

# ISOPROPYLMALATE SYNTHASE FROM NICOTIANA TABACUM AND METHODS AND USES THEREOF

## FIELD OF THE INVENTION

5

The present invention discloses the isopropylmalate synthase gene from *Nicotiana tabacum* and variants, homologues and fragments thereof. In particular there is described the modification of the expression of this gene or the activity of the protein encoded thereby in order to alter the sucrose ester composition of a plant - such as a tobacco plant.

10

## BACKGROUND OF THE INVENTION

Sucrose esters are widely recognized as having insecticidal properties in plants and have shown toxicity toward soft-bodied insects, including aphids, mites, pear psylla, and whiteflies, as well as antibiotic activity. Sucrose esters are also considered as flavour precursors and are accumulated on the leaf surface during the life of the plant. They are stable compounds that can be identified in cured and smoking material. Esterified small carboxylic acids are released from sucrose esters upon heating. These small carboxylic acids are very potent flavour molecules and are considered to be responsible in part for the oriental flavour of tobacco. Sucrose esters are produced in the glandular trichome cells of plants. The glandular trichome cells are also the place for synthesis of other leaf exudates - such as proteins (phyloplanins) and diterpenoids (cembrenoids and labdenoids). The sucrose and the small carboxylic acids that are esterified to the sucrose to give rise to the sucrose esters are produced via two distinct metabolic blocks. Tobacco varieties differ in their quantities and qualities of the sucrose ester present on the surface of the leaf. As a general rule, flue cured, burley and Maryland varieties accumulate low amounts of sucrose esters with acyl groups up to five carbon chain length whereas most oriental varieties and many cigar tobacco types accumulate high amounts of sucrose esters having acyl groups up to six carbon chain length, and to a lesser extent, esters having acyl groups with a seven carbon chain length can be observed. The dichotomy observed in the sucrose ester produced among tobacco varieties is linked to a single dominant locus called BMVSE (beta-methyvaleryl containing sucrose esters) for its functional allele and *bmvse* for its non functional allele. The BMVSE genomic locus has been positioned to chromosome A of tobacco.

Kroumova and Wagner (2009) *General and Applied Plant Physiology* 35, 3-4, p95-110 describe attempts to change sucrose ester acyl group content in various plants using a reverse genetics approach to knock down the expression of isopropylmalate synthase expression using double stranded interfering RNA. In this study, the isopropylmalate synthase gene from *Solanum pennelli* was used to amplify part of the genes from cDNA. They were then introduced in the sense and antisense orientation into a double stranded interfering RNA construct and

40

transformed via *Agrobacterium* into *Nicotiana tabacum* T.I. 1068, *Nicotiana glutinosa* cv. 24a and *S. pennellii*. The results obtained using *N. tabacum* plants showed some changes in the acyl group abundance of sugar esters through reduced beta-methylvaleryl and increased 2-methylbutyryl-acylation compared to non-transformed controls. However, the *N. tabacum* plants that were obtained were phenotypically impaired since they were chlorotic and some had curly leaves. Likewise, the transformed *N. glutinosa* and *S. pennellii* plants were phenotypically impaired. The authors conclude that isopropylmalate synthase is a critical enzyme and that its impairment probably would lead to an unhealthy plant.

- 10 There is a need in the art for plants in which the sucrose ester composition is modulated and whilst minimising undesirable effects on the plant. It is an object of the present invention to satisfy this need.

## SUMMARY

- 15 The present disclosure is based, at least in part, on the surprising finding that modulating the activity or expression of the isopropylmalate synthase gene from *Nicotiana tabacum* results in a change in sucrose ester composition with fewer undesirable effects on the overall metabolism of the plant. Most of the effects are concentrated in the trichomes of the plant and to secondary metabolism. Thus, modulating the activity or expression of the isopropylmalate synthase gene from *Nicotiana tabacum* does not substantially result in alterations in the visual appearance of the plant as compared to a control plant. This is advantageous because plants can be used for the commercial production of various products where alterations in visual appearance would either not be acceptable to the industry or could result in unacceptably reduced production yields. Advantageously the flavour profile of the aerosol generated from tobacco can be modified such that new flavour profiles can be created. Moreover, the pest resistance of the plants in which the sucrose ester profile has been modified can be changed. Compositions of sucrose esters can be extracted from the plants for various uses - such as in drugs, food additives smoking flavourants and as components of organic pesticides and the like.
- 25
- 30 In an aspect, there is provided a mutant, non-naturally occurring or transgenic tobacco plant, wherein the plant comprises a plant cell comprising:
- (i) at least one polynucleotide comprising, consisting or consisting essentially of a sequence encoding an isopropylmalate synthase and having at least 80% sequence identity to SEQ ID NO:1 or SEQ ID NO:10 or SEQ ID NO: 12 or SEQ ID NO:14; or
  - 35 (ii) a polypeptide encoded by the polynucleotide(s); or
  - (iii) a polypeptide having at least 80% sequence identity to SEQ ID NO:2 or SEQ ID NO:11 or SEQ ID NO:13 or SEQ ID NO:15; or

(iv) a construct, vector or expression vector comprising the polynucleotide sequence(s), optionally in which the construct, vector or expression vector additionally comprises a promoter comprising, consisting or consisting essentially of the sequence set forth in SEQ ID NO:8 or a variant thereof with at least 60% identity thereto or a trichome promoter or a naturally occurring isopropylmalate synthase promoter; and wherein the visual appearance of said tobacco plant is substantially the same as the control tobacco plant; and wherein the chlorophyll content and/or the colouration of the leaves of the mutant, non-naturally occurring or transgenic plant is substantially the same as the control plant at three months after transplant or 36 days after topping.

In another aspect, there is provided a method for modulating the quantity or type of sucrose esters in at least a part of a tobacco plant, comprising the steps of:

(i) modulating the expression or activity of isopropylmalate synthase in the plant, in which the isopropylmalate synthase comprises a polynucleotide comprising, consisting or consisting essentially of a sequence encoding isopropylmalate synthase and having at least 80% sequence identity to SEQ ID NO:1 or SEQ ID NO:10 or SEQ ID NO: 12 or SEQ ID NO:14; or is encoded by a polypeptide encoded by the polynucleotide(s); or in which the isopropylmalate synthase is encoded by a polypeptide having at least 80% sequence identity to SEQ ID NO:2 or SEQ ID NO:11 or SEQ ID NO:13 or SEQ ID NO:15;

(ii) optionally measuring the quantity or determining the type of one or more sucrose esters in at least a part of the mutant, non-naturally occurring or transgenic plant obtained in step (i); and

(iii) identifying a mutant, non-naturally occurring or transgenic plant in which the quantity or type of the one or more sucrose esters therein has changed in comparison to

a control plant in which the expression or activity of isopropyl malate synthase has not been modulated and, in which the visual appearance of the mutant, non-naturally occurring or transgenic plant is substantially the same as the control plant, and wherein the chlorophyll content and/or the colouration of the leaves of the mutant, non-naturally occurring or transgenic plant is substantially the same as the control plant at three months after transplant or 36 days after topping.

In another aspect, there is provided plant material including biomass, seed or leaves comprising cells or tissue from a plant which is disclosed herein.

In another aspect, there is provided a tobacco product comprising a part of a plant which is disclosed herein.

In another aspect, there is provided a method for producing a composition comprising one or more sucrose esters comprising the steps of:

(i) providing at least part of a mutant, non-naturally occurring or transgenic tobacco plant which is disclosed herein;

5 (ii) extracting the sucrose ester(s) therefrom; and

(iii) optionally, isolating or purifying the extracted sucrose ester(s).

In another aspect, there is provided a method for modulating the flavour of tobacco or a tobacco product comprising: (i) adding to tobacco or a tobacco product, a part of a plant, preferably

10 leaves, from a mutant, non-naturally occurring or transgenic plant which is disclosed herein; or (ii) adding to tobacco or a tobacco product a composition comprising sucrose esters obtained or obtainable by a method which is disclosed herein.

In another aspect, there is provided a method for producing beta-methylvaleric acid comprising the steps of:

15

(i) providing at least part of a mutant, non-naturally occurring or transgenic plant which is disclosed herein;

(ii) hydrolysing the material or an extract thereof provided in step (i); and

(iii) optionally isolating or purifying the beta-methylvaleric acid.

20

#### **ASPECTS AND EMBODIMENTS OF THE INVENTION**

Aspect and embodiments of the present invention are set forth in the accompanying claims. In a first aspect there is provided an isolated polynucleotide comprising consisting or consisting essentially of a sequence encoding an isopropylmalate synthase and having at least 60% sequence identity to SEQ ID NO: 1 or SEQ ID NO: 10 or SEQ ID NO: 12 or SEQ ID NO: 14. In another aspect, there is provided an isolated polypeptide encoded by the polynucleotide.

25

In another aspect, there is provided an isolated polypeptide having at least 60% sequence identity to SEQ ID NO: 2 or SEQ ID NO: 11 or SEQ ID NO: 13 or SEQ ID NO: 15.

A construct, vector or expression vector comprising the isolated polynucleotide sequence, optionally wherein said construct, vector or expression vector additionally comprises a promoter comprising, consisting or consisting essentially of the sequence set forth in SEQ ID NO:8 or a variant thereof with at least about 60% identity thereto or a trichome promoter or the natural isopropylmalate synthase promoter.

In another aspect, there is provided an isolated polynucleotide comprising, consisting or consisting essentially of SEQ ID NO:8 or a variant thereof with at least about 60% identity thereto.

In another aspect, there is provided a mutant, non-naturally occurring or transgenic plant cell comprising at least one of the polynucleotides, at least one of the polypeptides or the construct, vector or expression vector of the present invention. Suitably, the expression of isopropylmalate synthase or the activity of the protein encoded thereby is modulated and at least a part of the plant has a change in the composition of sucrose esters as compared to a control plant in which the expression or the activity of isopropylmalate synthase has not been modulated. Suitably, the visual appearance of the plant is substantially the same as the control plant.

In another aspect, there is provided a mutant, non-naturally occurring or transgenic plant comprising the plant cell of the present invention.

In another aspect, there is provided a method for modulating the quantity of sucrose esters in a part of a plant, comprising the steps of: (i) modulating the expression or activity of isopropylmalate synthase in the plant, preferably, wherein the isopropylmalate synthase comprises the polynucleotide sequence or the polypeptide sequence described herein; (ii) measuring the quantity of sucrose esters in at least a part of the mutant, non-naturally occurring or transgenic plant obtained in step (i); and (iii) identifying a mutant, non-naturally occurring or transgenic plant in which the quantity of sucrose esters therein has changed in comparison to a control plant in which the expression or activity of isopropylmalate synthase has not been modulated and, preferably, wherein the visual appearance of said mutant, non-naturally occurring or transgenic plant is substantially the same as the control plant.

In another aspect, there is provided a mutant, non-naturally occurring or transgenic plant that is obtained or obtainable by this method.

In another aspect, there is provided a mutant, non-naturally occurring or transgenic plant, wherein expression of isopropylmalate synthase or the activity of the protein encoded thereby is modulated and at least a part of the plant has a change in the composition of sucrose esters as compared to a control plant in which the expression or the activity of isopropylmalate synthase has not been modulated and wherein the visual appearance of said plant is substantially the same as the control plant.

In another aspect, there is provided plant material including biomass, seed or leaves comprising cells or tissue from the plant.

In another aspect, there is provided a tobacco product comprising a part of the plant or plant material according to the present invention.

In another aspect, there is provided a method for producing a composition of sucrose esters comprising the steps of: (i) providing at least part of a mutant, non-naturally occurring or transgenic plant, plant material or the tobacco product according to the present invention; (ii) extracting the (for example, one or more) sucrose esters therefrom; and (iii) optionally, purifying the extracted sucrose esters.

A composition of sucrose esters obtained or obtainable by said method is also provided.

Suitably, the one or more sucrose esters have the structure as shown in Figure 5 and wherein R3 is acetyl or hydrogen, preferably acetyl; one or more of R1, R2 and R4 comprise at least one acyl chain with 6 carbons, preferably beta-methylvaleryl; and R5 is an acetyl or hydrogen, preferably, hydrogen.

Suitably, the one or more sucrose esters are selected from the group consisting of a sucrose ester as shown in Figure 5 wherein R3=acetyl, R1=propionyl or an isomer thereof, R2=propionyl or an isomer thereof, R4=hexanoyl or an isomer thereof and R5 is hydrogen or acetyl; or a sucrose ester as shown in Figure 5 wherein R3=acetyl, R1=propionyl or an isomer thereof, R2=valeryl or an isomer thereof, R4=valeryl or an isomer thereof and R5 is hydrogen or acetyl; or a sucrose ester as shown in Figure 5 wherein R3=acetyl, R1=butyryl, R2=valeryl or an isomer thereof, R4=hexanoyl or an isomer thereof and R5 is hydrogen or acetyl (C2C14:0); a sucrose ester as shown in Figure 5 wherein R3=acetyl, R1=propionyl or an isomer thereof, R2=valeryl or an isomer thereof, R4=hexanoyl or an isomer thereof and R5 is hydrogen or acetyl; or a sucrose ester as shown in Figure 5 wherein R3=acetyl, R1=valeryl or an isomer thereof, R2=valeryl or an isomer thereof, R4=valeryl or an isomer thereof and R5 is hydrogen or acetyl; or a sucrose ester as shown in Figure 5 wherein R3=acetyl, R1=butyryl, R2=hexanoyl or an isomer thereof, R4=hexanoyl or an isomer thereof and R5 is hydrogen or acetyl (C2C15:0); a sucrose ester that has the general structure shown in Figure 5 and wherein R3 is acetyl, R1 is valeryl or an isomer thereof, R2 is valeryl or an isomer thereof, R4 is hexanoyl or an isomer thereof and R5 is either a hydrogen atom or an acetyl moiety (C2C16:0); or a sucrose ester that has the general structure shown in Figure 5 and wherein R3 is acetyl, R1 is propionyl or an isomer thereof, R2 is hexanoyl or an isomer thereof, R4 is hexanoyl or an isomer thereof and R5 is either a hydrogen atom or an acetyl moiety (C2C16:0); a sucrose ester that has the general structure shown in Figure 5 and wherein R3 is acetyl, R1 is valeryl or an isomer thereof, R2 is hexanoyl, R4 is hexanoyl and R5 is either a hydrogen atom or an acetyl moiety (C2C17:0); or a sucrose ester that has the general structure shown in Figure 5 and wherein R3 is acetyl, R1 is hexanoyl, R2 is hexanoyl, R4 is hexanoyl and R5 is either a hydrogen atom or an acetyl moiety (C2C18:0) or a combination thereof.

