELD OF THE INVENTION

5

The present invention is in the field of plant breeding. More in particular, the invention relates to novel tomato plants having improved growth patterns, and to methods for producing such tomato plants using marker assisted breeding tools.

10 BACKGROUND OF THE INVENTION

Tomatoes (*Solanum* spp.) exhibit sympodial growth. This is a pattern of growth wherein the apical or terminal bud dies or ends in an inflorescence, and growth (sympodial shoots) continues from axillary or lateral buds. What looks like the plant's main axis is actually a series of many lateral branches, each arising from the previous lateral branch. After the production of some leaves by the shoot apical meristem (SAM), the growth of the primary shoot is terminated by the initiation of the first inflorescence, which is displaced from its terminal position by activation of the meristem at the axil of the last initiated leaf. The latter so-called sympodial meristem (SYM) continues shoot growth, carrying up the subtending leaf until it occupies a position above the inflorescence, which then develops laterally. The SYM undergoes a vegetative phase—producing most often three leaves—then initiates the second inflorescence, which is once again displaced laterally by the active outgrowth of the next SYM. The process is indefinitely reiterated and growth is thus indeterminate. The shoot section between two successive inflorescences is called the sympodium, and the number of leaf nodes per sympodium is referred to as the sympodial index (spi).

15
20
25

Thus, vegetative and reproductive phases alternate regularly during sympodial growth in tomato. In wild-type 'indeterminate' plants, inflorescences are separated by three vegetative nodes. As a result of this pattern of growth, a distinctive feature of tomato is its spi value. All of the red-fruited species such as *S. lycopersicum* have a mean spi of 3, whilst all green-fruited spp., including *S. pennellii*, have an average spi of 2. The spi is therefore related to fruit color and is species-specific.

Of all tomato species, *S. lycopersicum* (formerly *L. esculentum*) is the only commercially valuable species, due to its appealing and tasty fruits.

SUMMARY OF THE INVENTION

In an attempt to solve the problem of improving yield in tomato, the present inventors have now recognized that it would be desirable to produce *S. lycopersicum* plants for protected cultivation (i.e. greenhouse growth) with a spi of 2 (herein after referred to by the designation spi²), thereby increasing the density of trusses along its shoot.

In an attempt to produce a spi² *S. lycopersicum* line, *S. pennellii* LA716 (PI 24650) was back-crossed to a proprietary *S. lycopersicum* breeding line. LA716 is a self-fertile, homozygous green fruited, indeterminate accession collected in Atico, Peru and obtainable from the Tomato Genetics Stock Centre, University of California, Davis (U.S.A.). During backcrossing, selection for the spi² trait was difficult due to the fact that the spi was very variable.

In two segregating populations, it was found that the SP3D gene (AY186735, 6819 bp) was fully linked with the spi variation. Plants with SP3D homozygous for the donor LA716, gave on average 2.1 leaves between subsequent inflorescences. Heterozygous plants or plants homozygous for the *S. Lycopersicum* allele, gave higher number of leaves between fruit clusters. The linkage between the marker and the trait was 100% in the studies performed.

The inventors thus discovered that it is possible to produce spi² *S. lycopersicum* plants by crossing a plant of a *S. lycopersicum* line with a plant of a green-fruited tomato such as *S. pennellii*. Moreover, the present inventors discovered that the production of additional *S. lycopersicum* lines with spi² can be accelerated by using newly discovered markers for the spi² trait. It is expected that the spi² trait in other *Solanum* spp. is also linked to the SP3D gene. Hence, suitable markers for other spi² donor plants can be developed by the skilled person. It is therefore contemplated in the context of the present invention that spi² *S. lycopersicum* plants can be produced by similar methods using other spi² *Solanum* species as donor plants, and using a marker based on polymorphic sequences in the SP3D gene of the donor plant.

During backcrossing studies with the newly developed spi marker it was found that the spi² trait was coupled to yellow fruit color. Plants homozygous for the *S. pennellii* allele spi² produced yellow fruits.

It was however discovered that the traits for spi² and fruit color could be uncoupled. In the segregating population several plants were discovered that combined spi² with red fruit color. However, in many cases, selfing of these plants resulted in segregation for fruit color. It proved virtually impossible to select the homozygous spi² combined with homozygous red color genotypes based on phenotypic characteristics. This greatly hampers successful breeding for spi² and red color in commercial varieties.

In search for a suitable marker to allow for detecting the uncoupling of the spi and color traits in segregating populations, the inventors discovered that the gene for phytoene synthase (PSY-1 X60441.1) was linked to SP3D in *S. pennellii*. Hence, the possibility to use this gene as a marker for uncoupling the spi and color trait was investigated. It was subsequently found that a marker based on the sequence polymorphisms between the phytoene synthase gene in *S. pennellii* and in *S. lycopersicum*

proved a very robust marker system for the development of red fruited, spi² *S. lycopersicum* lines.

Thus, the present invention in one embodiment provides a method for producing a spi² *S. lycopersicum* plant comprising a step of marker
5 assisted selection using a marker linked to the family of the spi genes and/or a marker linked to a gene from the carotenoid synthesis pathway. Using this method, plants with a low spi that bear red fruits can be produced. It was hitherto unknown that spi and fruit color were intricately linked and could be uncoupled. In fact, this problem has not been addressed in the prior art.

10 Now in a first aspect, the present invention provides a method for the production of an indeterminate or semi-determinate *S. lycopersicum* plant having a sympodial index of between 1.6 and 2.4, preferably 1.8 and 2.2, and producing red-colored fruits, said method comprising:

a) crossing a plant of a recipient breeding line of an indeterminate
15 or semi-determinate *S. lycopersicum* capable of producing red-colored fruits, with a plant of a donor line of a *Solanum* spp. having a sympodial index of between 1.6 and 2.4, preferably 1.8 and 2.2;

b) collecting the seeds resulting from the cross in step (a),

c) regenerating the seeds into plants;

20 d) providing one or more backcross generations by crossing the plants of step (c) or (optionally selfed) offspring thereof with one or more plants of said recipient breeding line of *S. lycopersicum* to provide backcross plants;

e) selfing plants of step (d) and growing the selfed seed into plants;

25 f) optionally repeating said steps of backcrossing and selfing of steps (d) and/or (e);

g) identifying and selecting from the plants grown in step (c), (e) or (f) plants having a sympodial index of between 1.6 and 2.4, preferably 1.8 and 2.2, and producing red-colored fruits.

In a preferred embodiment of said method, the step of identification and selection of step g) is performed by marker-assisted selection.

The breeding line of an indeterminate or semi-determinate *S. lycopersicum* is preferably a line with a yellow (non-transparent) skin.

5 The breeding line of an indeterminate or semi-determinate *S. lycopersicum* is preferably a cherry tomato, a cherry truss tomato or a cocktail tomato.

The breeding line of *S. lycopersicum* is preferably a line possessing resistance to tobacco mosaic virus (TMV).

10 The donor line is preferably not a line of *Solanum pimpinellifolium*.

The step of marker-assisted selection in a method of the invention preferably comprises the use of a marker linked to the SP3D gene, and/or a marker linked to the PSY1 gene.

15 A suitable donor line of a *Solanum* spp. having a sympodial index of between 1.6 and 2.4, preferably 1.8 and 2.2, is the wild tomato species *S. pennellii* (LA716), which is available from the Tomato Genetics Resource Center (TGRC), Department of Plant Sciences, University of California, Davis, USA.

20 It is contemplated that also other genes than the SP3D gene that are linked to the spi² trait can suitably be used as a basis for designing suitable markers for selection purpose. In general, it is envisioned by the present inventors that suitable genes linked to the spi² trait may be selected from the gene family for self pruning genes, including SP2G,
25 Genbank accession No. AY186734; SP3D, Genbank accession No. AY186735; SP5G, Genbank accession No. AY186736; SP6A, accession No. AY186737; and SP9D, Genbank accession No. AY186738. Markers based on said genes may comprise 20-30 or larger nucleotide fragments of said genes

30 It is contemplated that also other genes than the PSY1 gene that are linked to the color trait can suitably be used as a basis for designing

suitable markers for selection purpose. In general, it is envisioned by the present inventors that any gene (or fragment thereof) linked to fruit color is suitable for use as a marker. However, since the coupling between the spi and color traits, fruit color genes that are in coupling phase with a gene from the family of self-pruning genes are preferred. Genes encoding enzymes in the carotenoid synthesis pathways (the 1-deoxy-D-xylulose-5-phosphate (DOXP) isoprenoid biosynthetic pathway in plastids) are in principle also suitable. These include phytoene synthases (PSY1, Genbank accession no. EF157835.1 and PSY2, Genbank accession no. EU021055.1), phytoene desaturase (PDS, Genbank accession no. X71023.1), zeta-carotene desaturase (ZDS, Genbank accession no. AF195507.1), and carotene isomerase (CRTISO). Phytoene synthases such as PSY1, PSY2 and LOC778345 (accession no. DQ335097.1) are however preferred, as markers based on these genes (or fragments thereof) have provided very good association and result in proper selections. Again, markers based on said genes may comprise 20-30 or larger nucleotide fragments of said genes. The accession numbers indicated above refer to the Genbank entries in the database version of June 2009.

Contemplated as being suitable for use in aspects of the invention are markers having nucleic acid with at least 80%, more preferably at least 90% sequence identity to the sequences of the genes indicated above, as well as markers capable of hybridizing under stringent conditions to the genes indicated above, or to their complementary strands.

Thus, desirable recombinant plants may be found by using markers based on the gene sequences linked to spi² and based on the gene sequences linked to fruit color.

In another aspect, the present invention provides a plant of a *S. lycopersicum* breeding line having an average sympodial index of between 1.6 and 2.4, preferably 1.8 and 2.2, and producing red-colored fruits. The term breeding line as used herein refers to an elite line having amongst other beneficial traits multiple disease and/or pest resistance traits and high

yielding fruit production characteristics, and generally refers to a plant used as a parent in the production of commercial hybrid plants used to produce marketable tomato fruits.

A plant of the invention is preferably a recombinant plant. The
5 plant of the invention preferably comprises an introgression from a plant of a donor line of a *Solanum* spp. having an average sympodial index of between 1.6 and 2.4, preferably 1.8 and 2.2, said introgression comprising genes that result in an average sympodial index of between 1.6 and 2.4, preferably 1.8 and 2.2, in the recipient plant. Said donor plant may suitably be a plant of
10 the wild tomato species *S. pennellii*, preferably *S. pennellii* LA716.

The plants of the invention preferably exhibit higher yields than plants of the breeding line lacking the introgression responsible for the spi² trait. Preferably, the total fruit weight per plant is increased by at least 3-5%, more preferably by at least 10%, still more preferably by at least 20, 30, 40, or
15 even 50% relative to a plant of said tomato breeding line lacking said introgression.

A plant of the present invention is in one embodiment obtained by a method of producing a plant as described herein.