Suitably, the one or more sucrose esters are selected from the group consisting of a sucrose

ester as shown in Figure 5 wherein R3=acetyl, R1=propionyl or an isomer thereof, R2=valeryl or  
 an isomer thereof, R4=hexanoyl or an isomer thereof and R5 is hydrogen or acetyl; or a sucrose  
 ester as shown in Figure 5 wherein R3=acetyl, R1=valeryl or an isomer thereof, R2=valeryl or  
 an isomer thereof, R4=valeryl or an isomer thereof and R5 is hydrogen or acetyl; or a sucrose  
 5 ester as shown in Figure 5 wherein R3=acetyl, R1=butyl, R2= hexanoyl or an isomer thereof,  
 R4=hexanoyl or an isomer thereof and R5 is hydrogen or acetyl (C2C15:0); a sucrose ester  
 that has the general structure shown in Figure 5 and wherein R3 is acetyl, R1 is valeryl or an  
 isomer thereof, R2 is valeryl or an isomer thereof, R4 is hexanoyl or an isomer thereof and R5 is  
 either a hydrogen atom or an acetyl moiety (C2C16:0); or a sucrose ester that has the general  
 10 structure shown in Figure 5 and wherein R3 is acetyl, R1 is propionyl or an isomer thereof, R2 is  
 hexanoyl or an isomer thereof, R4 is hexanoyl or an isomer thereof and R5 is either a hydrogen  
 atom or an acetyl moiety (C2C16:0); a sucrose ester that has the general structure shown in  
 Figure 5 and wherein R3 is acetyl, R1 is valeryl or an isomer thereof, R2 is hexanoyl, R4 is  
 hexanoyl and R5 is either a hydrogen atom or an acetyl moiety (C2C17:0); or a sucrose ester  
 15 that has the general structure shown in Figure 5 and wherein R3 is acetyl, R1 is hexanoyl, R2 is  
 hexanoyl, R4 is hexanoyl and R5 is either a hydrogen atom or an acetyl moiety (C2C18:0) or a  
 combination thereof.

In a further aspect, there is provided a method for modulating the flavour of tobacco or a  
 tobacco product comprising: (i) adding to tobacco or a tobacco product, part of a plant,  
 20 preferably leaves, from a mutant, non-naturally occurring or transgenic plant or plant material as  
 described herein, preferably, wherein said mutant, non-naturally occurring or transgenic plant or  
 plant material belongs to the genus *Nicotinia*; or (ii) adding to tobacco or a tobacco product a  
 composition comprising beta-methylvaleryl containing sucrose esters obtained or obtainable by  
 the method described herein.

25 In a further aspect, there is provided a method for producing beta-methylvaleric acid comprising  
 the steps of: (i) providing at least part of the mutant, non-naturally occurring or transgenic plant,  
 the plant material or the tobacco product; (ii) hydrolysing the material provided in step (i) or an  
 extract thereof; and (iii) optionally isolating or purifying the beta-methylvaleric acid. Suitably, the  
 plant is a tobacco plant – such as a plant of the genus *Nicotiana* or of the species *Nicotiana*  
 30 *tabacum*.

Suitably, the visual appearance of the plant is substantially the same as the control plant.  
 Suitably, the visual appearance of the plant is substantially the same as the control plant three  
 months after field transplant or 36 days after topping, preferably, wherein the stalk height of the  
 mutant, non-naturally occurring or transgenic plants is substantially the same as the stalk height  
 35 of the control plants three months after field transplant or 36 days after topping and/or the  
 chlorophyll content of the mutant, non-naturally occurring or transgenic plants is substantially  
 the same as the chlorophyll content of the control plants three months after field transplant or 36

- days after topping. In other embodiments, any one or more of the following characteristics: degree of maturity, number of leaves per plant, stalk height, leaf insertion angle, leaf size (width and length), internode distance, lamina-midrib ratio, and colouration of the leaves of the mutant, *non-naturally occurring or transgenic plants is substantially the same as the control plants at*
- 5 three months after field transplant or 36 days after topping.
- A further aspect relates to biomass, seed or leaves comprising cells or tissue from the plant(s) described herein as is a tobacco product comprising a part of the plant or biomass, seed or leaves.
- Tobacco material identified or identifiable by this method is also provided in a further aspect of
- 10 the disclosure.
- The pest resistance of the plant may be modulated.
- Still further aspects of the present invention are set forth below.
- A chimeric gene comprising the isolated polynucleotide operably linked to one or more *regulatory sequences.*
- 15 An isopropylmalate synthase polynucleotide construct comprising, consisting or consisting essentially of at least 15-30 nucleotides, 30-50 nucleotides, 50-100 nucleotides, 100-150 nucleotides, 150-200 nucleotides, 200-300 nucleotides, 300-400 nucleotides, 400-500 nucleotides, 500-600 nucleotides, 600-700 nucleotides, 700-1000 nucleotides, 1000-1300 nucleotides, 1300-1500 nucleotides or 1500-1900 nucleotides.
- 20 A consumable product incorporating or utilising plant material, biomass, seed or leaves according to the present invention.
- A cell or cell line comprising the isolated polynucleotide, the chimeric gene, the polynucleotide construct, the double-stranded RNA, the conjugate or the expression vector and the like *according to the present invention.*
- 25 A method for modulating the expression of isopropylmalate synthase or the activity of the protein encoded thereby in a cell, said method comprising administering the chimeric gene, the polynucleotide construct, the double-stranded RNA, the conjugate or the expression vector described herein.
- A method for detecting, isolating, amplifying or analysing an isopropylmalate synthase
- 30 polynucleotide, the method comprising the step of providing a sample comprising a polynucleotide and hybridising said polynucleotide to a polynucleotide molecule comprising a nucleotide sequence of at least 10 contiguous nucleotides from the isolated nucleotide sequence according to the present invention.
- Use of an agent that modulates the expression of isopropylmalate synthase or the activity of the*
- 35 *protein encoded thereby for modulating the sucrose ester content in at least a part of a plant -- such as the leaves.*
- In one embodiment, the agent is or is derived from isopropylmalate synthase nucleic acid, a

chimeric isopropylmalate synthase gene, a polynucleotide construct comprising an isopropylmalate synthase polynucleotide, an antisense RNA, a double-stranded RNA, a cDNA, a conjugate comprising an isopropylmalate synthase polynucleotide and at least one non-nucleotide or non-polynucleotide moiety covalently attached thereto, a ribozyme, a mutagen, a zinc finger, a small molecule or a meganuclease.

In another embodiment, the polynucleotide fragment(s) encodes an antisense nucleic acid, a ribozyme, an RNA that effects spliceosome-mediated trans-splicing, an interfering RNA (RNAi), a guide RNA, or other non-translated RNA and the like. In another embodiment, the polynucleotide fragment(s) encodes an RNAi.

A further aspect relates to a method of producing a tobacco product comprising the steps of: (a) obtaining seed from the transgenic, mutant or non-naturally occurring plant described herein; (b) planting and growing the seed into a plant; (c) harvesting the plant; and (d) preparing a tobacco product from the harvested plant.

The above-mentioned embodiments are disclosed as embodiments of each of the aspects described above.

#### BRIEF DESCRIPTION OF THE DRAWING

**Figure 1** illustrates the estimated quantity of sucrose ester isomers measured on leaf exudates. The measurement is performed in acetone/methanol washes of green leaf discs from a *bmvse* tobacco variety Hicks Broadleaf (HBL) and from a BMVSE tobacco variety Red Russian (RR). The quantity is estimated from the addition of sucrose octaacetate as an external standard.  $n=4$  for each variety. C2C12:0 includes a (saturated) sucrose ester as shown in Figure 5 wherein R3=acetyl, R1=butyryl, R2=propionyl or an isomer thereof, R4=valeryl or an isomer thereof and R5 is hydrogen or acetyl; or a sucrose ester as shown in Figure 5 wherein R3=acetyl, R1=acetyl, R2=valeryl or an isomer thereof, R4=valeryl or an isomer thereof and R5 is hydrogen or acetyl. C2C13:0 includes a (saturated) sucrose ester as shown in Figure 5 wherein R3=acetyl, R1=butyryl, R2= butyryl, R4=valeryl or an isomer thereof and R5 is hydrogen or acetyl; or a sucrose ester as shown in Figure 5 wherein R3=acetyl, R1=butyryl, R2= propionyl or an isomer thereof, R4=hexanoyl or an isomer thereof and R5 is hydrogen or acetyl; or a sucrose ester as shown in Figure 5 wherein R3=acetyl, R1=propionyl or an isomer thereof, R2= propionyl or an isomer thereof, R4=hexanoyl or an isomer thereof and R5 is hydrogen or acetyl; or a sucrose ester as shown in Figure 5 wherein R3=acetyl, R1= butyryl, R2=valeryl or an isomer thereof, R4= valeryl or an isomer thereof and R5 is hydrogen or acetyl; or a sucrose ester as shown in Figure 5 wherein R3=acetyl, R1=butyryl, R2=valeryl or an isomer thereof, R4=hexanoyl or an isomer thereof and R5 is hydrogen or

acetyl. C2C15:0 includes a (saturated) sucrose ester as shown in Figure 5 wherein R3=acetyl, R1=propionyl or an isomer thereof, R2=valeryl or an isomer thereof, R4=hexanoyl or an isomer thereof and R5 is hydrogen or acetyl; or a sucrose ester as shown in Figure 5 wherein R3=acetyl, R1=valeryl or an isomer thereof, R2=valeryl or an isomer thereof, R4=valeryl or an isomer thereof and R5 is hydrogen or acetyl; or a sucrose ester as shown in Figure 5 wherein R3=acetyl, R1=butyryl, R2=hexanoyl or an isomer thereof, R4=hexanoyl or an isomer thereof and R5 is hydrogen or acetyl; C2C16:0 includes a (saturated) sucrose ester that has the general structure shown in Figure 5 and wherein R3 is acetyl, R1 is valeryl or an isomer thereof, R2 is valeryl or an isomer thereof, R4 is hexanoyl or an isomer thereof and R5 is either a hydrogen atom or an acetyl moiety; or wherein the sucrose ester has the general structure shown in Figure 5 and wherein R3 is acetyl, R1 is propionyl or an isomer thereof, R2 is hexanoyl or an isomer thereof, R4 is hexanoyl or an isomer thereof and R5 is either a hydrogen atom or an acetyl moiety. C2C17:0 includes a (saturated) sucrose ester as shown in Figure 5 wherein R3 is acetyl, R1 is valeryl or an isomer thereof, R2 is hexanoyl, R4 is hexanoyl and R5 is either a hydrogen atom or an acetyl moiety. C2C18:0 includes a (saturated) sucrose ester as shown in Figure 5 wherein R3 is acetyl, R1 is hexanoyl, R2 is hexanoyl, R4 is hexanoyl and R5 is either a hydrogen atom or an acetyl moiety.

**Figure 2** shows the major sucrose ester isomers sorted by molecular mass and their occurrence in tobacco varieties producing or not producing sucrose esters containing beta-methylvaleryl esters. No distinction is made between 2 and 3-methylbutyryl or 3 and 4 methylvaleryl esters listed as C5 and C6 respectively. The following nomenclature is used to define the acyl composition. Referring for example to C2C4C4C5 this refers to the number of carbons in the acyl group at R3, R1, R2 and R4, respectively. C2 may be acetyl; C3 may be butyryl; C4 may be propionyl or an isomer thereof or isobutyryl; C5 may be valeryl (pentanoyl) or an isomer thereof or 2-methyl-butyryl or isovaleryl or isopentenoyl or pentenoyl; C6 may be hexanoyl or an isomer thereof or 2-methylvaleryl or beta-methylvaleryl or 4-methylvaleryl. An acetyl or hydrogen may be present at R5 (not shown), suitably hydrogen.

**Figure 3** illustrates the branched chain amino acid synthesis pathway and generation of the acyl chains that are esterified on sucrose. Boxed steps represent the core branched chain amino acid synthesis pathway where carbon chains are extended by one carbon per cycle of condensation isomerization and decarboxylation. 1: Acetolactate synthase [2.2.1.6]; 2: Acetohydroxyacid isomeroreductase [1.1.1.86]; 3: 3,4-Dihydroxyacid dehydratase [4.2.1.9]; 4: Branched-chain aminoacid aminotransferase [2.6.1.42]; 5: Branched-chain keto acid dehydrogenase complex (BCKD) ([1.2.4.4] [2.3.1.168] [1.8.1.4]); 6: Isopropylmalate synthase [2.3.3.13]; 7: Isopropylmalate isomerase [4.2.1.33]; 8: Isopropylmalate dehydrogenase [1.1.1.85]; 9: Threonine deaminase [4.3.1.19]. C4: isobutanoyl ester; iC5: 3-methyl-butyryl ester (isovaleryl ester); aiC5: 2-methyl-butyryl ester (anteiso-valeryl ester); iC6: 4-methyl-valeryl ester

(iso-caproyl ester); aiC6: 3-methyl-valeryl ester (anteiso-caproyl ester),  $\beta$ -methylvaleryl ester, BMV). Carboxylic acids are presented in neutral form although names reflect soluble forms. The proposed IPMS2 catalyzed steps are circled.

**Figure 4** shows the estimation of sucrose tetraester abundance in surface leaf exudate of transformant plants. Leaves are collected from one month old regenerating plants resulting from independent transformation events. Internal standard (sucrose octaacetate) is added at the rate of 2.5 ng/mg fresh weight. The area of the signal recorded for the ion specific to each sucrose ester is measured and plotted as ratio of internal standard area. The ratio to internal standard signal area allows the normalisation of signals on leaf fresh weight. The sucrose ester nomenclature presented here is described in Figure 2. **A:** composition of leaf exudates of the parental variety K326 (Flue cured, n=3) and the transformants expressing SEQ ID NO:1 under a trichome specific promoter; **B:** composition of leaf exudates of the parental variety K326 and the transformants expressing SEQ ID NO:1 under the control of a viral promoter; **C:** composition of leaf exudates of the parental variety Basma (Oriental, n=4) and the transformants expressing the RNAi construct of SEQ ID NO: 9 under a trichome specific promoter; and **D:** composition of leaf exudates of the parental variety Basma and transformants expressing the RNAi construct of SEQ ID NO: 9 under a viral promoter. C2C14:0 includes a (saturated) sucrose ester as shown in Figure 5 wherein R3=acetyl, R1=propionyl or an isomer thereof, R2= propionyl or an isomer thereof, R4=hexanoyl or an isomer thereof and R5 is hydrogen or acetyl; or a sucrose ester as shown in Figure 5 wherein R3=acetyl, R1= propionyl or an isomer thereof, R2=valeryl or an isomer thereof, R4= valeryl or an isomer thereof and R5 is hydrogen or acetyl; or a sucrose ester as shown in Figure 5 wherein R3=acetyl, R1=butyryl, R2=valeryl or an isomer thereof, R4=hexanoyl or an isomer thereof and R5 is hydrogen or acetyl. C2C15:0 includes a (saturated) sucrose ester as shown in Figure 5 wherein R3=acetyl, R1=propionyl or an isomer thereof, R2=valeryl or an isomer thereof, R4=hexanoyl or an isomer thereof and R5 is hydrogen or acetyl; or a sucrose ester as shown in Figure 5 wherein R3=acetyl, R1=valeryl or an isomer thereof, R2=valeryl or an isomer thereof, R4=valeryl or an isomer thereof and R5 is hydrogen or acetyl; or a sucrose ester as shown in Figure 5 wherein R3=acetyl, R1=butyryl, R2=hexanoyl or an isomer thereof, R4=hexanoyl or an isomer thereof and R5 is hydrogen or acetyl; C2C16:0 includes a (saturated) sucrose ester that has the general structure shown in Figure 5 and wherein R3 is acetyl, R1 is valeryl or an isomer thereof, R2 is valeryl or an isomer thereof, R4 is hexanoyl or an isomer thereof and R5 is either a hydrogen atom or an acetyl moiety; or wherein the sucrose ester has the general structure shown in Figure 5 and wherein R3 is acetyl, R1 is propionyl or an isomer thereof, R2 is hexanoyl or an isomer thereof, R4 is hexanoyl or an isomer thereof and R5 is either a hydrogen atom or an acetyl moiety. C2C17:0 includes a (saturated) sucrose ester as shown in Figure 5 wherein R3 is acetyl, R1 is valeryl or an isomer thereof, R2 is hexanoyl, R4 is hexanoyl and R5 is either a hydrogen atom or an acetyl moiety.

C2C18:0 includes a (saturated) sucrose ester as shown in Figure 5 wherein R3 is acetyl, R1 is hexanoyl, R2 is hexanoyl, R4 is hexanoyl and R5 is either a hydrogen atom or an acetyl moiety.

**Figure 5** illustrates the structure of the sucrose ester scaffold. R1 to R5 are either hydrogen atoms or acyl moieties linked by an ester bond.

- 5 **Figure 6** illustrates the acyl chains esterified in sucrose esters. R' represents the sucrose molecule. 1. acetyl, 2. butyryl, 3. propionyl, 4. isobutyryl, 5. valeryl (pentanoyl), 6. 2-methylbutyryl, 7. isovaleryl, 8. isopentenoyl, 9. pentenoyl, 10. hexanoyl or an isomer thereof 11. 2-methylvaleryl, 12. beta-methylvaleryl, 13. 4-methylvaleryl. Isomers and derivatives thereof are also contemplated. Thus, by way of example, the isomer may be an isomer of propionyl – such as isobutyryl. By way of further example, the isomer may be an isomer of valeryl (pentanoyl) – such as 2-methylbutyryl or isovaleryl; or an isomer of valeryl (pentanoyl) with a double bond – such as isopentenoyl or pentenoyl. By way of further example, the isomer may be an isomer of hexanoyl – such as 2-methylvaleryl, beta-methylvaleryl or 4-methylvaleryl. The isomer may or may not comprise one or more double bonds.

15

#### DEFINITIONS

The technical terms and expressions used within the scope of this application are generally to be given the meaning commonly applied to them in the pertinent art of plant and molecular biology. All of the following term definitions apply to the complete content of this application. The word "comprising" does not exclude other elements or steps, and the indefinite article "a" or "an" does not exclude a plurality. A single step may fulfil the functions of several features recited in the claims. The terms "about", "essentially" and "approximately" in the context of a given

20

numerate value or range refers to a value or range that is within 20%, within 10%, or within 5%, 4%, 3%, 2% or 1% of the given value or range.

25

The term "isolated" refers to any entity that is taken from its natural milieu, but the term does not connote any degree of purification.

A "vector" refers to a nucleic acid vehicle that comprises a combination of nucleic acid components for enabling the transport of nucleic acid, nucleic acid constructs and nucleic acid conjugates and the like. Suitable vectors include episomes capable of extra-chromosomal replication such as circular, double-stranded nucleic acid plasmids; linearized double-stranded nucleic acid plasmids; and other vectors of any origin.

30

An "expression vector" is a nucleic acid vehicle that comprises a combination of nucleic acid components for enabling the expression of nucleic acid, nucleic acid constructs and nucleic acid conjugates and the like. Suitable expression vectors include episomes capable of extra-chromosomal replication such as circular, double-stranded nucleic acid plasmids; linearized double-stranded nucleic acid plasmids; and other functionally equivalent expression vectors of any origin. An expression vector comprises at least a promoter positioned upstream and

35

operably-linked to a nucleic acid, nucleic acid constructs or nucleic acid conjugate, as defined below.

The term "construct" refers to a double-stranded, recombinant nucleic acid fragment comprising one or more polynucleotides. The construct comprises a "template strand" base-paired with a complementary "sense or coding strand." A given construct can be inserted into a vector in two possible orientations, either in the same (or sense) orientation or in the reverse (or anti-sense) orientation with respect to the orientation of a promoter positioned within a vector – such as an expression vector. An exemplary construct is set forth in SEQ ID NO: 9.

A "promoter" refers to a nucleic acid element/sequence, typically positioned upstream and operably-linked to a double-stranded nucleic acid fragment. Promoters can be derived entirely from regions proximate to a native gene of interest, or can be composed of different elements derived from different native promoters or synthetic nucleic acid segments. An exemplary promoter is set forth in SEQ ID NO: 8.

The terms "homology, identity or similarity" refer to the degree of sequence similarity between two polypeptides or between two nucleic acid molecules compared by sequence alignment. The degree of homology between two discrete nucleic acid sequences being compared is a function of the number of identical, or matching, nucleotides at comparable positions. The percent identity may be determined by visual inspection and mathematical calculation. Alternatively, the percent identity of two nucleic acid sequences may be determined by comparing sequence information using a computer program such as – ClustalW, ClustalX, BLAST, FASTA or Smith-Waterman.

The term "plant" refers to any plant at any stage of its life cycle or development, and its progenies. In one embodiment, the plant is a tobacco plant, which refers to a plant belonging to the genus *Nicotiana*. Preferred species, cultivars, hybrids and varieties of tobacco plant are described herein.

A "plant cell" refers to a structural and physiological unit of a plant. The plant cell may be in the form of a protoplast without a cell wall, an isolated single cell or a cultured cell, or as a part of higher organized unit such as but not limited to, plant tissue, a plant organ, or a whole plant.

The term "plant material" refers to any solid, liquid or gaseous composition, or a combination thereof, obtainable from a plant, including biomass, leaves, leaf lamina, midrib, stems, roots, flowers or flower parts, fruits, pollen, egg cells, zygotes, seeds, cuttings, secretions, extracts, cell or tissue cultures, or any other parts or products of a plant. In one embodiment, the plant material comprises or consists of biomass, seed or leaves. In another embodiment, the plant material comprises or consists of leaves.

The term "variety" refers to a population of plants that share constant characteristics which separate them from other plants of the same species. While possessing one or more distinctive

traits, a variety is further characterized by a very small overall variation between individuals within that variety. A variety is often sold commercially.

The term "line" or "breeding line" as used herein denotes a group of plants that are used during plant breeding. A line is distinguishable from a variety as it displays little variation between  
5 individuals for one or more traits of interest, although there may be some variation between individuals for other traits.

The term "modulating" may refer to reducing, inhibiting, increasing or otherwise affecting the activity of a polypeptide. The term may also refer to reducing, inhibiting, increasing or otherwise affecting the activity of a gene encoding a polypeptide which can include, but is not limited to,  
10 modulating transcriptional activity.

The term "reduce" or "reduced" as used herein, refers to a reduction of from about 10% to about 99%, or a reduction of at least 10%, at least 20%, at least 25%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 75%, at least 80%, at least 90%, at least 95%, at least 98%, at least 99%, or at least 100% or more of a quantity or an activity, such as but not  
15 limited to polypeptide activity, transcriptional activity, and/or protein expression.

The term "inhibit" or "inhibited" as used herein, refers to a reduction of from about 98% to about 100%, or a reduction of at least 98%, at least 99%, but particularly of 100%, of a quantity or an activity, such as but not limited to polypeptide activity, transcriptional activity, and/or protein expression.

The term "increase" or "increased" as used herein, refers to an increase of from about 10% to about 99%, or an increase of at least 10%, at least 20%, at least 25%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 75%, at least 80%, at least 90%, at least 95%, at least 98%, at least 99%, at least 100%, at least 150%, or at least 200% or more or more of a quantity or an activity, such as but not limited to polypeptide activity, transcriptional  
25 activity, and/or protein expression.

The term "control" in the context of a control plant or control plant cells means a plant or plant cells in which the expression or activity of isopropylmalate synthase has not been modified (for example, increased or reduced) and so it can provide a comparison with a plant in which the expression or activity of isopropylmalate synthase has been modified. The control plant may  
30 comprise an empty vector. The control plant may correspond to a wild-type plant.

#### DETAILED DESCRIPTION

In one aspect, there is provided an isolated polynucleotide comprising, consisting or consisting essentially of a polynucleotide sequence and having at least 60% sequence identity to any of  
35 the sequences described herein, including any of polynucleotides shown in the sequence listing. Suitably, the isolated polynucleotides comprise, consist or consist essentially of a sequence having at least 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 75%, 80%, 85%,

87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identity thereto.

In one aspect, there is provided an isolated polynucleotide comprising, consisting or essentially of a polynucleotide sequence encoding an isopropylmalate synthase and  
 5 least 60% sequence identity to SEQ ID NO:1. Suitably, the isolated polynucleotides consist or consist essentially of a sequence having at least 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 75%, 80%, 85%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity to SEQ ID NO:1.

In another aspect, there is provided an isolated polynucleotide comprising, consisting or essentially of a polynucleotide sequence encoding an isopropylmalate synthase  
 10 having at least 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 75%, 80%, 85%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity to SEQ ID NO:10.

In another aspect, there is provided an isolated polynucleotide comprising, consisting or essentially of a polynucleotide sequence encoding an isopropylmalate synthase  
 15 having at least 60%, 61%, 62%, 63%, 64%, 65%, 70%, 75%, 80%, 85%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity to SEQ ID NO:10.

In another aspect, there is provided an isolated polynucleotide comprising, consisting or essentially of a polynucleotide sequence encoding an isopropylmalate synthase  
 20 having at least 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 75%, 80%, 85%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity to SEQ ID NO:14.

In another aspect, there is provided an isolated polynucleotide comprising, consisting or essentially of a polynucleotide sequence having at least 60%, 61%, 62%, 63%, 64%, 65%, 70%, 75%, 80%, 85%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity to SEQ ID NO: 8. SEQ ID NO: 8 encodes a promoter.

The term "polynucleotide" refers to a polymer of nucleotides, which may be unmodified deoxyribonucleic acid (DNA) or ribonucleic acid (RNA). Accordingly, a polynucleotide  
 30 can be, without limitation, a genomic DNA, complementary DNA (cDNA), mRNA, or RNA or a fragment(s) thereof. Moreover, a polynucleotide can be single-stranded or double-stranded. A polynucleotide can be a mixture of single-stranded and double-stranded regions.

cloned fragments of cDNA, genomic DNA, oligonucleotides, or individual nucleotides, or a combination of the foregoing. Although the polynucleotide sequences described herein are shown as DNA sequences, the sequences include their corresponding RNA sequences, and their complementary (for example, completely complementary) DNA or RNA sequences, including the reverse complements thereof.

The term "isopropylmalate synthase polynucleotide" includes polynucleotides encoding isopropylmalate synthase (IMPS) from *Nicotiana tabacum* and includes polynucleotides comprising, consisting or consisting essentially of polynucleotides with substantial homology (that is, sequence similarity) or substantial identity to SEQ ID NO:1, SEQ ID NO: 10, SEQ ID NO: 12 or SEQ ID NO: 14; polynucleotide variants that have at least about 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 75%, 80%, 85%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity to the sequence of SEQ ID NO:1, SEQ ID NO: 10, SEQ ID NO: 12 or SEQ ID NO: 14; fragments of the polynucleotides including fragments of SEQ ID NO:1, SEQ ID NO: 10, SEQ ID NO: 12 or SEQ ID NO: 14; and fragments of SEQ ID NO:1, SEQ ID NO: 10, SEQ ID NO: 12 or SEQ ID NO: 14 with substantial homology (that is, sequence similarity) or substantial identity thereto that have at least about 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 75%, 80%, 85%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity to the corresponding fragments of SEQ ID NO:1, SEQ ID NO: 10, SEQ ID NO: 12 or SEQ ID NO: 14. The isopropylmalate synthase polynucleotide also includes sequences comprising a sufficient or substantial degree of identity or similarity to SEQ ID NO:1, SEQ ID NO: 10, SEQ ID NO: 12 or SEQ ID NO: 14 to encode a polypeptide that functions as an isopropylmalate synthase. In one embodiment, the term "isopropylmalate synthase polynucleotide" refers to a polymer of nucleotides which comprises, consists or consists essentially of a polynucleotide designated herein as SEQ ID NO:1, SEQ ID NO: 10, SEQ ID NO: 12 or SEQ ID NO: 14.

A polynucleotide as described herein will generally contain phosphodiester bonds, although in some cases, polynucleotide analogs are included that may have alternate backbones, comprising, for example, phosphoramidate, phosphorothioate, phosphorodithioate, or O-methylphosphoroamidite linkages; and peptide polynucleotide backbones and linkages. Other analog polynucleotides include those with positive backbones; non-ionic backbones, and non-ribose backbones. Modifications of the ribose-phosphate backbone may be done for a variety of reasons, for example, to increase the stability and half-life of such molecules in physiological environments or as probes on a biochip. Mixtures of naturally occurring polynucleotides and analogs can be made; alternatively, mixtures of different polynucleotide analogs, and mixtures of naturally occurring polynucleotides and analogs may be made.

A variety of polynucleotide analogs are known, including, for example, phosphoramidate, phosphorothioate, phosphorodithioate, O-methylphosphoroamidite linkages and peptide

polynucleotide backbones and linkages. Other analog polynucleotides include those with positive backbones, non-ionic backbones and non-ribose backbones. Polynucleotides containing one or more carbocyclic sugars are also included.

Other analogs include peptide nucleic acids which are peptide polynucleotide analogs. These backbones are substantially non-ionic under neutral conditions, in contrast to the highly charged phosphodiester backbone of naturally occurring polynucleotides. This may result in advantages. First, the peptide polynucleotide backbone may exhibit improved hybridization kinetics. Peptide nucleic acids have larger changes in the melting temperature for mismatched versus perfectly matched basepairs. DNA and RNA typically exhibit a 2-4 °C drop in melting temperature for an internal mismatch. With the non-ionic peptide polynucleotide backbone, the drop is closer to 7-9 °C. Similarly, due to their non-ionic nature, hybridization of the bases attached to these backbones is relatively insensitive to salt concentration. In addition, peptide nucleic acids may not be degraded or degraded to a lesser extent by cellular enzymes, and thus may be more stable.