Yield per plant is generally dependent on the type of tomato, the
20 planting density and the number of stems per plant. A good grower may reach yields that are 10 % above average. At an average planting density of 2.3 plants/m², and a single stem per plant, the yield for a small fruited truss or cluster tomato is generally about 57-58 kg/m²/yr; the yield for a cherry tomato is generally about 40 kg/m²/yr; the yield for a beef tomato is generally
25 about 60 kg/m²/yr; and the yield for a large fruited truss or cluster tomato is generally about 60-65 kg/m²/yr. Hence it is preferred that yields for the tomatoes of the invention are at least 10-15% higher than these standard yields.

It is an important advantage of the plants of the invention that the
30 picking of leaves is no longer necessary. *S. lycopersicum* plants with an

average spi³ require manual removal of on average 1 leaf per sympodium in order to optimize productivity. This is no longer necessary when using the plants of the invention as production crop.

Other aspects of the invention include fruit of a plant according to
5 the invention described above as well as seed harvested from said plant.

The present invention also contemplates the use of the markers as described herein for selecting the spi² trait in tomato plants, and for monitoring the uncoupling between yellow fruit color and spi² in crossings between *S. lycopersicum* and wild tomato species, such as *S. pennellii*.

10 The invention further provides the use of a polymorphic sequence of the SP3D gene of *L. esculentum* as a marker for the sympodial index.

DESCRIPTION OF THE FIGURES

Figure 1 shows the complete amino acid sequence of the protein
15 product of the *Solanum lycopersicum* SP3D (SP3D) gene (GenBank accession AY186735)

Figure 2 shows the complete nucleotide sequence of the *Solanum lycopersicum* SP3D (SP3D) gene (GenBank accession AY186735). Underlined is the approximate position of the CAPS marker indicated herein showing,
20 respectively, the forward primer binding site (5'-caagggttgaagttggagga), the restriction site for EcoRV (5'-gatatc) and the reverse primer binding site (5'-attctggtacgtgaccgtc).

Figure 3 shows the complete amino acid sequence of *Solanum lycopersicum* phytoene synthase (Psy1) protein (GenBank accession
25 ABM45873.1).

Figure 4 shows the complete nucleotide sequence of the *Solanum lycopersicum* phytoene synthase (Psy1) gene (GenBank accession EF157835). Underlined is the approximate position of the CAPS marker indicated herein showing, respectively, the forward primer binding site (5'-

ggtggtggaaagcaaactaata), the restriction site for EcoRV (5'-cgcg) and the reverse primer binding site (5'-tattaccccgagccttag).

Figure 5 shows the result of the spi² marker selection. Lane 1 = 100bp ladder (the bright band is 600 bp); Marker associated with spi² = lane 9; Marker associated with heterozygous genotype = lanes 2, 3, 4, 6 and 7; Marker associated with spi³ = lanes 5, 8, 10 and 11.

Figure 6 shows the result of the red fruit color marker. Lane 1, 6, 7 = 100bp ladder (the bright band is 600bp) Score red fruit color = lanes 2, 8, 9, 10, 11, 12, 13; Score heterozygous = lanes 4, 5, 16; Score yellow skin color = lanes 3, 14, 15.

Figure 7 shows the location of the restriction site in the CAPS marker for spi².

DETAILED DESCRIPTION OF THE INVENTION

As used herein, the term "tomato" means any plant, line or population formerly known under the genus name of *Lycopersicon* including but not limited to *Lycopersicon cerasiforme*, *Lycopersicon cheesmanii*, *Lycopersicon chilense*, *Lycopersicon chmielewskii*, *Lycopersicon esculentum* (now *Solanum lycopersicum*), *Lycopersicon hirsutum*, *Lycopersicon parviflorum*, *Lycopersicon pennellii*, *Lycopersicon peruvianum*, *Lycopersicon pimpinellifolium*, or *Solanum lycopersicoides*. The newly proposed scientific name for *Lycopersicon esculentum* is *Solanum lycopersicum*. Similarly, the names of the wild species may be altered. *L. pennellii* has become *Solanum pennellii*, *L. hirsutum* may become *S. habrochaites*, *L. peruvianum* may be split into *S. 'N peruvianum'* and *S. 'Callejon de Huayles'*, *S. peruvianum*, and *S. corneliomuelleri*, *L. parviflorum* may become *S. neorickii*, *L. chmielewskii* may become *S. chmielewskii*, *L. chilense* may become *S. chilense*, *L. cheesmaniae* may become *S. cheesmaniae* or *S. galapagense*, and *L. pimpinellifolium* may become *S. pimpinellifolium* (Solanacea Genome

Network (2005) Spooner and Knapp;

http://www.sgn.cornell.edu/help/about/solanum_nomenclature.html).

The term "*S. lycopersicum*", as used herein, refers to any variety or cultivar of the garden tomato.

5 The term "sympodial index", as used herein, refers to the number of leaf nodes per sympodium (i.e., between successive inflorescences).

 The term "average sympodial index", as used herein, refers to the mean number of leaf nodes per sympodium for all plants in a population, generally the arithmetic average of all values for the population. This figure
10 is preferably around 2 for plants of the present invention.

 The term "red-colored fruits", as used herein, refers to fruits having red color (e.g. as determined by visual inspection), including those having yellow skin.

 Whether a fruit is red colored can be determined by any method
15 available to tone of skill in the art. Several suitable methods include:

 1) Determining fruit color phenotype by visual comparison with a standard tomato color chart or comparator (e.g. The Greenery color scale for Tomato Color Stages, The Greenery, 2004, Barendrecht, The Netherlands, wherein "red-colored fruits" as defined herein is a color comparable to the
20 ripening stage of at least the early stage (color no. 8), but preferably later than early stage (color no. >8), and more preferably a color comparable with a ripening stage between no. 9-12; or the USDA Visual Aid TM-L-1 (February 1975) color chart for classification requirements of surface color for tomatoes, USDA, North Highlands, CA, wherein "red-colored fruits" as defined herein is
25 a color comparable to at least (5) "Light-red", indicating that more than 60% of the surface, in the aggregate, shows pinkish-red or red, but preferably (6) "Red" indicating that more than 90% of the surface, in the aggregate, shows red color.

 2) Determining fruit color using L*, a*, b* color readings of the fruit
30 skin and the puree. Measuring color with L*a*b* values is a quantitative way

to indicate color. The three coordinates of CIELAB represent the lightness of the color ($L^* = 0$ yields black and $L^* = 100$ indicates diffuse white; specular white may be higher), its position between red/magenta and green (a^* , negative values indicate green while positive values indicate magenta) and its position between yellow and blue (b^* , negative values indicate blue and positive values indicate yellow). The possible range of a^* and b^* coordinates depends on the color space that one is converting from. "Red fruit-color" is indicated by $L^*a^*b^*$ values of $L^*=38 \pm 5\%$, $a^*=19 \pm 5\%$, and $b^*=21 \pm 5\%$ for fruit skin and $L^*=48 \pm 5\%$, $a^*=23 \pm 5\%$, and $b^*=21 \pm 5\%$ for values for puree; whereas yellow is indicated by $L^*a^*b^*$ values of $L^*=47 \pm 5\%$, $a^*=1 \pm 5\%$, and $b^*=36 \pm 5\%$ for fruit skin and $L^*=64 \pm 5\%$, $a^*=0 \pm 5\%$, and $b^*=32 \pm 5\%$ for values for puree.

3) Determining fruit color by measuring lycopene content of the fruit; wherein "red fruit-color" is indicated by a lycopene content of at least 5 mg, preferably at least 6 mg or even 7 mg of lycopene/100 g fresh weight, whereas yellow is indicated by <1 mg, preferably <0.75 mg of lycopene/100 g fresh weight. *Lycopersicon pennellii* fruits have undetectable levels of lycopene as expected of this green-fruited species.

The color of the fruit as indicated herein refers to the color of ripe or mature fruits (i.e. at maturity). The term "mature" as used herein means that the contents of two or more seed cavities have developed a jellylike consistency and the seeds are well developed. External color shows at least a definite break from green to tannish-yellow, pink or red color on not less than 10 percent of the surface, preferably on at least 60 most preferably at least 90 percent of the fruit surface.

The term "crossing" as used herein refers to the fertilization of female plants (or gametes) by male plants (or gametes). The term "gamete" refers to the haploid reproductive cell (egg or pollen) produced in plants by meiosis from a gametophyte and involved in sexual reproduction, during which two gametes of opposite sex fuse to form a diploid zygote. The term generally

includes reference to a pollen (including the sperm cell) and an ovule (including the ovum). "Crossing" therefore generally refers to the fertilization of ovules of one individual with pollen from another individual, whereas "selfing" refers to the fertilization of ovules of an individual with pollen from the same individual. When referring to crossing in the context of achieving the introgression of a genomic region or segment, the skilled person will understand that in order to achieve the introgression of only a part of a chromosome of one plant into the chromosome of another plant, random portions of the genomes of both parental lines recombine during the cross due to the occurrence of crossing-over events in the production of the gametes in the parent lines. Therefore, the genomes of both parents must be combined in a single cell by a cross, where after the production of gametes from said cell and their fusion in fertilization will result in an introgression event.

The term "recipient", as used herein, refers to the plant or plant line receiving the trait, introgression or genomic segment from a donor, and which recipient may or may not have the trait, introgression or genomic segment itself either heterozygous or homozygous.

The term "breeding line", as used herein, refers to a line of a cultivated cucumber having commercially valuable or agronomically desirable characteristics, as opposed to wild varieties or landraces. In particular, the breeding line is characterized by having an excellent fruit quality (e.g. red fruits with yellow skin) and is preferably resistant to TMV and other diseases. The term includes reference to elite breeding line or elite line, which represents an essentially homozygous, e.g. inbred or doubled haploid, line of plants used to produce F1 hybrids.

As used herein, the term "hybrid" means any offspring of a cross between two genetically unlike individuals, more preferably the term refers to the cross between two (elite) breeding lines which will not reproduce true to the parent from seed.

The term “donor”, as used herein, refers to the plant or plant line from which the trait, introgression or genomic segment originates, and which donor may have the trait, introgression or genomic segment itself either heterozygous or homozygous.

5 The term “seed” as used herein includes all tissues which result from the development of a fertilized plant egg; thus, it includes a matured ovule containing an embryo and stored nutrients, as well as the integument or integuments differentiated as the protective seed coat, or testa. The nutrients in seed tissues may be stored in the endosperm or in the body of the embryo,
10 notably in the cotyledons, or both.

 The term “plant”, as used herein, refers to the vegetative growth phase essentially consisting of a single shoot, or only a limited number (2, 3, 4, or 5) shoots which produce fruits in order to optimize yield. Tomato suckers, or side shoots, may be maintained to produce additional flowering
15 shoots as long as the main shoot is strong, but side shoots above 2 or 3 are preferably removed, especially in indeterminate plants, as such tomato suckers will compete for nutrients and may result in fruits of smaller size. The term “plant” includes reference to a plant part. The term “plant part” indicates a part of the tomato plant, including single cells and cell tissues
20 such as plant cells that are intact in plants, cell clumps and tissue cultures from which tomato plants can be regenerated. Examples of plant parts include, but are not limited to, single cells and tissues from pollen, ovules, leaves, embryos, roots, root tips, anthers, flowers, fruits, stems shoots, and seeds; as well as pollen, ovules, leaves, embryos, roots, root tips, anthers,
25 flowers, fruits, stems, shoots, scions, rootstocks, seeds, protoplasts, calli, and the like.

 The term “regenerating”, as used herein, with reference to a tomato plant refers to the formation of a plant that includes a rooted shoot.

 The term “backcross”, as used herein, refers to The term backcross
30 refers to the crossing an F1 hybrid with one of the original parents. A

backcross is used to maintain the identity of one parent (species) and to incorporate a particular trait from a second parent (species). The best strategy is to cross the F1 hybrid back to the parent possessing the most desirable traits. Two or more generations of backcrossing may be necessary,
5 but this is practical only if the desired characteristic or trait is present in the F1.

The term “backcross generation”, as used herein, refers to the offspring of a backcrossing.

The term “selfed”, as used herein, means self-pollinated and includes
10 the fertilization process wherein both the ovule and pollen are from the same plant or plant line.

The term “offspring”, as used herein, refers to any progeny generation resulting from a crossing or selfing.

The term “growing”, as used herein, refers to the growth of a plant, a
15 process wherein the plant biomass is increased and which coincides with a progressive development of the plant.

The term “identifying”, as used herein, refers to a process of establishing the identity or distinguishing character of a plant, such as exhibiting a certain trait.

20 The term “selecting”, as used herein, refers to a process of picking out a certain individual from a group of individuals, usually based on a certain identity of that individual.

The term “marker-assisted selection”, as used herein, refers to the diagnostic process of identifying, optionally followed by selecting a plant from
25 a group of plants using the presence of a molecular marker as the diagnostic characteristic or selection criterion. The process usually involves detecting the presence of a certain nucleic acid sequence or polymorphism in the genome of a plant.

The term “marker”, as used herein, refers to an indicator
30 that is used in methods for visualizing differences in characteristics of nucleic

acid sequences. Examples of such indicators are restriction fragment length polymorphism (RFLP) markers, amplified fragment length polymorphism (AFLP) markers, single nucleotide polymorphisms (SNPs), microsatellite markers (e.g. SSRs), sequence-characterized amplified region (SCAR) markers, cleaved amplified polymorphic sequence (CAPS) markers or isozyme markers or combinations of the markers described herein which defines a specific genetic and chromosomal location.

The term “linked”, as used herein, with reference to markers linked to a trait, refers to a marker the presence of which in the genome of the plant coincides with the presence of the trait. Usually the term refers to a genetic marker that falls within the physical boundaries of a genomic region spanned by at least two markers having established LOD scores above a certain threshold thereby indicating that no or very little recombination between these markers and the trait locus occurs in crosses; as well as any marker in linkage disequilibrium to the trait locus; as well as markers that represent the actual causal mutations within the trait locus. The term “linked” is used in its broadest sense and indicates that the marker and the gene are located within a continuous DNA sequence of several centiMorgan. The term is used herein with reference to the linkage between markers and phenotype and refers to a distance of preferably less than 20 cM, preferably less than 10 cM, still more preferably less than 6, 5, 4, 3, 2, or 1 cM.

The term “gene”, as used herein, refers to a hereditary unit consisting of a sequence of DNA that occupies a specific location on a chromosome and that contains the genetic instruction for a particular characteristics or trait in an organism. The term “gene” thus refers to a nucleic acid (for example, DNA or RNA) sequence that comprises coding sequences necessary for the production of an RNA, or a polypeptide or its precursor. A functional polypeptide can be encoded by a full length coding sequence or by any portion of the coding sequence as long as the desired activity or functional properties (for example, enzymatic activity, ligand binding, signal transduction, etc.) of

the polypeptide are retained. The term “gene” encompasses both cDNA and genomic forms of a gene.

The term “indeterminate”, as used herein, refers to a variety that grows in an upright or gangly fashion, producing fruit throughout the growing season, in contrast to a determinate tomato plant, which grows in a more bushy shape and is most productive for a single, larger harvest, then either tapers off with minimal new growth/fruit, or dies.

The term “yellow skin”, as used herein, refers to the pigment present in the cell walls of the epidermis of the fruits, in contrast to colorless skin, in which this pigment is absent. Red-colored fruits can have yellow skin.

The term “recombinant”, as used herein with reference to a plant refers to a plant carrying a foreign (donor) gene combined, in whole or in part, in recipient genome.

As used herein, the terms “introgression”, “introgressed” and “introgressing” refer to both a natural and artificial process, and the resulting events, whereby genes of one species, variety or cultivar are moved into the genome of another species, variety or cultivar, by crossing those species. The process may optionally be completed by backcrossing to the recurrent parent.

The term “*Solanum* spp. ”, as used herein, refers to tomatoes or other members of the genus, preferably tomato.

The term “total fruit weight per plant”, as used herein, refers to the average yield of fruits over a predetermined period of time, such as a harvest period or the lifetime of a plant.

The term “increased by at least 3-5%”, as used herein, or comparative expressions, refers to a significant increase in the average values for a plant population, preferably a population of hybrid plants generated from seed.

The term “fruit”, as used herein, refers to tomatoes including a tomato product, such as fruit pulp or processed fruit, wherein the cells in the fruit comprise a genome containing the spi^2 gene and/or containing the nucleic acid sequence of markers linked to the spi^2 trait as identified herein.

Producing plants with σ^2

Plant breeders and in particular seed companies employ elite breeding lines, generally referred to as “elite lines” to provide a constant quality product. The elite lines are the result of many years of inbreeding and combine multiple superior characteristics such as high yield, fruit quality, and resistance to pests, disease, or tolerance to abiotic stress. The average yield of these elite lines is generally much higher than the original wild (landrace) accessions from which many of the modern tomato varieties are descendants. The elite lines can be used directly as crop plant, but are typically used to produce so-called F1 or single-cross hybrids, produced by a cross between two (homozygous or inbred) elite lines. The F1 hybrids thus combine the genetic properties of the two parents into a single plant. An additional benefit of hybrids is that they express hybrid vigour or heterosis, the poorly understood phenomenon that hybrid plants grow better than either (inbred) parent and show higher yields.

Backcross or pedigree selection is one method by which breeders add desirable agronomic traits to their elite breeding lines. The method involves crossing the breeding line with a line that expresses the desirable trait followed by backcrossing offspring plants expressing the trait to the recurrent parent. As a result, the selection of an individual as a parent in a breeding program is based on the performance of its forebears. Such methods are most effective in breeding for qualitatively-inherited traits, i.e traits which are present or absent.

Recurrent selection is an alternative breeding method for improving breeding lines and involves systematic testing and selection of desirable progeny followed by recombination of the selected individuals to form a new population. Recurrent selection has proven effective for improving quantitative traits in crop plants. Recurrent selection, however, decreases the

rate of broadening genetic basis underlying the various traits in a breeding program, and its potential is therefore limited.

The present inventors discovered that the yield of a tomato plant may be increased by introgressing into an elite breeding line the trait of spi^2 .

5 A first method would comprise introgressing the trait from a tomato plant having a sympodial index of about 1.8-2.2, such as a plant of the wild tomato species *S. pennellii*, such as *S. pennellii* LA716, or an offspring plant thereof having said spi^2 , into a plant of a tomato line of interest. This may for instance be achieved by crossing a plant of a recipient breeding line of *S.*
10 *lycopersicum* capable of producing red-colored fruits, with a plant of a donor line of a tomato species, preferably a *S. lycopersicum* variety, having an average sympodial index of between 1.8 and 2.2. This will result in a situation wherein the spi^2 gene is in the genetic background of the tomato line of interest. The establishment of the proper introgression in offspring
15 plants may be monitored by using specific markers as defined herein.

Recombination is the exchange of information between two homologous chromosomes during meiosis. In a recombinant plant, DNA that is originally present on a specific location within the chromosome is exchanged for DNA from another plant (i.e. maternal for paternal or *vice versa*). In order to
20 exchange only the required material, and maintain the valuable original information on the chromosome as much as possible, will usually require two crossover events. The normal way to find such a recombinant is to screen a population of F₂-plants. This population must be of sufficient size in order to detect the rare (low frequency) double recombinants. The frequency of
25 recombination can be expressed in a genetic distance. For instance, a single recombinant in a 10 cM area can be found with a frequency of 10% (1 centimorgan is defined as 1% recombinant progeny in a testcross).

The present invention also provides methods of producing the plants of the invention using marker assisted selection (MAS). The invention therefore
30 relates to methods of plant breeding and to methods to select plants, in

particular tomato plants, particularly cultivated tomato plants as breeder plants for use in breeding programs or cultivated tomato plants for having desired genotypic or potential phenotypic properties, in particular related to producing quantities of valuable tomato fruits, also referred herein to as
5 agronomically desirable plants. Herein, a cultivated plant is defined as a plant being purposely selected or having been derived from a plant having been purposely selected in agricultural or horticultural practice for having desired genotypic or potential phenotypic properties, in particular a plant obtained by inbreeding.

10 Since the gene can only be properly identified phenotypically when the plant has produced several sympodia, it is of particular advantage that the establishment of the proper introgression in offspring plants may be monitored by using the gene-specific markers as provided herein, either in cis or in trans coupling as explained below. By using marker assisted selection
15 (MAS) or marker assisted breeding (MAB) methods, the skilled person is therefore provided with methods for selecting plants carrying the desired genotype loci and discarding plants lacking the potential of producing spi^2 progeny.

The present invention thus also provides methods for selecting a
20 tomato plant exhibiting a sympodial index of about 2, comprising detecting in said plant the presence of the spi^2 gene as defined herein. In a preferred method of the invention for selecting such a plant the method comprises:

- a) providing a sample of genomic DNA from a tomato plant;
- b) detecting in said sample of genomic DNA at least one molecular
25 marker linked to the gene for spi^2 .

The step of providing a sample of genomic DNA from a tomato plant may be performed by standard DNA isolation methods well known in the art.

The step of detecting a molecular marker (step b) may, in a preferred embodiment, comprise the use of CAPS markers, which constitute a set of bi-
30 directional primers in combination with a restriction enzyme. This allows for

the detection of specific SNPs linked to the trait. Bi-directional means that the orientation of the primers is such that one functions as the forward and one as the reverse primer in an amplification reaction of nucleic acid.

Alternatively, the step of detecting a molecular marker (step b) may in
5 another preferred embodiment, comprise the use of a nucleic acid probe
having a base sequence which is substantially complementary to the nucleic
acid sequence defining said molecular marker (e.g. said SNP) and which
nucleic acid probe specifically hybridizes under stringent conditions with a
nucleic acid sequence defining said molecular marker. A suitable nucleic acid
10 probe may for instance be a single strand oligonucleotide of the amplification
product corresponding to the marker.