Among the uses of the disclosed polynucleotides, and combinations of fragments thereof, is the use of fragments as probes in nucleic acid hybridisation assays or primers in nucleic acid amplification assays. Such fragments generally comprise at least about 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20 or more contiguous nucleotides of a DNA sequence. In other embodiments, a DNA fragment comprises at least about 10, 15, 20, 30, 40, 50 or 60 or more contiguous nucleotides of a DNA sequence. These techniques can also be used to identify *BMVSE* and *bmvse* plants since a functional isopropylmalate synthase gene is shown herein to be necessary for the formation of sucrose esters. The close by genetic markers and the alleles of the isopropylmalate synthase gene can be used as predictors of sucrose ester production by a plant. Such a test could be used to track plant varieties or to predict smoking quality or batch type homogeneity. Thus, in one aspect, there is also provided a method for detecting an isopropylmalate synthase polynucleotide as described herein comprising the use of probes or primers capable of specifically detecting or specifically amplifying said polynucleotide or specifically detecting and specifically amplifying said polynucleotide. In a further aspect, there is provided a method for identifying a plant that is capable of producing sucrose esters comprising the steps of: (a) providing a sample comprising nucleic acid from a plant of interest; and (b) determining the presence of an isopropylmalate synthase polynucleotide as described herein, wherein the presence of said polynucleotide in said plant is indicative that said plant is capable of producing the sucrose esters. Kits for detecting at least a portion of the isopropylmalate synthase polynucleotide are also provided which comprise one or more primers or probes for specifically detecting at least a portion of isopropylmalate synthase polynucleotide. The kit may comprise reagents for polynucleotide amplification - such as polymerase chain reaction (PCR) - or reagents for nucleic acid probe hybridization-detection technology - such as Southern Blots,

Northern Blots, in-situ hybridization, or microarray. The kit may comprise reagents for antibody binding-detection technology such as Western Blots, ELISAs, SELDI mass spectrometry or test strips. The kit may comprise reagents for DNA sequencing. The kit may comprise reagents and/or instructions for determining sucrose ester content. In some embodiments, a kit may  
5 comprise instructions for one or more of the methods described. The kits described may be useful for genetic identity determination, phylogenetic studies, genotyping, haplotyping, pedigree analysis or plant breeding particularly with co-dominant scoring. The methods and kits may be of use in a breeding approach relying on a cross between a BMVSE and a *bmvse* variety and a sequence-based selection of the plants in the offspring. Typically, plants showing  
10 the functional isopropylmalate synthase genetic region that was present in the BMVSE producing variety or plants showing the dysfunctional isopropylmalate synthase genetic region that was present in the *bmvse* non producing variety will be selected.

The basic parameters affecting the choice of hybridization conditions and guidance for devising suitable conditions are described by Sambrook, J., E. F. Fritsch, and T. Maniatis (1989,  
15 Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.). Using knowledge of the genetic code in combination with the amino acid sequences described herein, sets of degenerate oligonucleotides can be prepared. Such oligonucleotides are useful as primers, for example, in PCR, whereby DNA fragments are isolated and amplified. In certain embodiments, degenerate primers can be used as probes for  
20 genetic libraries. Such libraries would include but are not limited to cDNA libraries, genomic libraries, and even electronic express sequence tag or DNA libraries. Homologous sequences identified by this method would then be used as probes to identify homologues of the sequences identified herein.

Also of potential use are polynucleotides and oligonucleotides (for example, primers or probes)  
25 that hybridize under reduced stringency conditions, typically moderately stringent conditions, and commonly highly stringent conditions to the polynucleotide(s) as described herein. The basic parameters affecting the choice of hybridization conditions and guidance for devising suitable conditions are set forth by Sambrook, J., E. F. Fritsch, and T. Maniatis (1989, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.)  
30 and can be readily determined by those having ordinary skill in the art based on, for example, the length or base composition of the polynucleotide.

One way of achieving moderately stringent conditions involves the use of a prewashing solution containing 5x Standard Sodium Citrate, 0.5% Sodium Dodecyl Sulphate, 1.0 mM Ethylenediaminetetraacetic acid (pH 8.0), hybridization buffer of about 50% formamide, 6x  
35 Standard Sodium Citrate, and a hybridization temperature of about 55 °C (or other similar hybridization solutions, such as one containing about 50% formamide, with a hybridization temperature of about 42 °C), and washing conditions of about 60 °C, in 0.5x Standard Sodium

Citrate, 0.1% Sodium Dodecyl Sulphate. Generally, highly stringent conditions are defined as hybridization conditions as above, but with washing at approximately 68 °C, 0.2x Standard Sodium Citrate, 0.1% Sodium Dodecyl Sulphate. SSPE (1x SSPE is 0.15M sodium chloride, 10 mM sodium phosphate, and 1.25 mM Ethylenediaminetetraacetic acid, pH 7.4) can be substituted for Standard Sodium Citrate (1x Standard Sodium Citrate is 0.15M sodium chloride and 15 mM sodium citrate) in the hybridization and wash buffers; washes are performed for 15 minutes after hybridization is complete. It should be understood that the wash temperature and wash salt concentration can be adjusted as necessary to achieve a desired degree of stringency by applying the basic principles that govern hybridization reactions and duplex stability, as known to those skilled in the art and described further below (see, for example, Sambrook, J., E. F. Fritsch, and T. Maniatis (1989, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.). When hybridizing a polynucleotide to a target polynucleotide of unknown sequence, the hybrid length is assumed to be that of the hybridizing polynucleotide. When polynucleotides of known sequence are hybridized, the hybrid length can be determined by aligning the sequences of the polynucleotides and identifying the region or regions of optimal sequence complementarity. The hybridization temperature for hybrids anticipated to be less than 50 base pairs in length should be 5 to 10 °C less than the melting temperature of the hybrid, where melting temperature is determined according to the following equations. For hybrids less than 18 base pairs in length, melting temperature (°C)=2(number of A+T bases)+4(number of G+C bases). For hybrids above 18 base pairs in length, melting temperature (°C)=81.5+16.6(log<sub>10</sub> [Na<sup>+</sup>])+0.41(% G+C)-(600/N), where N is the number of bases in the hybrid, and [Na<sup>+</sup>] is the concentration of sodium ions in the hybridization buffer ([Na<sup>+</sup>] for 1x Standard Sodium Citrate=0.165M). Typically, each such hybridizing polynucleotide has a length that is at least 25% (commonly at least 50%, 60%, or 70%, and most commonly at least 80%) of the length of a polynucleotide to which it hybridizes, and has at least 60% sequence identity (for example, at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100%) with a polynucleotide to which it hybridizes.

As will be understood by the person skilled in the art, a linear DNA has two possible orientations: the 5'-to-3' direction and the 3'-to-5' direction. For example, if a reference sequence is positioned in the 5'-to-3' direction, and if a second sequence is positioned in the 5'-to-3' direction within the same polynucleotide molecule/strand, then the reference sequence and the second sequence are orientated in the same direction, or have the same orientation. Typically, a promoter sequence and a gene of interest under the regulation of the given promoter are positioned in the same orientation. However, with respect to the reference sequence positioned in the 5'-to-3' direction, if a second sequence is positioned in the 3'-to-5' direction within the same polynucleotide molecule/strand, then the reference sequence and the second sequence are orientated in anti-sense direction, or have anti-sense orientation. Two

sequences having anti-sense orientations with respect to each other can be alternatively described as having the same orientation, if the reference sequence (5'-to-3' direction) and the reverse complementary sequence of the reference sequence (reference sequence positioned in the 5'-to-3') are positioned within the same polynucleotide molecule/strand. The sequences set forth herein are shown in the 5'-to-3' direction.

Recombinant constructs provided herein can be used to transform plants or plant cells in order to modulate isopropylmalate synthase protein expression levels. A recombinant polynucleotide construct can comprise a polynucleotide encoding an isopropylmalate synthase polynucleotide as described herein, operably linked to a regulatory region suitable for expressing the isopropylmalate synthase polypeptide in the plant or cell. Thus, a polynucleotide can comprise a coding sequence that encodes the isopropylmalate synthase polypeptide as described herein. The isopropylmalate synthase polypeptide encoded by a recombinant polynucleotide can be a native isopropylmalate synthase polypeptide, or can be heterologous to the cell. In some cases, the recombinant construct contains a polynucleotide that reduces or inhibits expression of a isopropylmalate synthase-modulating polypeptide, operably linked to a regulatory region. Examples of suitable regulatory regions are described herein.

Vectors containing recombinant polynucleotide constructs such as those described herein also are provided. Suitable vector backbones include, for example, those routinely used in the art such as plasmids, viruses, artificial chromosomes, BACs, YACs, or PACs. Suitable expression vectors include, without limitation, plasmids and viral vectors derived from, for example, bacteriophage, baculoviruses, and retroviruses. Numerous vectors and expression systems are commercially available.

The vectors can also include, for example, origins of replication, scaffold attachment regions or markers. A marker gene can confer a selectable phenotype on a plant cell. For example, a marker can confer biocide resistance, such as resistance to an antibiotic (for example, kanamycin, G418, bleomycin, or hygromycin), or an herbicide (for example, glyphosate, chloresulfuron or phosphinothricin). In addition, an expression vector may include a tag sequence designed to facilitate manipulation or detection (for example, purification or localization) of the expressed polypeptide. Tag sequences, such as luciferase, beta-glucuronidase (GUS), green fluorescent protein (GFP), glutathione S-transferase (GST), polyhistidine, c-myc or hemagglutinin sequences typically are expressed as a fusion with the encoded polypeptide. Such tags can be inserted anywhere within the polypeptide, including at either the carboxyl or amino terminus.

A plant or plant cell can be transformed by having the recombinant polynucleotide integrated into its genome to become stably transformed. Stably transformed cells typically retain the introduced polynucleotide with each cell division. A plant or plant cell may also be transiently transformed such that the recombinant polynucleotide is not integrated into its genome.

Transiently transformed cells typically lose all or some portion of the introduced recombinant polynucleotide with each cell division such that the introduced recombinant polynucleotide cannot be detected in daughter cells after a sufficient number of cell divisions.

5 A number of methods are available in the art for transforming a plant cell which are all encompassed herein, including biolistics, gene gun techniques, Agrobacterium-mediated transformation, viral vector-mediated transformation and electroporation. The Agrobacterium system for integration of foreign nucleic acid into plant chromosomes has been extensively studied, modified, and exploited for plant genetic engineering. Naked recombinant nucleic acid molecules comprising nucleic acid sequences corresponding to the subject purified tobacco  
10 protein operably linked, in the sense or antisense orientation, to regulatory sequences are joined to appropriate T-DNA sequences by conventional methods. These are introduced into tobacco protoplasts by polyethylene glycol techniques or by electroporation techniques, both of which are standard. Alternatively, such vectors comprising recombinant nucleic acid molecules encoding the subject purified protein are introduced into live Agrobacterium cells, which then  
15 transfer the nucleic acid into the plant cells. Transformation by naked DNA without accompanying T-DNA vector sequences can be accomplished via fusion of tobacco protoplasts with nucleic acid -containing liposomes or via electroporation. Naked DNA unaccompanied by T-DNA vector sequences can also be used to transform tobacco cells via inert, high velocity microprojectiles.

20 If a cell or cultured tissue is used as the recipient tissue for transformation, plants can be regenerated from transformed cultures if desired, by techniques known to those skilled in the art.

The choice of regulatory regions to be included in a recombinant construct depends upon several factors, including, but not limited to, efficiency, selectability, inducibility, desired  
25 expression level, and cell- or tissue-preferential expression. It is a routine matter for one of skill in the art to modulate the expression of a coding sequence by appropriately selecting and positioning regulatory regions relative to the coding sequence. Transcription of a polynucleotide can be modulated in a similar manner. Some suitable regulatory regions initiate transcription only, or predominantly, in certain cell types. Methods for identifying and characterizing  
30 regulatory regions in plant genomic DNA are known in the art.

Suitable promoters include tissue-specific promoters recognized by tissue-specific factors present in different tissues or cell types (for example, root-specific promoters, shoot-specific promoters, xylem-specific promoters), or present during different developmental stages, or present in response to different environmental conditions. Suitable promoters include  
35 constitutive promoters that can be activated in most cell types without requiring specific inducers. Examples of suitable promoters for controlling isopropylmalate synthase RNAi polypeptide production include the cauliflower mosaic virus 35S (CaMV/35S), SSU, OCS, lib4,

usp, STLS1, B33, nos or ubiquitin- or phaseolin-promoters. Persons skilled in the art are capable of generating multiple variations of recombinant promoters.

Tissue-specific promoters are transcriptional control elements that are only active in particular cells or tissues at specific times during plant development, such as in vegetative tissues or reproductive tissues. Tissue-specific expression can be advantageous, for example, when the expression of polynucleotides in certain tissues is preferred. Examples of tissue-specific promoters under developmental control include promoters that can initiate transcription only (or primarily only) in certain tissues, such as vegetative tissues, for example, roots or leaves, or reproductive tissues, such as fruit, ovules, seeds, pollen, pistils, flowers, or any embryonic tissue. Reproductive tissue-specific promoters may be, for example, anther-specific, ovule-specific, embryo-specific, endosperm-specific, integument-specific, seed and seed coat-specific, pollen-specific, petal-specific, sepal-specific, or combinations thereof.

Suitable leaf-specific promoters include pyruvate, orthophosphate dikinase (PPDK) promoter from C4 plant (maize), cab-m1Ca+2 promoter from maize, the *Arabidopsis thaliana* myb-related gene promoter (Atmyb5), the ribulose biphosphate carboxylase (RBCS) promoters (for example, the tomato RBCS 1, RBCS2 and RBCS3A genes expressed in leaves and light-grown seedlings, RBCS1 and RBCS2 expressed in developing tomato fruits or ribulose biphosphate carboxylase promoter expressed almost exclusively in mesophyll cells in leaf blades and leaf sheaths at high levels).

Suitable senescence-specific promoters include a tomato promoter active during fruit ripening, senescence and abscission of leaves, a maize promoter of gene encoding a cysteine protease. Suitable anther-specific promoters can be used. Suitable root-preferred promoters known to persons skilled in the art may be selected. Suitable seed-preferred promoters include both seed-specific promoters (those promoters active during seed development such as promoters of seed storage proteins) and seed-germinating promoters (those promoters active during seed germination). Such seed-preferred promoters include, but are not limited to, Cim1 (cytokinin-induced message); cZ19B1 (maize 19 kDa zein); milps (myo-inositol-1-phosphate synthase); mZE40-2, also known as Zm-40; nuc1c; and celA (cellulose synthase). Gamma-zein is an endosperm-specific promoter. Glob-1 is an embryo-specific promoter. For dicots, seed-specific promoters include, but are not limited to, bean beta-phaseolin, napin,  $\beta$ -conglycinin, soybean lectin, cruciferin, and the like. For monocots, seed-specific promoters include, but are not limited to, a maize 15 kDa zein promoter, a 22 kDa zein promoter, a 27 kDa zein promoter, a gamma-zein promoter, a 27 kDa  $\gamma$ -zein promoter (such as gzw64A promoter, see Genbank Accession number S78780), a waxy promoter, a shrunken 1 promoter, a shrunken 2 promoter, a globulin 1 promoter (see Genbank Accession number L22344), an ltp2 promoter, cim1 promoter, maize end1 and end2 promoters, nuc1 promoter, Zm40 promoter, eep1 and eep2; lec1, thioredoxin H promoter, mlip15 promoter, PCNA2 promoter; and the shrunken-2 promoter.