The step of detecting a molecular marker (step b) may also comprise
the performance of a unique nucleic acid amplification reaction on said
genomic DNA to detect said gene. This can suitable be done by performing a
15 PCR reaction using a pair of marker-specific primers based on the internal or
adjacent (up to 500 kilo base) sequence. In a preferred embodiment, said step
b) comprises the use of at least one pair of primers defining a marker for said
gene (e.g. being complementary to said marker or hybridizing specifically to
said marker or allowing polymerase chain extension to occur when bound to
20 said marker), or a pair of primers which specifically hybridize under
stringent conditions with the nucleic acid sequence of a marker for said gene.

The step of detecting an amplified DNA fragment having a certain
predicted length or a certain predicted nucleic acid sequence may be
performed such that the amplified DNA fragment has a length that
25 corresponds (plus or minus a few bases, e.g. a length of one, two or three
bases more or less) to the expected length as based on the nucleotide
sequence of the genes and markers identified herein. The skilled person is
aware that markers that are absent in plants having the introgression as
defined herein (donor plants), while they are present in the plants receiving
30 the introgression (recipient plants) (so-called trans-markers), may also be

useful in assays for detecting the introgression among offspring plants, although detecting the presence of a specific introgression is not optimally demonstrated by the absence of a marker.

The step of detecting an amplified DNA fragment having the predicted
5 length or the predicted nucleic acid sequence may be performed by standard gel-electrophoresis techniques, real time PCR, or by using DNA sequencers. The methods need not be described here as they are well known to the skilled person. It should be noted that a marker is usually defined based on its nucleotide sequences in combination with its position relative to other
10 markers on a linkage map.

Molecular Markers and genes

Molecular markers are used for the visualisation of differences in nucleic acid sequences. This visualisation is possible due to DNA-DNA
15 hybridisation techniques after digestion with a restriction enzyme (RFLP) and/or due to techniques using the polymerase chain reaction (e.g. STS, microsatellites, AFLP). All differences between two parental genotypes will segregate in a mapping population (e.g., BC₁, F₂) based on the cross of these parental genotypes. The segregation of the different markers may be
20 compared and recombination frequencies can be calculated. The recombination frequencies of molecular markers on different chromosomes is generally 50%. Between molecular markers located on the same chromosome the recombination frequency depends on the distance between the markers. A low recombination frequency corresponds to a short genetic distance between
25 markers on a chromosome. Comparing all recombination frequencies will result in the most logical order of the molecular markers on the chromosomes. This most logical order can be depicted in a linkage map. A group of adjacent or contiguous markers on the linkage map that is associated with π^2 , pinpoints the position of a gene associated with π^2 .

The markers identified herein may be used in various aspects of the invention as will now be illustrated. Aspects of the invention are not limited to the use of the markers identified herein. It is stressed that the aspects may also make use of markers not explicitly disclosed herein or even yet to be
5 identified.

In the present invention amplified fragment length polymorphism (AFLP) markers, single nucleotide polymorphisms (SNPs), and insertion deletions (INDELs), microsatellite markers, restriction fragment length
polymorphism (RFLP) markers, sequence-characterized amplified region
10 (SCAR) markers, cleaved amplified polymorphic sequence (CAPS) markers or isozyme markers or combinations of these markers might be used.
In general, a gene may span a region of several hundreds to thousands of bases. Although the sequence of the spi^2 gene has not yet been
elucidated, the plants that have the genetic potential for exhibiting a
15 particular phenotypic trait (spi^2) may be traced amongst a population of offspring plants through the observed correlation between the presence of a (string of contiguous) genomic marker(s) and the presence of the
phenotypic trait. By providing a non-limiting list of markers, the present invention thus provides for the effective utility of the genes in a breeding
20 program.

It is further important to note that the contiguous genomic markers can also be used to indicate the presence of the gene (and thus of the phenotype) in an individual plant, i.e. they can be used in marker assisted selection (MAS) procedures. In principle, the number of potentially useful
25 markers is limited but may be very large, and the skilled person may easily identify additional markers to those mentioned in the present application.
Any marker that is linked to the gene, e.g. falling within the physical boundaries of the genomic region spanned by the markers having established
Lod scores above a certain threshold thereby indicating that no or very little
30 recombination between the marker and the gene occurs in crosses; as well as

any marker in linkage disequilibrium to the gene; as well as markers that represent the actual causal mutations within the gene, may be used in MAS procedures.

5 A Lod score ("logarithmic odds") is a measure of the likelihood of two loci being within a measurable distance of each other.

This means that the markers identified herein, are mere examples of markers suitable for use in MAS procedures. Moreover, when the gene, or the specific trait-conferring part thereof, is introgressed into another genetic background (i.e. into the genome of another plant line), then some markers
10 may no longer be found in the offspring although the trait is present therein, indicating that such markers are outside the genomic region that represents the specific trait-conferring part of the gene in the original parent line only and that the new genetic background has a different genomic organisation. Such markers of which the absence indicates the successful introduction of
15 the genetic element in the offspring are called "trans markers" and may be equally suitable in MAS procedures under the present invention.

Upon the identification of the gene, the gene effect (spi^2) is confirmed by determining the sympodial index of progenies respectively recombinant or segregating for the genes under investigation. Preferably, detecting the
20 presence of a gene of the invention is performed with at least one of the markers for a gene as defined herein. The present invention therefore also relates to a method for detecting the presence of a gene for spi^2 as defined herein in tomato by the use of the said markers.

The nucleotide sequence of the genes of the present invention may be
25 resolved by determining the nucleotide sequence of one or more markers associated with said gene and designing internal primers for said marker sequences that may then be used to further determine the sequence the gene adjacent to said marker sequences. For instance the nucleotide sequence of CAPS markers may be obtained by isolating said markers from the
30 electrophoresis gel used in the determination of the presence of said markers

in the genome of a subject plant, and determining the nucleotide sequence of said markers by for instance Sanger or pyro sequencing methods, well known in the art.

In embodiments of methods for detecting the presence of a gene in a
5 tomato plant, the method may also comprise the steps of providing an
oligonucleotide or polynucleotide capable of hybridizing under stringent
hybridization conditions to a nucleic acid sequence of a marker linked to said
gene, contacting said oligonucleotide or polynucleotide with nucleic acid of a
tomato plant, and determining the presence of specific hybridization of said
10 oligonucleotide or polynucleotide to said nucleic acid.

Preferably said method is performed on a nucleic acid sample obtained
(isolated) from said tomato plant, although *in situ* hybridization methods may
also be employed. Alternatively, and in a more preferred embodiment, the
skilled person may, once the nucleotide sequence of the gene has been
15 determined, design specific hybridization probes or oligonucleotides capable
of hybridizing under stringent hybridization conditions to the nucleic acid
sequence of said gene and may use such hybridization probes in methods for
detecting the presence of a gene of the invention in a tomato plant.

20 Production of tomato plants exhibiting spi^2 by transgenic methods

According to another aspect of the present invention, a nucleic acid
(preferably DNA) sequence comprising one or more of the genes as defined
herein may be used for the production of a tomato plant exhibiting spi^2 . In
this aspect, the invention provides for the use of genes as defined herein or
25 spi^2 -conferring parts thereof, for producing a spi^2 tomato plant as defined
herein, which use involves the introduction of a nucleic acid sequence
comprising said gene in a suitable recipient plant. As stated, said nucleic acid
sequence may be derived from a suitable donor plant. A suitable source
according to the present invention for the spi^2 genes is tomato line S.

pennellii LA716 (PI 246502 available from the Agricultural Research Service (ARS-GRIN) of the US Department of Agriculture, Washington DC, USA).

The nucleic acid sequence that comprises a gene for spi^2 , or a spi^2 -conferring part thereof, may be transferred to a suitable recipient plant by
5 any method available. For instance, the said nucleic acid sequence may be transferred by crossing a plant of line PI 246502 with a selected breeding line which is spi^3 or of which the spi is to be improved, i.e. by introgression, by transformation, by protoplast fusion, by a doubled haploid technique or by embryo rescue or by any other nucleic acid transfer system, optionally
10 followed by selection of offspring plants comprising the spi^2 gene (as assessed by markers) and/or exhibiting spi^2 . For transgenic methods of transfer a nucleic acid sequence comprising a gene for spi^2 may be isolated from said donor plant by using methods known in the art and the thus isolated nucleic acid sequence may be transferred to the recipient plant by
15 transgenic methods, for instance by means of a vector, in a gamete, or in any other suitable transfer element, such as a bombardment with a particle coated with said nucleic acid sequence.

Plant transformation generally involves the construction of a vector with an expression cassette that will function in plant cells. In the present
20 invention, such a vector consists of a nucleic acid sequence that comprises a gene for spi^2 , which vector may comprise a spi^2 gene that is under control of or operably linked to a regulatory element, such as a promoter. The expression vector may contain one or more such operably linked gene/regulatory element combinations, provided that at least one of the genes
25 contained in the combinations confers spi^2 . The vector(s) may be in the form of a plasmid, and can be used, alone or in combination with other plasmids, to provide transgenic plants that exhibit spi^2 , using transformation methods known in the art, such as the *Agrobacterium* transformation system.

Expression vectors can include at least one marker gene, operably
30 linked to a regulatory element (such as a promoter) that allows transformed

cells containing the marker to be either recovered by negative selection (by inhibiting the growth of cells that do not contain the selectable marker gene), or by positive selection (by screening for the product encoded by the marker gene). Many commonly used selectable marker genes for plant transformation are known in the art, and include, for example, genes that code for enzymes that metabolically detoxify a selective chemical agent which may be an antibiotic or a herbicide, or genes that encode an altered target which is insensitive to the inhibitor. Several positive selection methods are known in the art, such as mannose selection. Alternatively, marker-less transformation can be used to obtain plants without mentioned marker genes, the techniques for which are known in the art.

One method for introducing an expression vector into a plant is based on the natural transformation system of *Agrobacterium* (See *e.g.* Horsch *et al.*, 1985). *A. tumefaciens* and *A. rhizogenes* are plant pathogenic soil bacteria that genetically transform plant cells. The Ti and Ri plasmids of *A. tumefaciens* and *A. rhizogenes*, respectively, carry genes responsible for genetic transformation of the plant. Methods of introducing expression vectors into plant tissue include the direct infection or co-cultivation of plant cells with *Agrobacterium tumefaciens*. Descriptions of *Agrobacterium* vector systems and methods for *Agrobacterium*-mediated gene transfer are provided in US Pat. No. 5,591,616. General descriptions of plant expression vectors and reporter genes and transformation protocols and descriptions of *Agrobacterium* vector systems and methods for *Agrobacterium*-mediated gene transfer can be found in Gruber and Crosby, 1993. General methods of culturing plant tissues are provided for example by Miki *et al.*, 1993 and by Phillips, *et al.*, 1988. A proper reference handbook for molecular cloning techniques and suitable expression vectors is Sambrook and Russell, 2001.

Another method for introducing an expression vector into a plant is based on microprojectile-mediated transformation (particle bombardment) wherein DNA is carried on the surface of microprojectiles. The expression

vector is introduced into plant tissues with a biolistic device that accelerates the microprojectiles to speeds of 300 to 600 m/s which is sufficient to penetrate plant cell walls and membranes. Another method for introducing DNA to plants is via the sonication of target cells. Alternatively, liposome or spheroplast fusion has been used to introduce expression vectors into plants. Direct uptake of DNA into protoplasts using CaCl_2 precipitation, polyvinyl alcohol or poly-L-ornithine has also been reported. Electroporation of protoplasts and whole cells and tissues has also been described.

Other well known techniques such as the use of BACs, wherein parts of the tomato genome are introduced into bacterial artificial chromosomes (BACs), i.e. vectors used to clone DNA fragments (100- to 300-kb insert size; average, 150 kb) in *Escherichia coli* cells, based on naturally occurring F-factor plasmid found in the bacterium *E. coli* may for instance be employed in combination with the BIBAC system to produce transgenic plants.

Following transformation of tomato target tissues, expression of the above described selectable marker genes allows for preferential selection of transformed cells, tissues and/or plants, using regeneration and selection methods now well known in the art.

20 Production of tomato plants exhibiting spi^2 by non-transgenic methods

In an alternative embodiment for producing a tomato plant exhibiting spi^2 , protoplast fusion can be used for the transfer of nucleic acids from a donor plant to a recipient plant. Protoplast fusion is an induced or spontaneous union, such as a somatic hybridization, between two or more protoplasts (cells of which the cell walls are removed by enzymatic treatment) to produce a single bi- or multi-nucleate cell. The fused cell, that may even be obtained with plant species that cannot be interbred in nature, is tissue cultured into a hybrid plant exhibiting the desirable combination of traits. More specifically, a first protoplast can be obtained from a tomato plant of accession PI 246502. A second protoplast can be obtained from a second

tomato plant variety, preferably a tomato line that comprises commercially valuable characteristics, such as, but not limited to disease resistance, insect resistance, valuable fruit characteristics, etc. The protoplasts are then fused using traditional protoplast fusion procedures, which are known in the art.

5 Alternatively, embryo rescue may be employed in the transfer of a nucleic acid comprising the gene as described herein from a donor plant to a recipient plant. Embryo rescue can be used as a procedure to isolate embryo's from crosses wherein plants fail to produce viable seed. In this process, the fertilized ovary or immature seed of a plant is tissue cultured to create new
10 plants.

 The present invention also relates to a method for improving the spi^2 of a plant of a tomato breeding line, comprising the steps of:

 a) crossing a plant of a tomato breeding line with a plant of tomato line PI 246502 or an offspring plant thereof harbouring the gene for spi^2 as
15 described herein;

 b) selecting a progeny tomato plant resulting from said crossing having an introgression from tomato accession PI 246502 or an offspring plant thereof associated with spi^2 ;

 c) selfing and/or backcrossing said progeny tomato plant selected in
20 step (b) using said tomato breeding line as a recurrent parent;

 d) selecting a progeny tomato plant resulting from the selfing or backcrossing in step (c) having an introgression from tomato accession PI 246502 or an offspring plant thereof associated with spi^2 ,

 e) repeating said steps of selfing and/or backcrossing and selection of
25 steps (c) and (d) to provide a plant of a tomato breeding line essentially homozygous for said introgression,
 wherein preferably at least one selection as performed in steps (b) or (d) is performed by marker-assisted selection.

 In a preferred embodiment of such a method, said tomato breeding
30 line is an elite line.

In an alternative preferred embodiment of the above method, the marker-assisted selection procedure comprises the selection for at least one marker as exemplified in the Examples below.

The introgression of the nucleic acid sequence comprising the spi² gene as described herein may suitably be accomplished by using traditional breeding techniques. The gene is preferably introgressed into commercial tomato varieties by using marker-assisted selection (MAS) or marker-assisted breeding (MAB). MAS and MAB involves the use of one or more of the molecular markers for the identification and selection of those offspring plants that contain one or more of the genes that encode for the desired trait. In the present instance, such identification and selection is based on selection of the gene of the present invention or markers associated therewith. MAS can also be used to develop near-isogenic lines (NIL) harboring the gene of interest, or the generation of gene isogenic recombinants (QIRs), allowing a more detailed study of each gene effect and is also an effective method for development of backcross inbred line (BIL) populations. Tomato plants developed according to this embodiment can advantageously derive a majority of their traits from the recipient plant, and derive spi² from the donor plant.

Crossing can be achieved by mechanically pollinating the female flower of one parent plant with pollen obtained from male flowers of another parent plant.

As discussed briefly above, traditional breeding techniques can be used to introgress a nucleic acid sequence encoding a gene for spi² into a recipient tomato plant requiring spi². In one method, which is referred to as pedigree breeding, a donor tomato plant that exhibits spi² and comprising a nucleic acid sequence encoding for the gene associated with spi² as defined herein is crossed with a recipient tomato plant (preferably a plant of an elite line) that exhibits agronomically desirable characteristics, such as, but not limited to, disease (e.g. TMV) resistance, insect resistance, valuable fruit

characteristics, etc., but which is spi^3 , or which requires improvement of spi towards spi^2 . The resulting plant population (representing the F_1 hybrids) is then self-pollinated and set seeds (F_2 seeds). The F_2 plants grown from the F_2 seeds are then screened for spi^2 . The population can be screened in a
5 number of different ways.

First, the population can be screened using a visual inspection of the number of sympodia. Second, marker-assisted selection can be performed using one or more of the hereinbefore-described molecular markers to identify those progeny that comprise a nucleic acid sequence encoding for spi^2 as
10 defined herein. Other methods, described above by methods for detecting the presence of a gene may be used. Also, marker-assisted selection can be used to confirm the results obtained from the spi^2 phenotype scores, and therefore, several methods may also be used in combination.

Inbred tomato plant lines exhibiting spi^2 can be developed using the
15 techniques of recurrent selection and backcrossing, selfing and/or dihaploids or any other technique used to make parental lines. In a method of recurrent selection and backcrossing, the spi^2 -conferring genetic element as disclosed herein can be introgressed into a target recipient plant (the recurrent parent) by crossing the recurrent parent with a first donor plant, which differs from
20 the recurrent parent and is referred to herein as the "non-recurrent parent". The recurrent parent is a plant of which the spi is to be improved and possesses agronomically desirable characteristics, such as, but not limited to disease resistance, insect resistance, valuable fruit characteristics, etc. The non-recurrent, or donor, parent may suitably be a plant of line PI 246502
25 which comprises a nucleic acid sequence that encodes for spi^2 . Alternatively, the donor parent can be any plant variety or inbred line that is cross-fertile with the recurrent parent and has acquired the gene for spi^2 in an earlier cross with a plant of line PI 246502, or a different accession having this trait. The progeny resulting from a cross between the recurrent parent and non-
30 recurrent parent is backcrossed to the recurrent parent. The resulting plant

population is then screened for the desired characteristics, which screening may occur in a number of different ways. For instance, the population can be screened using phenotypic screens as described herein. As an alternative to phenotypic assays, marker-assisted selection (MAS) can be performed using
5 one or more of the hereinbefore described molecular markers, hybridization probes or polynucleotides to identify progeny that comprise a nucleic acid sequence encoding the gene responsible for spi^2 .

Following screening, the F_1 hybrid plants that exhibit a spi^2 phenotype or, more preferably, genotype and thus comprise the requisite
10 nucleic acid sequence encoding for spi^2 are then selected and backcrossed to the recurrent parent for a number of generations in order to allow for the tomato plant to become increasingly elite. This process can be performed for two to five or more generations. In principle the progeny resulting from the process of crossing the recurrent parent with the non-recurrent parent are
15 heterozygous for one or more genes that encode for spi^2 .

In a preferred embodiment, a method of introducing a desired trait into a hybrid tomato variety comprises the steps of:

(a) crossing an inbred tomato parent with another tomato plant that comprises one or more desired traits, to produce F_1 progeny plants, wherein
20 the desired trait is spi^2 as conferred by the gene from PI 246502, or an offspring plant thereof;

(b) selecting said F_1 progeny plants that have the desired trait to produce selected F_1 progeny plants, preferably using molecular markers as defined herein;

25 (c) backcrossing the selected progeny plants with said inbred tomato parent plant to produce backcross progeny plants;

(d) selecting for backcross progeny plants that have the desired trait and morphological and physiological characteristics of said inbred tomato parent plant, wherein said selection preferably comprises the isolation of

genomic DNA and testing said DNA for the presence of at least one molecular marker for the gene as defined above;

(e) repeating steps (c) and (d) two or more times in succession to produce selected third or higher backcross progeny plants;

5 (f) optionally selfing selected backcross progeny in order to identify homozygous plants;

(g) crossing at least one of said backcross progeny and/or selfed plants with another inbred tomato parent plant to generate a hybrid tomato variety with the desired trait and all of the morphological and physiological
10 characteristics of hybrid tomato variety when grown in the same environmental conditions.

As indicated, the last backcross generation may be selfed in order to provide for homozygous pure breeding (inbred) progeny exhibiting spi^2 . Thus, the result of recurrent selection, backcrossing and selfing is the
15 generation of lines that are genetically homozygous for the genes associated with spi^2 as well as other genes associated with traits of commercial interest.

It should be noted that heterozygous plants having the gene for spi^2 may also be of interest as intermediate products, and such plants are
20 therefore also an aspect of the present invention.

Tomato plants and seeds

The goal of plant breeding is to combine various desirable traits in a single variety or hybrid. For commercial crops, these traits may include
25 resistance to diseases and insects, tolerance to heat and drought, reducing the time to crop maturity, greater yield, and better agronomic quality. Uniformity of plant characteristics such as germination, growth rate, maturity, and plant height may also be of importance.

Commercial crops are bred through techniques that take advantage of
30 the plant's method of pollination. A plant is self-pollinated if pollen from one

flower is transferred to the same or another flower of the same plant. A plant is sibling mated when individuals within the same family or line are used for pollination. A plant is cross-pollinated if the pollen comes from a flower on a different plant from a different family or line.

5 Plants that have been self-pollinated and selected for type for many generations become homozygous at almost all gene loci and produce a uniform population of true-bred progeny. A cross between two different homozygous lines produces a uniform population of hybrid plants that may be heterozygous for many gene loci. A cross of two plants each heterozygous at a
10 number of gene loci will produce a population of heterogeneous plants that differ genetically and will not be uniform.

 The development of a hybrid tomato variety in a tomato plant breeding program involves three steps: (1) the selection of plants from various germplasm pools for initial breeding crosses; (2) the selfing of the selected
15 plants from the breeding crosses for several generations to produce a series of inbred lines, which, individually breed true and are highly uniform; and (3) crossing a selected inbred line with an unrelated inbred line to produce the hybrid progeny (F1). After a sufficient amount of inbreeding successive filial generations will merely serve to increase seed of the developed inbred.
20 Preferably, an inbred line should comprise homozygous alleles at about 80% or more of its loci.