Examples of inducible promoters include promoters responsive to pathogen attack, conditions, elevated temperature, light, drought, cold temperature, or high salt concentration. Pathogen-inducible promoters include those from pathogenesis-related proteins (PR) which are induced following infection by a pathogen (for example, PR proteins, SA-inducible beta-1,3-glucanase, chitinase).

In addition to plant promoters, other suitable promoters may be derived from bacteria (for example, the octopine synthase promoter, the nopaline synthase promoter and other promoters derived from Ti plasmids), or may be derived from viral promoters (for example, 35S RNA promoters of cauliflower mosaic virus (CaMV), constitutive promoters of tobacco etch virus, cauliflower mosaic virus (CaMV) 19S and 35S promoters, or figwort mosaic promoter).

The term "isopropylmalate synthase polypeptide" refers to a polypeptide isopropylmalate synthase (IPMS) from *Nicotiana tabacum* and includes polypeptides comprising, consisting or consisting essentially of an amino acid sequence encoded by a polynucleotide variant with at least about 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 75%, 80%, 85%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity to SEQ ID NO:2, SEQ ID NO:11, SEQ ID NO:13 or SEQ ID NO:15; fragments of the isopropylmalate synthase polypeptide(s); and fragments of SEQ ID NO:11, SEQ ID NO:13 or SEQ ID NO:15 that have at least about 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 75%, 80%, 85%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 100% sequence identity to the corresponding fragments of SEQ ID NO:2, SEQ ID NO:11, SEQ ID NO:13 and SEQ ID NO:15. The isopropylmalate synthase polypeptide also includes sequences comprising a sufficient or substantial degree of identity or similarity to SEQ ID NO:11, SEQ ID NO:13 and SEQ ID NO:15 to function as an isopropylmalate synthase. The fragments of the isopropylmalate synthase polypeptide typically retain isopropylmalate synthase activity. Isopropylmalate synthase polypeptides also include mutants prepared by introducing any type of alterations (for example, insertions, deletions, or substitution of amino acids; changes in glycosylation states; changes that affect refolding or isomerization of three-dimensional structures, or self-association states), which can be deliberately engineered or isolated naturally provided that they still function as an isopropylmalate synthase. Isopropylmalate synthase polypeptides may be in linear form or cyclized using known methods. The term "isopropylmalate synthase polypeptide" can also refer to a polypeptide encoded by a nucleic acid sequence.

In another aspect, there is provided an isolated polypeptide comprising, consisting or consisting essentially of a polypeptide sequence having at least 60% sequence identity to any of the sequences described herein, including any of polypeptides shown in the sequence listing. Suitably, the isolated polypeptides comprise, consist or consist essentially of a sequence having

5 at least 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 75%, 80%, 85%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity thereto.

Polypeptides include variants produced by introducing any type of alterations (for example, insertions, deletions, or substitutions of amino acids; changes in glycosylation states; changes that affect refolding or isomerizations, three-dimensional structures, or self-association states),

10 which can be deliberately engineered or isolated naturally. The variant may have alterations which produce a silent change and result in a functionally equivalent protein. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity and the amphipathic nature of the residues as long as the

15 secondary binding activity of the substance is retained. For example, negatively charged amino acids include aspartic acid and glutamic acid; positively charged amino acids include lysine and arginine; and amino acids with uncharged polar head groups having similar hydrophilicity values include leucine, isoleucine, valine, glycine, alanine, asparagine, glutamine, serine, threonine, phenylalanine, and tyrosine. Conservative substitutions may be made, for example according

20 to the Table below. Amino acids in the same block in the second column and preferably in the same line in the third column may be substituted for each other:

|           |                   |                            |
|-----------|-------------------|----------------------------|
| ALIPHATIC | Non-polar         | Gly Ala Pro<br>Ile Leu Val |
|           | Polar - uncharged | Cys Ser Thr Met<br>Asn Gly |
|           | Polar - charged   | Asp Glu<br>Lys Arg         |
| AROMATIC  |                   | His Phe TrpTyr             |

The polypeptide may be a mature protein or an immature protein or a protein derived from an

25 immature protein. Polypeptides may be in linear form or cyclized using known methods. Polypeptides typically comprise at least 10, at least 20, at least 30, or at least 40 contiguous amino acids.

In another aspect, there is provided an isolated isopropylmalate synthase polypeptide comprising, consisting or consisting essentially of a sequence encoding an isopropylmalate

30 synthase and having at least about 60%, 65%, 70%, 75%, 80%, 85%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity to SEQ ID NO:2.

In another aspect, there is provided an isolated isopropylmalate synthase polypeptide comprising, consisting or consisting essentially of a sequence encoding an isopropylmalate synthase and having at least about 60%, 65%, 70%, 75%, 80%, 85%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95% 96%, 97%, 98%, 99% or 100% sequence identity to SEQ ID NO:11.

- 5 In another aspect, there is provided an isolated isopropylmalate synthase polypeptide comprising, consisting or consisting essentially of a sequence encoding an isopropylmalate synthase and having at least about 60%, 65%, 70%, 75%, 80%, 85%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95% 96%, 97%, 98%, 99% or 100% sequence identity to SEQ ID NO:13.
- In another aspect, there is provided an isolated isopropylmalate synthase polypeptide
- 10 comprising, consisting or consisting essentially of a sequence encoding an isopropylmalate synthase and having at least about 60%, 65%, 70%, 75%, 80%, 85%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95% 96%, 97%, 98%, 99% or 100% sequence identity to SEQ ID NO:15.
- Fragments of the polypeptide sequences are also disclosed herein, suitably, such fragments retain the activity of the full length sequence.

- 15 Mutant polypeptides can be used to create mutant plants, non-naturally occurring plants or transgenic plants comprising the mutant polypeptide. Suitably, the mutant polypeptide retains the activity of the unmutated polypeptide. The activity of the mutant polypeptide may be higher, lower or about the same as the unmutated polypeptide.

- A polypeptide may be prepared by culturing transformed or recombinant host cells under culture
- 20 conditions suitable to express a polypeptide. The resulting expressed polypeptide may then be purified from such culture using known purification processes. The purification of the polypeptide may include an affinity column containing agents which will bind to the polypeptide; one or more column steps over such affinity resins; one or more steps involving hydrophobic interaction chromatography; or immunoaffinity chromatography. Alternatively, the polypeptide
- 25 may also be expressed in a form that will facilitate purification. For example, it may be expressed as a fusion polypeptide, such as those of maltose binding polypeptide, glutathione-S-transferase or thioredoxin. Kits for expression and purification of fusion polypeptides are commercially available. The polypeptide may be tagged with an epitope and subsequently purified by using a specific antibody directed to such epitope. One or more liquid
- 30 chromatography steps – such as reverse-phase high performance liquid chromatography can be employed to further purify the polypeptide. Some or all of the foregoing purification steps, in various combinations, can be employed to provide a substantially homogeneous recombinant polypeptide. The polypeptide thus purified may be substantially free of other polypeptides and is defined herein as an "substantially purified polypeptide"; such purified polypeptides include
- 35 polypeptides, fragments, variants, and the like. Expression, isolation, and purification of the polypeptides and fragments can be accomplished by any suitable technique, including but not limited to the methods described herein.

It is also possible to utilise an affinity column such as a monoclonal antibody generated against polypeptides, to affinity-purify expressed polypeptides. These polypeptides can be removed from an affinity column using conventional techniques, for example, in a high salt elution buffer and then dialyzed into a lower salt buffer for use or by changing pH or other components depending on the affinity matrix utilized, or be competitively removed using the naturally occurring substrate of the affinity moiety.

A polypeptide may also be produced by known conventional chemical synthesis. Methods for constructing the polypeptides or fragments thereof by synthetic means are known to those skilled in the art. The synthetically-constructed polypeptide sequences, by virtue of sharing primary, secondary or tertiary structural or conformational characteristics with a native polypeptides may possess biological properties in common therewith, including biological activity.

The term 'non-naturally occurring' as used herein describes an entity (for example, a polynucleotide, a genetic mutation, a polypeptide, a plant, a plant cell and plant material) that is not formed by nature or that does not exist in nature. Such non-naturally occurring entities or artificial entities may be made, synthesized, initiated, modified, intervened, or manipulated by methods described herein or that are known in the art. Thus, by way of example, a non-naturally occurring plant, a non-naturally occurring plant cell or non-naturally occurring plant material may be made using traditional plant breeding techniques - such as backcrossing - or by genetic manipulation technologies - such as antisense RNA, interfering RNA, meganuclease and the like. By way of further example, a non-naturally occurring plant, a non-naturally occurring plant cell or non-naturally occurring plant material may be made by introgression of or by transferring one or more genetic mutations (for example one or more polymorphisms) from a first plant or plant cell into a second plant or plant cell (which may itself be naturally occurring), such that the resulting plant, plant cell or plant material or the progeny thereof comprises a genetic constitution (for example, a genome, a chromosome or a segment thereof) that is not formed by nature or that does not exist in nature. The resulting plant, plant cell or plant material is thus artificial or non-naturally occurring. Accordingly, an artificial or non-naturally occurring plant or plant cell may be made by modifying a genetic sequence in a first artificial or naturally occurring plant or plant cell, resulting in a modified genetic sequence (i) that is artificial; or (ii) that occurs naturally in a second plant or plant cell, provided that the second plant or plant cell comprises a different genetic background from the first plant or plant cell. Differences in genetic background of a plant or plant cell can be detected by phenotypic differences or by molecular biology techniques known in the art - such as nucleic acid sequencing, presence or absence of genetic markers (for example, microsatellite RNA markers).

A number of polynucleotide based methods can be used to increase gene expression in plants. By way of example, a construct, vector or expression vector that is compatible with the plant to

be transformed can be prepared which comprises the gene of interest together with an upstream promoter that is capable of overexpressing the gene in the plant. Exemplary promoters are described herein. Following transformation and when grown under suitable conditions, the promoter can drive expression in order to modulate (for example, increase) the levels of this enzyme in the plant, or in a specific tissue thereof. In one exemplary embodiment, a vector carrying isopropylmalate synthase is generated to overexpress the gene in a plant. The vector carries a suitable promoter upstream of the isopropylmalate synthase gene driving its expression in all tissues of the plant. The vector may also carry an antibiotic resistance gene in order to confer selection of the transformed calli and cell lines. In one embodiment, the promoter promotes transcriptional activity in the trichomes of a plant. In one embodiment, the promoter has transcriptional activity in the trichomes of a plant. In another embodiment, the promoter has transcriptional activity specifically in the trichomes of a plant with no or substantially no activity in any other parts of a plant. Suitably, the promoter comprises, consists or consists essentially of the sequence set forth in SEQ ID NO: 8 or a trichome promoter or a natural isopropylmalate synthase promoter. The natural isopropylmalate synthase promoter will be located at the 5' end of the isopropylmalate synthase gene. In SEQ ID NO:6 for example, the endogenous isopropylmalate synthase promoter will be positioned before nucleotide 2826. Accordingly, the isopropylmalate synthase gene may be over expressed in the trichome of a plant. Surprisingly, the inventors have discovered that constitutive expression does not result in as high an accumulation of sucrose esters as compared to the levels achieved using trichome specific expression (see Figure 4 and the description thereof).

According to one embodiment, the expression or activity of isopropylmalate synthase is increased in a plant already capable of producing beta-methylvaleryl containing sucrose esters thereby resulting in the increased accumulation of isopropylmalate synthase therein and thus increased beta-methylvaleryl containing sucrose ester production. According to another embodiment, the expression or activity of isopropylmalate synthase is increased in a plant that is not already capable of producing beta-methylvaleryl containing sucrose esters thereby resulting in the accumulation of isopropylmalate synthase therein and thus beta-methylvaleryl containing sucrose ester production.

A number of polynucleotide based methods, including antisense RNA, ribozyme directed RNA cleavage, post-transcriptional gene silencing, for example, RNA interference and transcriptional gene silencing, can be used to modulate gene expression in plants. Whilst such approaches are classically used to reduce or inhibit gene expression, they can also be used to increase gene expression by increasing cellular productivity and the quality of proteins that are expressed. Such approaches can include the silencing of apoptosis-associated gene expression, protein glycosylation-associated gene expression, lactate dehydrogenase involved in cellular metabolism, and dihydrofolate reductase used for gene amplification. Other

approaches can be extended to reducing or inhibiting multiple targets involved in different cellular pathways for changing gene regulation in plant cells.

In some embodiments, a complement of the full-length polynucleotide or a fragment thereof can be used. Typically, a fragment is at least 10 contiguous nucleotides, for example, at least 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 30, 35, 40, 50, 80, 100, 200, 500 contiguous nucleotides or more. In one embodiment, the polynucleotide fragment(s) is able to decrease expression or activity and may encode, for example, an antisense nucleic acid, a ribozyme, an RNA that effects spliceosome-mediated trans-splicing, an interfering RNA, a guide RNA, or other non-translated RNA and the like as described herein. In another embodiment, the polynucleotide fragment(s) encodes an interfering RNA.

Compositions that can reduce the expression or the activity of isopropylmalate synthase include, but are not limited to, sequence-specific polynucleotides that can interfere with the transcription of one or more endogenous isopropylmalate synthase gene(s); sequence-specific polynucleotides that can interfere with the translation of isopropylmalate synthase RNA transcripts (for example, double-stranded RNAs, siRNAs, ribozymes); sequence-specific polypeptides that can interfere with the stability of isopropylmalate synthase proteins; sequence-specific polynucleotides that can interfere with the enzymatic activity of isopropylmalate synthase protein or the binding activity of isopropylmalate synthase protein with respect to substrates or regulatory proteins; antibodies that exhibit specificity for isopropylmalate synthase protein; small molecule compounds that can interfere with the stability of isopropylmalate synthase protein or the enzymatic activity of isopropylmalate synthase protein or the binding activity of isopropylmalate synthase protein; zinc finger proteins that bind isopropylmalate synthase polynucleotide; and meganucleases that have activity towards isopropylmalate synthase polynucleotide.