 An important consequence of the homozygosity and homogeneity of the inbred lines is that the hybrid created by crossing a defined pair of inbreds will always be the same. Once the inbreds that create a superior hybrid have
25 been identified, a continual supply of the hybrid seed can be produced using these inbred parents and the hybrid tomato plants can then be generated from this hybrid seed supply.

 Using the methods as described above, the skilled person will be able to produce the required inbred lines and from those produce the commercial
30 (F1) hybrid seeds by crossing said inbred lines.

The present invention will now be explained in more detail by way of the following non-limiting Examples.

EXAMPLES

5

Yield and spi^2 .

In indeterminate tomato plants (protected tomato crops) the number of leaves between trusses is on average 3. The trait is called sympodial index (spi), and *Solanum lycopersicum* has $\text{spi}=3$. Fruit yield in tomato crops in the greenhouse is determined by the number of fruits per m^2 and their weight. The number of fruits per m^2 is determined by the number of fruits per truss and the number of trusses. It is expected that the yield can be increased in indeterminate tomato plants by increasing the number of trusses per m^2 , and reducing the total number of leaves. An experiment was performed wherein the yield was measured (as total fruit weight and as number of fruits), comparing plants of a *S. lycopersicum* breeding line having spi^3 with a plant according to the present invention comprising the introgression from *S. pennellii* LA716 conferring spi^2 as described herein. It was demonstrated that the yield was increased significantly, indicated by a higher number of fruits and total yield for spi^2 plants (see Table 1).

20

Table 1.

	total weight of fruits (in grams)	number of fruits
spi^3 common	948	11
spi^2 less leaves	1626	22

Spi^2 source.

Solanum Pennellii (formerly *Lycopersicum pennellii*) and other green fruited show spi^2 , unlike the greenhouse tomato which has spi^3 (Rick, 1986

25

Report of the Tomato Genetics Cooperative (TGC)). The Tomato Genetics Resource Center (TGRC) stocklists includes the *S. pennellii* LA716 as spi². Damiaux (1985) described tomato lines with spi² using *S. Peruvianum* (formerly *L. peruvianum*). *S. pennellii* LA716 was back crossed to a *S.*
 5 *lycopersicum* line. During backcrossing, selection for the spi² trait was difficult due to the fact that the spi was very variable.

Spi marker.

Carmel-Goren et al. (Plant Molecular Biology Plant Sciences
 10 52(6):1215-1222 (2003)) published the sequences of the self-pruning gene family. We used the available sequences for sequencing LA 716, and developed CAPS markers based on the SNP differences, herein referred to as a spi-markers.

One such marker to select for spi consisted of a forward primer (5'-
 15 CAAGGGTTGAAGTTGGAGGA -3') and a reverse primer (5'-GACGGTCAGCGTACCAGAAT -3') in combination with a restriction enzyme EcoR V (GAT|ATC). It results in a banding pattern (scores) for Spi³ of two bands (289bp + 452bp) and for spi² of one band (742bp). (See fig. 5). The marker sequence is indicated in more detail in Fig. 7. The marker scores are:

20 Score 1 = 289 + 452
 Score 2 = 289 + 452 + 742
 Score 3 = 742

In two segregating populations, we determined that the SP3D gene
 25 (AY186735, 6819 bp) was fully linked with the spi variation (See Figure 5). Plants with SP3D homozygous for the donor LA716, gave on average 2.1 leaves between clusters. Heterozygous plants or plants homozygous for the *S. Lycopersicum*, gave higher number of leaves between clusters (see Table 1). The spi² trait shows to be more stable with a coefficient of variation of 5%

than the spi^3 with 15%. The linkage between the marker and the trait was 100% in the 106 plants shown in table 2.

5

Table 2.

genotype *	average number of leaves between trusses	coefficient of variation in %
SLSL	2.9	15%
SLSP	2.3	8%
SPSP	2.1	5%
*SL <i>S. lycopersicum</i> (spi^3), SP <i>S. pennellii</i> (spi^2)		

Spi linkage drag: yellow fruit color.

During back crossing the spi marker was used. We found that the trait was
 10 linked to yellow fruit color (see table). Homozygous for the *S. pennellii* allele
 spi^2 is combined with yellow fruits. We found in the segregating population
 2 plants that combined spi^2 with red fruit color. However, red fruit color is
 dominant over yellow, resulting in heterozygous plants with red fruits, and
 resulting in the next generation segregation for fruit color. The gene for
 15 yellow fruit color phytoene synthase (PSY-1 GenBank Accession X60441,
 from *L. esculentum* GTom5) is linked to SP3D in *S. pennellii*. We developed a
 marker based on the sequence differences, based on specific restriction
 endonuclease digestion of DNA of the plant or a part thereof.

One such marker to select for red color consisted of a forward
 20 primer (5'-GAGGTGGTGGAAAGCAAATA-3') and a reverse primer (5'-
 CTAAGGCTGCCGGGGTAATA-3') in combination with the restriction
 enzyme Bsh1236 I (CGCG). It results in a banding pattern (scores) for red
 color of two bands (463 + 482bp), for heterozygous genotype of three bands
 (463 bp + 482bp + 945bp) and for yellow color of one band (945 bp). (See fig.
 25 6), or, when bp countings are based on the PSY-1 sequence of GenBank

accession EF157835, marker scores are for red color (460 + 474bp), for heterozygous genotype (460 bp + 474bp + 934bp) and for yellow color (934 bp). The marker sequence is indicated in more detail in Fig. 8.

5 **Table 3.** Number of plants

genotype *	Red fruit	Yellow fruit
SLSL	24	0
SLSP	65	1
SPSP	2	22
*SL <i>S. lycopersicum</i> (<i>spi</i> ³), SP <i>S. pennelili</i> (<i>spi</i> ²)		

Spi² – red fruit color recombinant.

To combine *spi*² and red fruit color using visual evaluation was not
 10 successful. The *spi* trait was too variable to select for, and we could not
 distinguish homozygous red fruit color from heterozygous red. Therefore, we
 used the *spi* marker and the color marker to select for recombinants that
 combined *spi*² and red fruit color. Table shows that plant 19, 20, 25 and 26
 are homozygous for both *spi*² and red color, based on marker and visual
 15 data. Sequences of one recombinant plant are shown below.

Table 4.

Spi geno- type	spi pheno- type	Fruit color
SPSP	2.1	yellow
SPSP	2.2	yellow
SPSP	2.1	yellow
SPSP	2.2	yellow
SPSP	1.8	yellow
SPSP	2.2	yellow
SLSP	2.2	red
SLSP	2.4	red
SLSP	2.3	red
SLSP	2.3	red
SLSP	2.2	red
SLSP	2.2	red
SLSP	2.2	red
SLSP	2.2	red
SLSP	2.2	red
SLSP	2.4	red
SLSP	2.2	red
SLSP	2.3	red
SLSP	2.2	red
SLSP	2.2	red
SLSP	2.6	red
SLSP	2.2	red
SLSP	2.2	red
SLSP	2.3	red
SLSP	2.4	red
SLSP	2.3	red
SLSP	2.3	red
SLSP	2.4	red
SLSP	2.2	red
SLSP	2.2	red
SLSP	2.2	red
SLSP	2.0	red
SLSP	2.2	red
SLSP	2.5	red

SLSP	2.2	red
SLSP	2.4	red
SLSP	2.2	red
SLSP	2.2	red
SLSP	2.3	red
SLSP	2.2	red
SLSP	2.2	red
SLSL	2.8	red
SLSL	2.5	red
SLSL	2.5	red
SLSL	2.5	red
SLSL	3.7	red
SLSP	2.0	red
SLSL	2.5	red
SLSL	3.0	red
SLSL	3.0	red
SLSL	2.5	red
SPSP	2.1	yellow
SPSP	2.0	yellow
SPSP	2.2	yellow
SPSP	2.1	yellow
SPSP	2.1	yellow
SPSP	2.1	yellow
SPSP	1.9	yellow
SPSP	2.0	yellow
SPSP	2.1	yellow
SPSP	2.0	yellow
SPSP	2.1	red
SPSP	2.1	red
SLSP	2.7	red
SLSP	2.3	red
SLSP	2.3	red
SLSP	2.5	red
SLSP	2.3	red
SLSP	3.0	red
SLSP	2.1	red

SLSP	2.4	red
SLSP	2.7	red
SLSP	2.7	red
SLSP	2.2	red
SLSP	2.3	red
SLSP	2.2	red
SLSP	2.6	red
SLSP	2.5	red
SLSP	2.3	red
SLSP	2.3	red
SLSP	2.3	red
SLSP	2.4	red
SLSP	2.1	yellow
SLSP	2.3	red
SLSP	2.3	red
SLSP	2.2	red
SLSL	3.5	red
SLSL	3.0	red
SLSL	3.0	red
SLSL	2.8	red
SLSP	2.3	red
SLSL	3.0	red
SLSL	2.8	red
SLSP	2.4	red
SLSP	2.7	red
SLSP	2.4	red
SLSL	2.7	red
SLSL	2.4	red
SLSL	3.0	red
SLSL	3.3	red
SLSL	3.0	red
SLSL	2.6	red
SLSL	4.3	red
SLSL	3.0	red
SLSL	3.0	red

Table 5.

Plant nr.	Spi	Fruit color	spi marke r	color marker
1	2.4	Red	–	red?
2	2.8	Red	SLSP	red
3	2.2	Red	–	–
4	2.4	Red	SLSP	red?
5	2.3	Red	–	–
6	2.0	ND	–	–
7	2.2	Red	SLSP	red
8	2.2	Red	–	–
9	2.2	Red	–	–
10	2.4	ND	–	–
11	2.0	Yellow	SPSP	yellow
12	2.0	Yellow	SPSP	yellow
13	2.0	Yellow	SPSP	yellow
14	2.0	Yellow	SPSP	yellow
15	2.4	Red	SLSP	red
16	2.2	Red	SLSP	red
17	2.0	Red	–	–
18	2.4	Red	SLSP	red
19	2.0	Red	SPSP	red

20	2.0	Red	SPSP	red
21	2.0	Yellow	SPSP	–
22		Yellow	–	–
23	2.5	Red	SLSP	red
24	2.0	Red	SLSP	red
25	2.0	Red	SPSP	red
26	2.0	Red	SPSP	red
27	2.0	Red	SLSP	red
28	3.7	Red	SLSP	red
29	2.3	Red	–	–
30	2.4	Red?	SLSP	red
31	2.4	Red	–	red
32	2.0	Red?	–	–
33	2.2	Red	SLSP	red
34	2.2	Red	SLSP	red
35	2.1	Red	SLSP	red
36	2.0	Red	SLSP	red
37	3.0	Red	SLSL	red
38	2.8	Red	SLSL	red
39	2.0	Red	–	–
40	2.0	Red	SLSP	red

Claims

1. A method for the production of a *S. lycopersicum* plant having an average sympodial index of between 1.8 and 2.2 and producing red-colored fruits, said method comprising:
 - a) crossing a plant of a recipient breeding line of *S. lycopersicum* capable of producing red-colored fruits, with a plant of a donor line of the wild tomato species *S. pennellii* having an average sympodial index of between 1.8 and 2.2;
 - b) collecting the seeds resulting from the cross in step (a),
 - c) regenerating the seeds into plants;
 - d) providing one or more backcross generations by crossing the plants of step (c) or optionally selfed offspring thereof with one or more plants of said recipient breeding line of *S. lycopersicum* to provide backcross plants;
 - e) selfing plants of step (d) and growing the selfed seed into plants;
 - f) optionally repeating said steps of backcrossing and selfing of steps (d) and/or (e); and
 - g) identifying and selecting from the plants grown in step (c), (e) or (f) plants having an average sympodial index of between 1.8 and 2.2 and producing red-colored fruits, wherein the step of identification and selection of step g) is performed by marker-assisted selection comprising the use of a marker linked to the SP3D gene from the family of self-pruning genes of *S. pennellii*, and a marker linked to the phytoene synthase PSY1 gene from the carotenoid synthesis pathway of *S. lycopersicum*.
2. The method according to claim 1, wherein said SP3D gene from the family of self-pruning genes is the SP3D gene of *Solanum pennellii* LA716.
3. The method according to claim 1 or 2, wherein the marker linked to the SP3D gene from the family of self-pruning genes is provided by the presence of the sequence provided in SEQ ID NO:12 or SEQ ID NO:17; or wherein said marker linked to said SP3D gene consists of a primer having SEQ ID NO:1 and a primer having SEQ ID NO:5, in combination with the restriction enzyme EcoR V, whereby a banding pattern of one band of 742 bp is indicative for a plant having an average sympodial index of between 1.8 and 2.2.

4. The method according to any one of claims 1-3, wherein the marker linked to the phytoene synthase PSY1 gene is provided by the presence of the sequence provided in SEQ ID NO:11 or SEQ ID NO:14; or wherein said marker linked to the phytoene synthase PSY1 gene consists of a primer having SEQ ID NO:6 and a primer having SEQ ID NO:7, in combination with the restriction enzyme Bsh1236 I, whereby a banding pattern of two bands of 463 bp and 482 bp is indicative of a plant producing red colored fruits.
5. The method according to any one of claims 1-4, wherein said *S. lycopersicum* plant is an indeterminate plant.
6. The method according to any one of claims 1-5, wherein said donor line of *S. pennellii* having a sympodial index of between 1.8 and 2.2 is the wild tomato accession *S. pennellii* LA716.
7. A cell of a *S. lycopersicum* plant, said plant having an average sympodial index of between 1.8 and 2.2 and having red-colored fruits with yellow skin, whereby said cell comprises a SP3D gene from the family of self-pruning genes of *S. pennellii*, and a phytoene synthase PSY1 gene of *S. lycopersicum*.
8. The cell according to claim 7, wherein said cell comprises the SP3D gene from *S. pennellii* LA716.
9. The cell according to claim 7 or 8, which is a fruit cell.
10. The cell according to claim 7 or 8, which is a seed cell.

Figure 1

MPRERDPLVVGRVVDVLDPFTRTIGLRVIYRDREVNNGCEL RPSQVINQPRVEVGGDDL
RTFFTLVMVDPDAPSPSDPNLREYLHWLVTDIPATTGSSFGQEIVSYESPRPSMGIHRFV
FVLFRQLGRQTVYAPGWRQNFNTRDFAELYNLGLPVAAVYFNCQRESGSGGRRRSAD

[illegible]

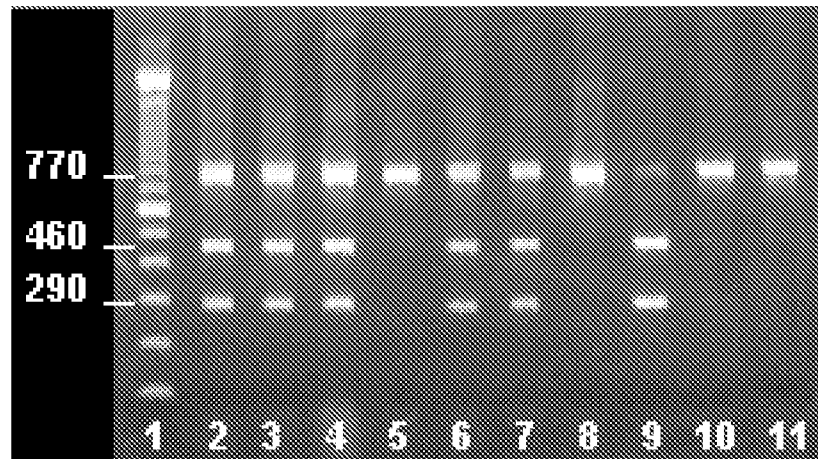
Figure 3

MSVALLWVVS PCDVSNSTSF MESVREGNRF FDSSRHRNLV SNERINRGGG KQTNNGRKFS
VRSAILATPS GERTMTSEQM VYDVVLRQAA LVKRQLRSTN ELEVKPDIPI PGNLGLLSEA
YDRCGEVCAE YAKTFNLGTM LMTPEERRAI WAIYVWCRT DELVDGPNAS YITPAALDRW
ENRLEDVFNG RPFDMLDGAL SDTVSNFPVD IQPFRDMIEG MRMDLRKSR Y KNFDELYLYC
YYVAGTVGLM SVPIMGIAPE SKATTESVYN AALALGIANQ LTNILRDVGE DARRGRVYLP
QDELAQAGLS DEDIFAGRVT DKWRIFMKKQ IHRARKFFDE AEKGVTELSS ASRFPVWASL
VLYRKILDEI EANDYNNFTK RAYVSKSKKL IALPIAYAKS LVPPTKTASL QR

Figure 4

tottgatttcttgaacaaagggtttgtttcccttcaacttcttgatagttaaagttgcaatctttataactttctattgtttgtctagt
 tttttgttatatacaggggttgaggttagagggtaagttacgcattttagtcgttaacttttagtcaaaacttctgtaataatttagtaagtta
 aaatatattagaaattttcagaattcataaaactttaaatttttaattttgacttcgctttgtgtgactatacaattacagaaattcaga
 gtggccattgttgaaagagaggggtggaattttgtgtaagttttgttcccttccagttcttgatataaaagttgcaatcttttaacattct
 ttgttcactttctataggtttgtctaggttcggttaaattcagtagcttttagtttaaaccctatgcgggaatagagaatgtgtaaacttta
 aacttcaaattttggctcgcgcatacgcactagcgcactatataataataggaattgagcaacttggctttgtatagcttctatgtgtac
 caaaattagaaaaatcagggcattattataatcttgttgactaaatatagaatgcattacccccaaaaagtggtgattccactgtca
 taggaggtttttttttttttttttttgtgtcttcaataatgtagagtagtttttcaaaagatoccttctttgtgacacatgggtagta
 atattgtctgattttgtctgtagttttgggttataaaagttcaaattatttctctggagggtaggggtgggggtgtctataatgccaggt
 tatggttttaacgtgaacctcaataaattattggtagaatctaagaaatccactcagtggtccggtgcggtggtcttctgtttgatctcag
 catcactggtagttgattgtgttttagattatcacattattctgtggctgtaactgtatccctgttagttgctttgtttctacactgttg
 ttttccctcttttataactattttgatattgtgtactcgaaacaggggtcactcggggaacaaactcttaacctcogtgaggtagagctat
 ggtctgtgtccactctacccctcccagatccctctttaggtatttctactatattgtaatttaacttgaggtcactataggagctcaaa
 aacttctaattttgaaatcaatgtctggttatactttttttgtcataactgtatctcaaatgtggtgtttgggttatctcattttgcaga
 agtcaagaaacaggttactcctgtttgagtgaggaaaggttgggttgcctgtctgtggtctttttataatctttttctacagaagagaa
 agtggttaattttttttttttttttttgtgtcttcaataatctctagttgggttgggtgagtagtaattttttctataaaactaaagttgga
 aggtgacaaaaagaaagacaaaaatcttggaaattgttttagacaaccaaggttttctgtctcagaatgtctgttgccttgttatgggtt
 gtttctccttgtgacgtctcaaatgggacaaagtttcatggaaatcagtcogggagggaaacccgtttttttgattcatcgaggcataggaa
 tttggtgtccaatgagagaatcaatagaggtggtggaagcaactaataatggacggaaattttctgtacgggtctgtctattttggcta
 ctccatctggagaacggacgatgacatcggaacagatgggtctatgatgtggttttgaggcagcgagccttggtagagggcaactgaga
 tctaccaatgagtttagaagtgagccggatatacctattccggggaatttgggttctgtgagtgagcatatgataggtgtggtgaagt
 atgtgcagagtgatgcaagacgttttaacttaggttagcttctcaatctattcattcgtttaccaaataattatttggtaagcactaatt
 atgaatatatatatgttcatgttattgatgaagacaaaatttgatctttgtttgttttatttcaggaactatgctaactgactcccgagaga
 agaagggctatctgggcaaatatggtgaggtttctagaggttataatcaacagtagtgcagcgacacaaacacatagtaaatcggggacgag
 aaaaaagaaatgaagttgagttttaggttcatatgtaataggtaaatccgagcgttgactagcttgagatgtttattgtcatatcat
 gctcaatactaaccaaaacactgaaaaagaacttgattatattacataactaatttttcatttgcgttgcgtgttcacatttttaccta
 tggaaactgggtttttgtgattttgttatacttcatattcgatgttaataaaatatatcattcctccctttttctccacttcaagctttact
 gtagtgtgaaaggggaaactccttttaattgattgcataataaaacgaacttcttgagttgaatagtttctcatttatgatctgtttaaa
 cagttatgggtgcagaagaaacagatgaactgttggatggcccaaacgcacatataattacccccggcagccttagataggtgggaaatagg
 ctagaagatgttttcaatgggogggccattttgacatgctogatgggtgctttgtccgatcacagtttctaactttccagttgatattcaggt
 tagtctaccaattctatgggtctttatatttgttcaatttgcgtttgatgtcacttttgcgtgagggcttttctaatagttacttcagcc
 tagcggaaatgtttgtagttgaactcttagttctgtctcctatattctgtttctctcgtcctagatactacacatacttcatattctgttt
 taactttttattogtcttttgggtgtttttgtatgtgaatcatataatttggacagaatcattattagttcacatgtttcatttgc
 ttcttcaatagogtaattgtctaaccttccaatataatgttgacgcattcagagatatgattgaaggaatgcgtatggacttgagaaaa
 togagatacaaaaacttcgacgaactatacctttattgttattatgttgcgtggtacgggtgggttgatgagtggtccaattatgggtat
 cgccctgaatcaaaggcaacaacagagagcgtatataatgctgctttggctcgtgggatcgcaaatcaattaactaacatactcagag
 atgttgggaaagagtagtaagtaacaaagctgtgtttacgcacataaattttttgtctaataatttaccatacaaaaatagggaaatgagct
 cttcgggttatccggtttatattttttttatgtcaacataatagtataaaagtaattagtagtcagtcgttctgggaataaaattgcagAAC
 tcaatttagccgtgttgtgaaatcctgctgttttgagagcttaaggtcatttagttagtcgttagagacgaagaaattcttcgttgtc
 catctttattccactttaaagttgtgatattttcattattggtacatttggcaaaaacacctgaacaaatttatgacgggatgccttttg
 aaagtcaactatacctgtctagtcggcgtttatcacatttctttgacatattgaaactttgaaacatgatcacgctctagacagtgacga
 gccatgatcaattttcttcttattcttcttcttggaaagtgcogtatttaggcttccgttgttcttatatattgctttccctgcagtg
 cagaagaggaagagtgctacttgcctcaagatgaattagcacagggcaggtctatccgatgaagatatatttgcgtggaagggtagccgata
 aatggagaatctttatgaagaaacaaatatacagggcaagaaagttctttagtgaggcagagaaagggcgtgacagaattgagctcagct
 agtagattccctgtaagcattctgtaaacctctttagttttatgaaatgattcttttttgcgttattagatgaataggttgcctgtgtt
 gagtatttctaggtcgatgaagttgagacaaggggttttaagttttaacgacttttacggggtgoccatgttatctgctacctaacttta
 ggtagttgacccgaagggctagaatttttaacctcatgttcacctaccaaccaagaaatgaacctcgcatagagctcgttagttatgaat
 atttgccttggcatgacattgtgoggatcatgaaatgtcttagattatggaaaaaatcattctattacatcgaatagatacatttagat
 ctaagaagcacgcgctgttgaatgagaaattctatagctcagatcttttagttttctctgaacgacctacaaaccaagggataacctt
 gtattgagcttgcgttctcagttattgcaactaacattacgtcgtgtggatcctgaaatggcttggattgctattattctggatatggc
 aaaaccattttattagtagatagatagaataactacatttgacctacaagtaacctgggttggagttacaatatcccatacctcgta
 tctttagtggttctctattttacactttgtctactattctggcaaaaataacctcactcgttactcgtgttttccaggtatgggcatc
 tttggtcttgaacgcacaaatactagatgagattgaagocaaatgactacaacaaacttcaaaaagagagcatatgtgagcaaatcaaga
 agttgattgcattacattgtcatatgcaaaatctcttgtgcctcctacaaaaactgcctctcttcaaagataaagcatgaaatgaaga
 tatatatatatatatatagcaatatatacattagaagaaaaaaggaagaagaatgttgttattgatataaatgtatatcataaat
 attaggtttagtaaacattcaatataattatctctttagttgttcttcactttatctc