Antisense technology is one well-known method that can be used to modulate (for example, reduce or inhibit) the expression of a polypeptide. A polynucleotide to be repressed is cloned and operably linked to a regulatory region and a transcription termination sequence so that the antisense strand of RNA is transcribed. The recombinant construct is then transformed into plants and the antisense strand of RNA is produced. The polynucleotide need not be the entire sequence of the gene to be repressed, but typically will be substantially complementary to at least a portion of the sense strand of the gene to be repressed.

A polynucleotide may be transcribed into a ribozyme, or catalytic RNA, that affects expression of an mRNA. Ribozymes can be designed to specifically pair with virtually any target RNA and cleave the phosphodiester backbone at a specific location, thereby functionally inactivating the target RNA. Heterologous polynucleotides can encode ribozymes designed to cleave particular mRNA transcripts, thus preventing expression of a polypeptide. Hammerhead ribozymes are useful for destroying particular mRNAs, although various ribozymes that cleave mRNA at site-

specific recognition sequences can be used. Hammerhead ribozymes cleave mRNAs at locations dictated by flanking regions that form complementary base pairs with the target mRNA. The sole requirement is that the target RNA contains a 5'-UG-3' nucleotide sequence. The construction and production of hammerhead ribozymes is known in the art. Hammerhead  
5 ribozyme sequences can be embedded in a stable RNA such as a transfer RNA to increase cleavage efficiency *in vivo*.

In one embodiment, the sequence-specific polynucleotides that can interfere with the translation of RNA transcripts is interfering RNA. RNA interference or RNA silencing is an evolutionarily conserved process by which specific mRNAs can be targeted for enzymatic degradation. A  
10 double-stranded RNA must be introduced or produced by a cell (for example, double-stranded RNA virus, or interfering RNA polynucleotides) to initiate the interfering RNA pathway. The double-stranded RNA can be converted into multiple small interfering RNA duplexes of 21-23 bp length by RNases III, which are double-stranded RNA-specific endonucleases. The small interfering RNAs can be subsequently recognized by RNA-induced silencing complexes that  
15 promote the unwinding of small interfering RNA through an ATP-dependent process. The unwound antisense strand of the small interfering RNA guides the activated RNA-induced silencing complexes to the targeted mRNA comprising a sequence complementary to the small interfering RNA anti-sense strand. The targeted mRNA and the anti-sense strand can form an A-form helix, and the major groove of the A-form helix can be recognized by the activated RNA-  
20 induced silencing complexes. The target mRNA can be cleaved by activated RNA-induced silencing complexes at a single site defined by the binding site of the 5'-end of the small interfering RNA strand. The activated RNA-induced silencing complexes can be recycled to catalyze another cleavage event.

Interfering RNA expression vectors may comprise interfering RNA constructs encoding  
25 interfering RNA polynucleotides that exhibit RNA interference activity by reducing the expression level of mRNAs, pre-mRNAs, or related RNA variants. The expression vectors may comprise a promoter positioned upstream and operably-linked to a interfering RNA construct, as further described herein. Interfering RNA expression vectors may comprise a suitable minimal core promoter, an interfering RNA construct of interest, an upstream (5') regulatory region, a  
30 downstream (3') regulatory region, including transcription termination and polyadenylation signals, and other sequences known to persons skilled in the art, such as various selection markers.

The polynucleotides can be produced in various forms, including as double stranded structures (that is, a double-stranded RNA molecule comprising an antisense strand and a complementary  
35 sense strand), double-stranded hairpin-like structures, or single-stranded structures (that is, a single-stranded RNA molecule comprising just an antisense strand). The structures may comprise a duplex, asymmetric duplex, hairpin or asymmetric hairpin secondary structure,

having self-complementary sense and antisense strands. The double stranded interfering RNA can be enzymatically converted to double-stranded small interfering RNAs. One of the strands of the small interfering RNA duplex can anneal to a complementary sequence within the target mRNA and related RNA variants. The small interfering RNA/mRNA duplexes are recognized by

5 RNA-induced silencing complexes that can cleave RNAs at multiple sites in a sequence-dependent manner, resulting in the degradation of the target mRNA and related RNA variants. The double-stranded RNA molecules may include small interfering RNA molecules assembled from a single oligonucleotide in a stem-loop structure, wherein self-complementary sense and antisense regions of the small interfering RNA molecule are linked by means of a polynucleotide

10 based or non-polynucleotide-based linker(s), as well as circular single-stranded RNA having two or more loop structures and a stem comprising self-complementary sense and antisense strands, wherein the circular RNA can be processed either *in vivo* or *in vitro* to generate an active small interfering RNA molecule capable of mediating interfering RNA.

The use of small hairpin RNA molecules is also contemplated herein and comprise a specific

15 antisense sequence in addition to the reverse complement (sense) sequence, typically separated by a spacer or loop sequence. Cleavage of the spacer or loop provides a single-stranded RNA molecule and its reverse complement, such that they may anneal to form a double-stranded RNA molecule (optionally with additional processing steps that may result in addition or removal of one, two, three or more nucleotides from the 3' end or the 5' end of either

20 or both strands). The spacer can be of a sufficient length to permit the antisense and sense sequences to anneal and form a double-stranded structure (or stem) prior to cleavage of the spacer (and, optionally, subsequent processing steps that may result in addition or removal of one, two, three, four, or more nucleotides from the 3' end or the 5' end of either or both strands). The spacer sequence is typically an unrelated nucleotide sequence that is situated between two

25 complementary nucleotide sequence regions which, when annealed into a double-stranded polynucleotide, comprise a small hairpin RNA. The spacer sequence generally comprises between about 3 and about 100 nucleotides.

Any RNA polynucleotide of interest can be produced by selecting a suitable sequence composition, loop size, and stem length for producing the hairpin duplex. A suitable range for

30 designing stem lengths of a hairpin duplex, includes stem lengths of at least about 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20 nucleotides – such as about 14-30 nucleotides, about 30-50 nucleotides, about 50-100 nucleotides, about 100-150 nucleotides, about 150-200 nucleotides, about 200-300 nucleotides, about 300-400 nucleotides, about 400-500 nucleotides, about 500-600 nucleotides, and about 600-700 nucleotides. A suitable range for designing loop lengths of

35 a hairpin duplex, includes loop lengths of about 4-25 nucleotides, about 25-50 nucleotides, or longer if the stem length of the hair duplex is substantial. A double-stranded RNA or single-stranded RNA molecule is between about 15 and about 40 nucleotides in length. The small

interfering RNA molecule may be a double-stranded RNA or single-stranded RNA molecule between about 15 and about 35 nucleotides in length. The small interfering RNA molecule may be a double-stranded RNA or single-stranded RNA molecule between about 17 and about 30 nucleotides in length. The small interfering RNA molecule may be a double-stranded RNA or single-stranded RNA molecule between about 19 and about 25 nucleotides in length. The small interfering RNA molecule is a double-stranded RNA or single-stranded RNA molecule between about 21 to about 23 nucleotides in length. Hairpin structures with duplexed regions longer than 21 nucleotides may promote effective small interfering RNA-directed silencing, regardless of loop sequence and length.

The target mRNA sequence may be between about 14 to about 50 nucleotides in length. The target mRNA can, therefore, be scanned for regions between about 14 and about 50 nucleotides in length that preferably meet one or more of the following criteria for a target sequence: an A+T/G+C ratio of between about 2:1 and about 1:2; an AA dinucleotide or a CA dinucleotide at the 5' end of the target sequence; a sequence of at least 10 consecutive nucleotides unique to the target mRNA (that is, the sequence is not present in other mRNA sequences from the same plant); and no "runs" of more than three consecutive guanine (G) nucleotides or more than three consecutive cytosine (C) nucleotides. These criteria can be assessed using various techniques known in the art, for example, computer programs such as BLAST can be used to search publicly available databases to determine whether the selected target sequence is unique to the target mRNA. Alternatively, a target sequence can be selected (and a small interfering RNA sequence designed) using computer software available commercially (for example, OligoEngine, Target Finder and the small interfering RNA Design Tool which are commercially available).

Target mRNA sequences may be selected that are between about 14 and about 30 nucleotides in length that meet one or more of the above criteria, or target sequences may be selected that are between about 16 and about 30 nucleotides in length that meet one or more of the above criteria. Target sequences may be selected that are between about 19 and about 30 nucleotides in length that meet one or more of the above criteria. Target sequences are selected that are between about 19 and about 25 nucleotides in length that meet one or more of the above criteria.

The small interfering RNA molecules may comprise a specific antisense sequence that is complementary to at least 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, or more contiguous nucleotides of any one of the polynucleotide sequences described herein.

The specific antisense sequence comprised by the small interfering RNA molecule may be identical or substantially identical to the complement of the target sequence. The specific antisense sequence comprised by the small interfering RNA molecule may be at least about

75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identical to the complement of the target mRNA sequence. Methods of determining sequence identity are known in the art and can be determined, for example, by using the BLASTN program of the University of Wisconsin Computer Group (GCG) software or provided on the NCBI website.

5 The specific antisense sequence of the small interfering RNA molecules may exhibit variability by differing (for example, by nucleotide substitution, including transition or transversion) at one, two, three, four or more nucleotides from the sequence of the target mRNA. When such nucleotide substitutions are present in the antisense strand of a double-stranded RNA molecule, the complementary nucleotide in the sense strand with which the substitute nucleotide would  
10 typically form hydrogen bond base-pairing may or may not be correspondingly substituted. double-stranded RNA molecules in which one or more nucleotide substitution occurs in the sense sequence, but not in the antisense strand, are also contemplated. When the antisense sequence of an small interfering RNA molecule comprises one or more mismatches between the nucleotide sequence of the small interfering RNA and the target nucleotide sequence, as  
15 described above, the mismatches may be found at the 3' terminus, the 5' terminus or in the central portion of the antisense sequence.

The small interfering RNA molecules may comprise a specific antisense sequence that is capable of selectively hybridizing under stringent conditions to a portion of a naturally occurring target gene or target mRNA. As known to those of ordinary skill in the art, variations in  
20 stringency of hybridization conditions may be achieved by altering the time, temperature or concentration of the solutions used for the hybridization and wash steps. Suitable conditions can also depend in part on the particular nucleotide sequences used, for example the sequence of the target mRNA or gene.

One method for inducing double stranded RNA-silencing in plants is transformation with a gene  
25 construct producing hairpin RNA (see Smith *et al.* (2000) *Nature*, 407, 319-320). Such constructs comprise inverted regions of the target gene sequence, separated by an appropriate spacer. The insertion of a functional plant intron region as a spacer fragment additionally increases the efficiency of the gene silencing induction, due to generation of an intron spliced hairpin RNA (Wesley *et al.* (2001) *Plant J.*, 27, 581-590). Suitably, the stem length is about 50  
30 nucleotides to about 1 kilo bases in length. Methods for producing intron spliced hairpin RNA are well described in the art (see for example, *Bioscience, Biotechnology, and Biochemistry* (2008) 72, 2, 615-617).

Interfering RNA molecules having a duplex or double-stranded structure, for example double-stranded RNA or small hairpin RNA, can have blunt ends, or can have 3' or 5' overhangs. As  
35 used herein, "overhang" refers to the unpaired nucleotide or nucleotides that protrude from a duplex structure when a 3'-terminus of one RNA strand extends beyond the 5'-terminus of the other strand (3' overhang), or vice versa (5' overhang). The nucleotides comprising the

overhang can be ribonucleotides, deoxyribonucleotides or modified versions thereof. In one embodiment, at least one strand of the interfering RNA molecule has a 3' overhang from about 1 to about 6 nucleotides in length. In other embodiments, the 3' overhang is from about 1 to about 5 nucleotides, from about 1 to about 3 nucleotides and from about 2 to about 4 nucleotides in length.

When the interfering RNA molecule comprises a 3' overhang at one end of the molecule, the other end can be blunt-ended or have also an overhang (5' or 3'). When the interfering RNA molecule comprises an overhang at both ends of the molecule, the length of the overhangs may be the same or different. In one embodiment, the interfering RNA molecule comprises 3' overhangs of about 1 to about 3 nucleotides on both ends of the molecule. In a further embodiment, the interfering RNA molecule is a double-stranded RNA having a 3' overhang of 2 nucleotides at both ends of the molecule. In yet another embodiment, the nucleotides comprising the overhang of the interfering RNA are TT dinucleotides or UU dinucleotides.

When determining the percentage identity of the interfering RNA molecule comprising one or more overhangs to the target mRNA sequence, the overhang(s) may or may not be taken into account. For example, the nucleotides from a 3' overhang and up to 2 nucleotides from the 5'- or 3'-terminus of the double strand may be modified without significant loss of activity of the small interfering RNA molecule.

The interfering RNA molecules can comprise one or more 5' or 3'-cap structures. The interfering RNA molecule can comprise a cap structure at the 3'-end of the sense strand, the antisense strand, or both the sense and antisense strands; or at the 5'-end of the sense strand, the antisense strand, or both the sense and antisense strands of the interfering RNA molecule. Alternatively, the interfering RNA molecule can comprise a cap structure at both the 3'-end and 5'-end of the interfering RNA molecule. The term "cap structure" refers to a chemical modification incorporated at either terminus of an oligonucleotide, which protects the molecule from exonuclease degradation, and may also facilitate delivery or localisation within a cell.

Another modification applicable to interfering RNA molecules is the chemical linkage to the interfering RNA molecule of one or more moieties or conjugates which enhance the activity, cellular distribution, cellular uptake, bioavailability or stability of the interfering RNA molecule.

The polynucleotides may be synthesized or modified by methods well established in the art. Chemical modifications may include, but are not limited to 2' modifications, introduction of non-natural bases, covalent attachment to a ligand, and replacement of phosphate linkages with thiophosphate linkages. In this embodiment, the integrity of the duplex structure is strengthened by at least one, and typically two, chemical linkages. Chemical linking may be achieved by any of a variety of well-known techniques, for example by introducing covalent, ionic or hydrogen bonds; hydrophobic interactions, van der Waals or stacking interactions; by means of metal-ion coordination, or through use of purine analogues.

In yet another embodiment, the nucleotides at one or both of the two single strands may be modified to reduce or inhibit the activation of cellular enzymes, such as, for example, without limitation, certain nucleases. Techniques for reducing or inhibiting the activation of cellular enzymes are known in the art including, but not limited to, 2'-amino modifications, 2'-fluoro  
 5 modifications, 2'-alkyl modifications, uncharged backbone modifications, morpholino modifications, 2'-O-methyl modifications, and phosphoramidate. Thus, at least one 2'-hydroxyl group of the nucleotides on a double-stranded RNA is replaced by a chemical group. Also, at least one nucleotide may be modified to form a locked nucleotide. Such locked nucleotide contains a methylene or ethylene bridge that connects the 2'-oxygen of ribose with the 4'-carbon  
 10 of ribose. Introduction of a locked nucleotide into an oligonucleotide improves the affinity for complementary sequences and increases the melting temperature by several degrees.