Figure 5



Sequence spi²

```

AGGAGATGACCTACGTACCTTTTTCACTTTGGTAATATTTCTTATATTTTTTGTTTGGAATATAATTAAG
TTATTATTTCTATGATTTTCATAAGCAAAGTAAAAAGTATTTTTGTCTTTTGTAAAGGTTATGGTGGACC
CTGATGCTCCAAGTCCGAGTGATCCAAATCTGAGAGAATACCTTCACTGGTCCGTATTTTTCTTATTCTC
TCTTCTTTTAAATCTCTTTCTTTTTTGACCTTTTCACTTTCCATAATAATTATATTCTTTAGTAATTATA
TATCCTTTTATTTTATTTTAAAAATTGGAAAGGAGAAACGAAGAGGAGATTTTACATGTGAGGGATTTA
ATTGTAATGCAAATGGTAGAAATATATAAATGTGAAGATATATTCTTGAACCTAAAAACAACTACTAAAA
TAAAAATGAATAAAATATTTACTCTGTCAATATTCTGTACTATATTGGTCAATGAATATTTATATTATTCA
TGACTTTAAAAATAGTCAAACCGAGACATAAGGTAAAAGTCAAATACGTTTAAGCTCATTATATAAATG
AATATTTTTAAATTTTGTTCATCCATCAAATATCTACTTTTTAAGAATGATATTTATTTTATAATATTC
ATATTTGATTTCGTTGATGGATAGATTTTATTCTGTAAGAAATTAATAAAAAATAAAAAATTTAGGCCTAGTC
ATATCCATCTAAAATGGGTGAGATT

```

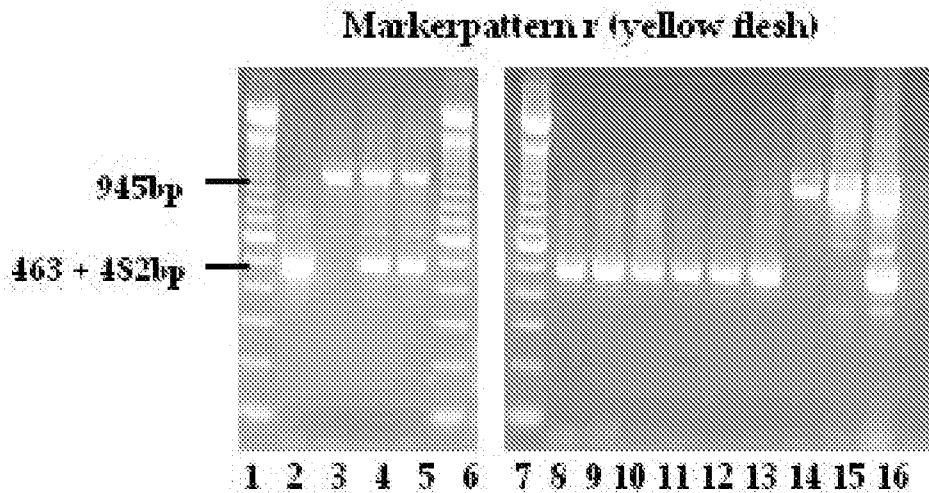
Sequence spi³

```

GAGGAGATGACCTACGTACCTTTTTCACTTTGGTAATATTTCTTATATTTTTTGTTTGGAATATAGTTAA
GTTGATTTTCATAAGCAAAGTAAAAAGTATTTTTGTCTTTTGTAAAGGTTATGGTGGACCCTGATGCTCC
AAGTCCGAGTGATCCAAATCTGAGAGAATACCTTCACTGGTCCGTATTTTTCTTATTCTCTCTTTT
CATCTCTTTCTTTTTTGACCTTTTTACTTAATTATATTCTTTAGTAATAATATATGATGATATCCTTTTAA
AAAATTGGAATACGAAAAGGAGAAATGAAGAGGAGATTTACATGTGAGGGAGCAGATGGTAGAAATATAT
AAATGTGAAGATATATATTCTTGAACCTAAAAACAAGCTACTAAAATAAAAAATGAATAAAATATTTACTCT
GTCAATATTCTGTACTATATTGGTCAATGAATATTTATATTATTCATGACTTTAAAAATAGTCAAACGAGA
CATAACGTAAAAGTCAAATACGTTTAAGCTCATTATATAAATGAATATTTTTAAATTTGTTCATCCA
TCAAATATCTACTTTTTAAGGAATGATATTTATTTTATAATATTCATATTTGATTTCGTTGATGGATAGAT
TTTATTCTTTAAAAAATTAATAAAAAAATAAAATTGGCCTAGTCATATCCATCTAAAATGGGTGA

```

Figure 6



Sequence red color

TCTGGAGAACGGACGATGACATCGGAACAGATGGTCTATGATGTGGTTTTGAGGCAGGCAGCCTTGGTGAA
GAGGCAACTGAGATCTACCAATGAGTTAGAAGTGAAGCCGGATATACCTATTCGGGGGAATTTGGGCTTGT
TGAGTGAAGCATATGATAGGTGTGGTGAAGTATGTGCAGAGTATGCAAAGACGTTTAACTTAGGTTAGCTT
CTTCAATCTATTCATTCGTTTACCAAATATTATTTGGTAAGCACTAATTATGAATATATATATGTTTCATGT
TATTGATGAAGACAAAATTTGATCTTTGTTTGTATTTCAGGAACATATGCTAATGACTCCCGAGAGAAGAA
GGGCTATCTGGGCAATATATGGTGAGGTTTCTAGCCATTTAATAACAGTTACGCGCACAAACACATATGAT
TAATCGGGGACGAGAAAAAAGAAATGAAGTTTGAGTTTTGAGGGTCATATGTAATAGGTAAATCCGAGCT
TGACTAGCTTGAGATGTTTATTGTCATATCATGCTCAATAACCAAACACTGAAAAAGAACCTTGATTA
TATTTACATACTAATAATTTTCATTTGCGTTGCTGTTTCACATTTTTTACCTATGGAACGGTTTTTGTGATTT
GTTATACCTCATATTCGATGTTAATAAAATATATCATTCCTCCCTTTTTCTCCACTTCAAGCTTTACTGTA
GTGTTGAAAGGGGAAACTCCTTTTAATGATTGCATATATAAACGAACCTTCTTGAGTTGAATAGTTTCTCAT
TATGATCTGTT

Sequence yellow color

TCTGGTGAACGGACGATGACATCGGAACAGATGGTCTATGATGTGGTTTTGAGGCAGGCAGCCTTGGTGAA
GAGGCAGCTGAGATCTACCAATGAGTTAGAAGTGAAGCCGGATATACCTATTCGGGGGAATTTGGGCTTGT
TGAGTGAAGCATATGATAGGTGTGGTGAAGTATGTGCAGAGTATGCAAAGACGTTTAACTTAGGTTAGCTT
CTTTAATCTATTCATTTGTTTACCAAATATTATTTGGTAAGCACTAATTATGACTATATATATATATATAT
ATATATATATATATATATATATATATCTGTTTCATGTTATTGATGAAGACAAAATTTGATCTTTGTTTGTAT
TCAGGAACATATGCTAATGACTCCCGAGAGAAGAAGGGCTATCTGGGCAATATATGGTGAGGTTTCTAGCCA
TTTAATAACAGATACGCACACAAACACATATGATTAATCGGAGACGAGAAAAAAGAAATGAAGTTTGAGT
TTGAGGGTCATATATAATAGGTAAATCCGAGCTTGACTAGCTTGAGATGTTTATTGTCATATCATGCTCAA
TACTAACCAAACACTGAAAAAGAACATGATTATATTTACATACTAATAATTTTCATTTGCGTTGCTGTTCA
CATTTTTACCTATGGAACGGTTTTTTTTGATTTGTTATACCTTCATATTCGATGTTAATAAACTATATCAT
TCCTCCCTTTTTCTCCACTTCAAGCTTTACTGTAGTGTGAAAGGGGAAACTCCTTTTAATGATTGCATAT
ATAAACGAACCTTCTTGAGTTGAAAAATTTCTCATTATGATCTGTTAAACAGTATG