Ligands may be conjugated to a interfering RNA molecule, for example, to enhance its cellular absorption. In certain embodiments, a hydrophobic ligand is conjugated to the molecule to facilitate direct permeation of the cellular membrane. These approaches have been used to  
 15 facilitate cell permeation of antisense oligonucleotides. In certain instances, conjugation of a cationic ligand to oligonucleotides often results in improved resistance to nucleases. Representative examples of cationic ligands include propylammonium and dimethylpropylammonium. Anti-sense oligonucleotides can retain their high binding affinity to mRNA when the cationic ligand is dispersed throughout the oligonucleotide.

20 According to one embodiment, the expression or activity of isopropylmalate synthase is decreased in a plant already capable of producing beta-methylvaleryl containing sucrose esters thereby resulting in the decreased accumulation of isopropylmalate synthase therein and thus decreased beta-methylvaleryl containing sucrose ester production.

The molecules and nucleotides described herein may be prepared using well-known techniques  
 25 of solid-phase synthesis. Any other means for such synthesis known in the art may additionally or alternatively be employed.

Various embodiments are directed to expression vectors comprising one or more of the polynucleotides described herein or interfering RNA constructs that comprise one or more polynucleotides.

30 Various embodiments are directed to expression vectors comprising one or more polynucleotides or one or more interfering RNA constructs.

Various embodiments are directed to expression vectors comprising one or more polynucleotides or one or more interfering RNA constructs encoding one or more interfering RNA polynucleotides capable of self-annealing to form a hairpin structure, in which the construct  
 35 comprises (a) one or more of the polynucleotides described herein; (b) a second sequence encoding a spacer element that includes an intron, which when spliced, forms a loop of the hairpin structure; and (c) a third sequence comprising a reverse complementary sequence of

the first sequence, positioned in the same orientation as the first sequence, wherein the second sequence is positioned between the first sequence and the third sequence, and the second sequence is operably-linked to the first sequence and to the third sequence.

The disclosed sequences can be utilized for constructing various polynucleotides that do not form hairpin structures. For example, a double-stranded RNA can be formed by (1) transcribing a first strand of the DNA by operably-linking to a first promoter, and (2) transcribing the reverse complementary sequence of the first strand of the DNA fragment by operably-linking to a second promoter. Each strand of the polynucleotide can be transcribed from the same expression vector, or from different expression vectors. The RNA duplex having RNA interference activity can be enzymatically converted to small interfering RNAs to reduce RNA levels.

Thus, various embodiments are directed to expression vectors comprising polynucleotide or interfering RNA constructs encoding interfering RNA polynucleotides capable of self-annealing, in which the construct comprises (a) one or more of the polynucleotides described herein; and (b) a second sequence comprising a complementary (for example, reverse complementary) sequence of the first sequence, positioned in the same orientation as the first sequence.

Various compositions and methods are provided for reducing the endogenous expression levels of isopropylmalate synthase by promoting co-suppression of gene expression. The phenomenon of co-suppression occurs as a result of introducing multiple copies of a transgene into a plant cell host. Integration of multiple copies of a transgene can result in reduced expression of the transgene and the targeted endogenous gene. The degree of co-suppression is dependent on the degree of sequence identity between the transgene and the targeted endogenous gene. The silencing of both the endogenous gene and the transgene can occur by extensive methylation of the silenced loci (that is, the endogenous promoter and endogenous gene of interest) that can preclude transcription. Alternatively, in some cases, co-suppression of the endogenous gene and the transgene can occur by post transcriptional gene silencing, in which transcripts can be produced but enhanced rates of degradation preclude accumulation of transcripts. The mechanism for co-suppression by post-transcriptional gene silencing is thought to resemble RNA interference, in that RNA seems to be both an important initiator and a target in these processes, and may be mediated at least in part by the same molecular machinery, possibly through RNA-guided degradation of mRNAs.

Co-suppression of nucleic acids can be achieved by integrating multiple copies of the nucleic acid or fragments thereof, as transgenes, into the genome of a plant of interest. The host plant can be transformed with an expression vector comprising a promoter operably-linked to the nucleic acid or fragments thereof. Various embodiments are directed to expression vectors for promoting co-suppression of endogenous genes comprising a promoter operably-linked to a polynucleotide.

Methods for obtaining conservative mutant polynucleotides and polypeptides are also described. Any plant of interest, including a plant cell or plant material, can be genetically modified by various methods known to induce mutagenesis, including site-directed mutagenesis, oligonucleotide-directed mutagenesis, chemically-induced mutagenesis, irradiation-induced mutagenesis, mutagenesis utilizing modified bases, mutagenesis utilizing gapped duplex DNA, double-strand break mutagenesis, mutagenesis utilizing repair-deficient host strains, mutagenesis by total gene synthesis, DNA shuffling and other equivalent methods. Alternatively, genes can be targeted for inactivation by introducing transposons (for example, IS elements) into the genomes of plants of interest. These mobile genetic elements can be introduced by sexual cross-fertilization and insertion mutants can be screened for loss in protein activity. The disrupted gene in a parent plant can be introduced into other plants by crossing the parent plant with plant not subjected to transposon-induced mutagenesis by, for example, sexual cross-fertilization. Any standard breeding techniques known to persons skilled in the art can be utilized. In one embodiment, one or more genes can be inactivated by the insertion of one or more transposons. Mutations can result in homozygous disruption of one or more genes, in heterozygous disruption of one or more genes, or a combination of both homozygous and heterozygous disruptions if more than one gene is disrupted. Suitable transposable elements include retrotransposons, retroposons, and SINE-like elements. Such methods are known to persons skilled in the art.

Alternatively, genes can be targeted for inactivation by introducing ribozymes derived from a number of small circular RNAs that are capable of self-cleavage and replication in plants. These RNAs can replicate either alone (viroid RNAs) or with a helper virus (satellite RNAs). Examples of suitable RNAs include those derived from avocado sunblotch viroid and satellite RNAs derived from tobacco ringspot virus, lucerne transient streak virus, velvet tobacco mottle virus, solanum nodiflorum mottle virus, and subterranean clover mottle virus. Various target RNA-specific ribozymes are known to persons skilled in the art.

In some embodiments, the expression of an isopropylmalate synthase polypeptide can be modulated by non-transgenic means, such as creating a mutation in a gene. Methods that introduce a mutation randomly in a gene sequence can include chemical mutagenesis, EMS mutagenesis and radiation mutagenesis. Methods that introduce one or more targeted mutations into a cell include but are not limited to genome editing technology, particularly zinc finger nuclease-mediated mutagenesis, tilling (targeting induced local lesions in genomes), homologous recombination, oligonucleotide-directed mutagenesis, and meganuclease-mediated mutagenesis.

Some non-limiting examples of mutations are deletions, insertions and missense mutations of at least one nucleotide, single nucleotide polymorphisms and a simple sequence repeat. After mutation, screening can be performed to identify mutations that create premature stop codons

or otherwise non-functional genes. After mutation, screening can be performed to identify mutations that create functional genes that are capable of being expressed at elevated levels. Screening of mutants can be carried out by sequencing, or by the use of probes or primers specific to the gene or protein. Specific mutations in polynucleotides can also be created that can result in reduced or increased gene expression, reduced or increased stability of mRNA, or reduced or increased stability of protein. Such plants are referred to herein as mutant plants.

The mutant plants can have any combination of one or more mutations which results in modulated polypeptide levels. For example, the mutant plants may have a single mutation in a single gene; multiple mutations in a single gene; a single mutation in two or more or three or more genes; or multiple mutations in two or more or three or more genes. Accordingly, mutant plants comprising the mutant polypeptide variants are disclosed.

In one embodiment, seeds from plants are mutagenised and then grown into first generation mutant plants. The first generation plants are then allowed to self-pollinate and seeds from the first generation plant are grown into second generation plants, which are then screened for mutations in their loci. Though the mutagenized plant material can be screened for mutations, an advantage of screening the second generation plants is that all somatic mutations correspond to germline mutations. One of skill in the art would understand that a variety of plant materials, including but not limited to, seeds, pollen, plant tissue or plant cells, may be mutagenised in order to create the mutant plants. However, the type of plant material mutagenised may affect when the plant nucleic acid is screened for mutations. For example, when pollen is subjected to mutagenesis prior to pollination of a non-mutagenized plant, the seeds resulting from that pollination are grown into first generation plants. Every cell of the first generation plants will contain mutations created in the pollen; thus these first generation plants may then be screened for mutations instead of waiting until the second generation.

Mutagens that create point mutations, short deletions, insertions, transversions, or transitions, including chemical mutagens or radiation, may be used to create the mutations. Mutagens include, but are not limited to, ethyl methanesulfonate, methylmethane sulfonate, N-ethyl-N-nitrosourea, triethylmelamine, N-methyl-N-nitrosourea, procarbazine, chlorambucil, cyclophosphamide, diethyl sulfate, acrylamide monomer, melphalan, nitrogen mustard, vincristine, dimethylnitrosamine, N-methyl-N'-nitro-Nitrosoguanidine, nitrosoguanidine, 2-aminopurine, 7,12 dimethyl-benz(a)anthracene, ethylene oxide, hexamethylphosphoramide, bisulfan, diepoxyalkanes (diepoxyoctane, diepoxybutane, and the like), 2-methoxy-6-chloro-9[3-(ethyl-2-chloro-ethyl)aminopropylamino]acridine dihydrochloride and formaldehyde. Spontaneous mutations in the locus that may not have been directly caused by the mutagen are also contemplated provided that they result in the desired phenotype. Suitable mutagenic agents also include, for example, ionising radiation – such as X-rays, gamma rays, fast neutron irradiation and UV radiation.

Any method of plant nucleic acid preparation known to those of skill in the art may be used to prepare the plant nucleic acid for mutation screening.

Prepared nucleic acid from individual plants, plant cells, or plant material can optionally be pooled in order to expedite screening for mutations in the isopropylmalate synthase gene of the population of plants originating from the mutagenized plant tissue, cells or material. One or more subsequent generations of plants, plant cells or plant material can be screened. The size of the optionally pooled group is dependent upon the sensitivity of the screening method used.

After the nucleic acid samples are optionally pooled, they can be subjected to polynucleotide-specific amplification techniques, such as Polymerase Chain Reaction (PCR). Any one or more primers or probes specific to the isopropylmalate synthase gene or the sequences immediately adjacent thereto may be utilized to amplify the sequences within the optionally pooled nucleic acid sample. Preferably, the one or more primers are designed to amplify the regions of the isopropylmalate synthase locus where useful mutations are most likely to arise. Most preferably, the primer is designed to detect mutations within regions of the isopropylmalate synthase polynucleotide. Additionally, it is preferable for the primer(s) to avoid known polymorphic sites in order to ease screening for point mutations. To facilitate detection of amplification products, the one or more primers or probes may be labelled using any conventional labelling method.

Primer(s) or probes can be designed based upon the sequences described herein using methods that are well understood in the art. Polymorphisms may be identified by means known in the art. In a further aspect there is provided a method of preparing a mutant plant. The method involves providing at least one cell of a plant comprising a gene encoding a functional isopropylmalate synthase polypeptide. Next, the at least one cell of the plant is treated under conditions effective to modulate the activity of the isopropylmalate synthase gene. The at least one mutant plant cell is then propagated into a mutant plant, where the mutant plant has a modulated level of isopropylmalate synthase polypeptide as compared to that of a control plant. In one embodiment of this method of making a mutant plant, the treating step involves subjecting the at least one cell to a chemical mutagenising agent as described above and under conditions effective to yield at least one mutant plant cell. In another embodiment of this method, the treating step involves subjecting the at least one cell to a radiation source under conditions effective to yield at least one mutant plant cell. The term "mutant plant" includes mutant plants in which the genotype is modified as compared to a control plant, suitably by means other than genetic engineering or genetic modification.

In certain embodiments, the mutant plant, mutant plant cell or mutant plant material may comprise one or more mutations that have occurred naturally in another plant, plant cell or plant material and confer a desired trait. This mutation can be incorporated (for example, introgressed) into another plant, plant cell or plant material (for example, a plant, plant cell or plant material with a different genetic background to the plant from which the mutation was

derived) to confer the trait thereto. Thus by way of example, a mutation that occurred naturally in a first plant may be introduced into a second plant – such as a second plant with a different genetic background to the first plant. The skilled person is therefore able to search for and identify a plant carrying naturally in its genome one or more mutant alleles of the isopropylmalate synthase gene which confer a desired trait. The mutant allele(s) that occurs naturally can be transferred to the second plant by various methods including breeding, backcrossing and introgression to produce a lines, varieties or hybrids that have one or more mutations in the isopropylmalate synthase gene. Plants showing a desired trait may be screened out of a pool of mutant plants. Suitably, the selection is carried out utilising the knowledge of the isopropylmalate synthase nucleotide sequences as described herein. Consequently, it is possible to screen for a genetic trait being indicative of changes in sucrose ester composition as compared to a control. Such a screening approach may involve the application of conventional nucleic acid amplification and/or hybridization techniques as discussed herein. Thus, a further aspect of the present invention relates to a method for identifying a mutant plant comprising the steps of: (a) providing a sample comprising a isopropylmalate synthase polynucleotide from a plant; and (b) determining the nucleic acid sequence of the isopropylmalate synthase polynucleotide, wherein a difference in the sequence of the isopropylmalate synthase polynucleotide as compared to the isopropylmalate synthase polynucleotide of a control plant is indicative that said plant is an isopropylmalate synthase mutant plant. In another aspect there is provided a method for identifying a mutant plant which has a change in the composition of sucrose esters as compared to a control plant comprising the steps of: (a) providing a sample from a plant to be screened; (b) determining if said sample comprises one or more mutations in the isopropylmalate synthase polynucleotide; and (c) determining the composition of the sucrose esters in said plant; wherein if said sample comprises one or more mutations in the isopropylmalate synthase polynucleotide that modulate the expression or the activity of the protein encoded as compared to a control plant and a part of the plant has a change in the composition of sucrose esters as compared to a control plant in which the expression or the activity of isopropylmalate synthase has not been reduced is indicative of a mutant plant which has a change in the composition of sucrose esters. In another aspect there is provided a method for preparing a mutant plant which has a change in the composition of sucrose esters as compared to a control plant comprising the steps of: (a) providing a sample from a first plant; (b) determining if said sample comprises one or more mutations in the isopropylmalate synthase polynucleotide that result in a change in the composition of sucrose esters therein; and (c) transferring the one or more mutations into a second plant. The mutation(s) can be transferred into the second plant using various methods – that are known in the art – such as by genetic engineering, genetic manipulation, introgression, plant breeding, backcrossing and the like. In one embodiment, the first plant is a naturally

occurring plant. In one embodiment, the second plant has a different genetic background to the first plant. In another aspect there is provided a method for preparing a mutant plant which has a change in the composition of sucrose esters as compared to a control plant comprising the steps of: (a) providing a sample from a first plant; (b) determining if said sample comprises one or more mutations in the isopropylmalate synthase polynucleotide that results in a change in the composition of sucrose esters therein; and (c) introgressing the one or more mutations from the first plant into a second plant. In one embodiment, the step of introgressing comprises plant breeding, optionally including backcrossing and the like. In one embodiment, the first plant is a naturally occurring plant. In one embodiment, the second plant has a different genetic background to the first plant. In one embodiment, the first plant is not a cultivar or an elite cultivar. In one embodiment, the second plant is a cultivar or an elite cultivar. A further aspect relates to a mutant plant (including a cultivar or elite cultivar mutant plant) obtained or obtainable by the methods described herein. In certain embodiments, the mutant plants may have one or more mutations localised only to a specific region of the plant – such as within the sequence of the isopropylmalate synthase polynucleotide. According to this embodiment, the remaining genomic sequence of the mutant plant will be the same or substantially the same as the plant prior to the mutagenesis.

In certain embodiments, the mutant plant may have one or more mutations localised in more than one region of the plant – such as within the sequence of the isopropylmalate synthase polynucleotide – and in one or more further regions of the genome. According to this embodiment, the remaining genomic sequence of the mutant plant will not be the same or will not be substantially the same as the plant prior to the mutagenesis. In certain embodiments, the mutant plant does not have one or more mutations in one or more, two or more, three or more, four or more or five or more exons of the isopropylmalate synthase polynucleotide; or in one or more, two or more, three or more, four or more or five or more introns of the isopropylmalate synthase polynucleotide; in a promoter of the isopropylmalate synthase polynucleotide; in the 3' untranslated region of the isopropylmalate synthase polynucleotide; in the 5' untranslated region of the isopropylmalate synthase polynucleotide; in the coding region of the isopropylmalate synthase polynucleotide; or in the non-coding region of the isopropylmalate synthase polynucleotide; or any combination of two or more, three or more, four or more, five or more; or six or more thereof parts thereof. In a further aspect there is provided a method of identifying a plant, a plant cell or plant material comprising a mutation in a gene encoding isopropylmalate synthase comprising: (a) subjecting a plant, a plant cell or plant material to mutagenesis; (b) obtaining a nucleic acid sample from said plant, plant cell or plant material or descendants thereof; and (c) determining the nucleic acid sequence of the gene encoding isopropylmalate synthase or a variant or a fragment thereof, wherein a difference in said sequence is indicative of one or more mutations therein.

Zinc finger proteins can be used to modulate expression or the activity of isopropylmalate synthase. In various embodiments, a genomic DNA sequence comprising a part of or all of the coding sequence of a polynucleotide is modified by zinc finger nuclease mediated mutagenesis. The genomic DNA sequence is searched for a unique site for zinc finger protein binding.

5 Alternatively, the genomic DNA sequence is searched for two unique sites for zinc finger protein binding wherein both sites are on opposite strands and close together, for example, 1, 2, 3, 4, 5, 6 or more basepairs apart. Accordingly, zinc finger proteins that bind to the polynucleotides described herein are provided.

A zinc finger protein may be engineered to recognize a selected target site in the isopropylmalate synthase gene. A zinc finger protein can comprise any combination of motifs derived from natural zinc finger DNA-binding domains and non-natural zinc finger DNA-binding domains by truncation or expansion or a process of site-directed mutagenesis coupled to a selection method such as, but not limited to, phage display selection, bacterial two-hybrid selection or bacterial one-hybrid selection. The term "non-natural zinc finger DNA-binding domain" refers to a zinc finger DNA-binding domain that binds a three-basepair sequence within the target DNA and that does not occur in the cell or organism comprising the DNA which is to be modified. Methods for the design of zinc finger protein which binds specific nucleotide sequences which are unique to a target gene are known in the art.

A zinc finger nuclease may be constructed by making a fusion of a first polynucleotide coding for a zinc finger protein that binds to a polynucleotide, and a second polynucleotide coding for a non-specific endonuclease such as, but not limited to, those of a Type IIS endonuclease. A fusion protein between a zinc finger protein and the nuclease may comprise a spacer consisting of two basepairs or alternatively, the spacer can consist of three, four, five, six or seven or more basepairs. In various embodiments, a zinc finger nuclease introduces a double stranded break in a regulatory region, a coding region, or a non-coding region of a genomic DNA sequence of a polynucleotide and leads to a reduction of the level of expression of a polynucleotide, or a reduction in the activity of the protein encoded thereby. Cleavage by zinc finger nuclease frequently results in the deletion of DNA at the cleavage site following DNA repair by non-homologous end joining.

30 In other embodiments, a zinc finger protein may be selected to bind to a regulatory sequence of a isopropylmalate synthase polynucleotide. More specifically, the regulatory sequence may comprise a transcription initiation site, a start codon, a region of an exon, a boundary of an exon-intron, a terminator, or a stop codon. Accordingly, the disclosure provides a mutant, non-naturally occurring or transgenic plant or plant cells, produced by zinc finger nuclease-mediated mutagenesis in the vicinity of or within the isopropylmalate synthase gene, and methods for making such a plant or plant cell by zinc finger nuclease-mediated mutagenesis. Methods for

delivering zinc finger protein and zinc finger nuclease to a plant are similar to those described below for delivery of meganuclease.

In another aspect, methods for producing recombinant, mutant, non-naturally occurring or transgenic or otherwise genetically-modified plants using meganucleases – such as  
5 meganuclease I-Crel – are provided. Naturally occurring meganucleases as well as recombinant meganucleases may be used to specifically cause a double-stranded break at a single site or at relatively few sites in the genomic DNA of a plant to allow for the disruption of a isopropylmalate synthase polynucleotide. The meganuclease may be an engineered  
10 meganuclease with altered DNA-recognition properties. Meganuclease proteins can be delivered into plant cells by a variety of different mechanisms known in the art.

The disclosure encompasses the use of meganucleases to inactivate one or more of the polynucleotides described herein in a plant cell or plant. Aspects also relate to a method for inactivating a polynucleotide in a plant using a meganuclease comprising: (a) providing a plant  
15 cell comprising one or more polynucleotides described herein; (b) introducing a meganuclease or a construct encoding a meganuclease into said plant cell; and (c) allowing the meganuclease to substantially inactivate the polynucleotide.

Meganucleases can be used to cleave meganuclease recognition sites within the coding regions of a polynucleotide. Such cleavage frequently results in the deletion of DNA at the meganuclease recognition site following mutagenic DNA repair by non-homologous end joining.  
20 Such mutations in the gene coding sequence are typically sufficient to inactivate the gene. This method to modify a plant cell involves, first, the delivery of a meganuclease expression cassette to a plant cell using a suitable transformation method. For highest efficiency, it is desirable to link the meganuclease expression cassette to a selectable marker and select for successfully transformed cells in the presence of a selection agent. This approach will result in the  
25 integration of the meganuclease expression cassette into the genome, however, which may not be desirable if the plant is likely to require regulatory approval. In such cases, the meganuclease expression cassette (and linked selectable marker gene) may be segregated away in subsequent plant generations using conventional breeding techniques. Alternatively, plant cells may be initially be transformed with a meganuclease expression cassette lacking a  
30 selectable marker and may be grown on media lacking a selection agent. Under such conditions, a fraction of the treated cells will acquire the meganuclease expression cassette and will express the engineered meganuclease transiently without integrating the meganuclease expression cassette into the genome. Because it does not account for transformation efficiency, this latter transformation procedure requires that a greater number of treated cells be  
35 screened to obtain the desired genome modification. The above approach can also be applied to modify a plant cell when using a zinc finger protein or zinc finger nuclease.

Following delivery of the meganuclease expression cassette, plant cells are grown, initially, under conditions that are typical for the particular transformation procedure that was used. This may mean growing transformed cells on media at temperatures below 26.degree.C., frequently in the dark. Such standard conditions can be used for a period of time, preferably 1-4 days, to allow the plant cell to recover from the transformation process. At any point following this initial recovery period, growth temperature may be raised to stimulate the activity of the engineered meganuclease to cleave and mutate the meganuclease recognition site.

For certain applications, it may be desirable to precisely remove the polynucleotide from the genome of the plant. Such applications are possible using a pair of engineered meganucleases, each of which cleaves a meganuclease recognition site on either side of the intended deletion.

Meganucleases and other DNA-repair strategies may also be used to trigger a heterologous recombination between a DNA contained in a plant cell and the target DNA. The target DNA typically comprises at least one, preferably two regions that are homologous to the DNA contained in the plant cell. Following DNA repair by the cell, homologous ends are joined resulting in the introduction of the target DNA into the cell.

In a further embodiment, expression levels of isopropylmalate synthase in a plant may be increased by inserting the gene directly into the chloroplast genome since isopropylmalate synthase is a chloroplastic protein encoded in the plant cell nucleus. Likewise, increasing the transport or the import of isopropylmalate synthase protein into chloroplasts can also modulate the activity pattern.

One object is to provide mutant, transgenic or non-naturally occurring plants that exhibit modulated sucrose ester levels whilst maintaining substantially the same visual appearance as compared to a control plant. Accordingly, there is described herein mutant, transgenic or non-naturally occurring plants or cells that have modulated levels of sucrose esters as compared to control cells or control plants. The mutant, transgenic or non-naturally occurring plants or cells have been modified to modulate the synthesis or activity of isopropylmalate synthase described herein by modulating the expression of one or more polypeptides encoding the polynucleotide sequences described herein.

A further aspect, relates to a mutant, non-naturally occurring or transgenic plant or cell, wherein the expression of or the activity of isopropylmalate synthase is modulated and a part of the plant has an increase or a decrease in sucrose ester levels of at least 5% as compared to a control plant in which the expression or the activity said enzyme has not been modulated. A still further aspect, relates to a mutant, non-naturally occurring or transgenic plant or cell, wherein expression of isopropylmalate synthase or the activity of the protein encoded thereby is modulated and wherein the sucrose ester levels in aerosol is increased or decreased by at least 5% as compared to the aerosol from the control plant.

The change in the sucrose ester content as compared to the control plant may be at least about 5 %, at least about 10 %, at least about 20 %, at least about 25 %, at least about 30 %, at least about 40 %, at least about 50 %, at least about 60 %, at least about 70 %, at least about 75 %, at least about 80 %, at least about 90 %, at least about 95 %, at least about 96 %, at least about 97 %, at least about 98 %, at least about 99 %, or about 100 % or more.

The plant may be heated to 100°C or above – such as at least 125°C, at least 150°C, at least 175°C or at least 200°C - to release the aerosol.

Plants suitable for use in genetic modification include monocotyledonous and dicotyledonous plants and plant cell systems, including species from one of the following families: *Acanthaceae*,

10 *Alliaceae*, *Alstroemeriaceae*, *Amaryllidaceae*, *Apocynaceae*, *Arecaceae*, *Asteraceae*, *Berberidaceae*, *Bixaceae*, *Brassicaceae*, *Bromeliaceae*, *Cannabaceae*, *Caryophyllaceae*, *Cephalotaxaceae*, *Chenopodiaceae*, *Colchicaceae*, *Cucurbitaceae*, *Dioscoreaceae*, *Ephedraceae*, *Erythroxylaceae*, *Euphorbiaceae*, *Fabaceae*, *Lamiaceae*, *Linaceae*, *Lycopodiaceae*, *Malvaceae*, *Melanthiaceae*, *Musaceae*, *Myrtaceae*, *Nyssaceae*, *Papaveraceae*,  
15 *Pinaceae*, *Plantaginaceae*, *Poaceae*, *Rosaceae*, *Rubiaceae*, *Salicaceae*, *Sapindaceae*, *Solanaceae*, *Taxaceae*, *Theaceae*, or *Vitaceae*.

Suitable species may include members of the genera *Abelmoschus*, *Abies*, *Acer*, *Agrostis*, *Allium*, *Alstroemeria*, *Ananas*, *Andrographis*, *Andropogon*, *Artemisia*, *Arundo*, *Atropa*, *Berberis*, *Beta*, *Bixa*, *Brassica*, *Calendula*, *Camellia*, *Camptotheca*, *Cannabis*, *Capsicum*, *Carthamus*,  
20 *Catharanthus*, *Cephalotaxus*, *Chrysanthemum*, *Cinchona*, *Citrullus*, *Coffea*, *Colchicum*, *Coleus*, *Cucumis*, *Cucurbita*, *Cynodon*, *Datura*, *Dianthus*, *Digitalis*, *Dioscorea*, *Elaeis*, *Ephedra*, *Erianthus*, *Erythroxylum*, *Eucalyptus*, *Festuca*, *Fragaria*, *Galanthus*, *Glycine*, *Gossypium*, *Helianthus*, *Hevea*, *Hordeum*, *Hyoscyamus*, *Jatropha*, *Lactuca*, *Linum*, *Lolium*, *Lupinus*, *Lycopersicon*, *Lycopodium*, *Manihot*, *Medicago*, *Mentha*, *Miscanthus*, *Musa*, *Nicotiana*, *Oryza*,  
25 *Panicum*, *Papaver*, *Parthenium*, *Pennisetum*, *Petunia*, *Phalaris*, *Phleum*, *Pinus*, *Poa*, *Poinsettia*, *Populus*, *Rauwolfia*, *Ricinus*, *Rosa*, *Saccharum*, *Salix*, *Sanguinaria*, *Scopolia*, *Secale*, *Solanum*, *Sorghum*, *Spartina*, *Spinacea*, *Tanacetum*, *Taxus*, *Theobroma*, *Triticosecale*, *Triticum*, *Uniola*, *Veratrum*, *Vinca*, *Vitis*, and *Zea*.

Suitable species may include *Panicum* spp., *Sorghum* spp., *Miscanthus* spp., *Saccharum* spp.,  
30 *Erianthus* spp., *Populus* spp., *Andropogon gerardii* (big bluestem), *Pennisetum purpureum* (elephant grass), *Phalaris arundinacea* (reed canarygrass), *Cynodon dactylon* (bermudagrass), *Festuca arundinacea* (tall fescue), *Spartina pectinata* (prairie cord-grass), *Medicago sativa* (alfalfa), *Arundo donax* (giant reed), *Secale cereale* (rye), *Salix* spp. (willow), *Eucalyptus* spp. (eucalyptus), *Triticosecale*, bamboo, *Helianthus annuus* (sunflower), *Carthamus tinctorius* (safflower), *Jatropha curcas* (jatropha), *Ricinus communis* (castor), *Elaeis guineensis* (palm),  
35 *Linum usitatissimum* (flax), *Brassica juncea*, *Beta vulgaris* (sugarbeet), *Manihot esculenta* (cassava), *Lycopersicon esculentum* (tomato), *Lactuca sativa* (lettuce), *Musa paradisiaca*

- (banana), *Solanum tuberosum* (potato), *Brassica oleracea* (broccoli, cauliflower, Brussels sprouts), *Camellia sinensis* (tea), *Fragaria ananassa* (strawberry), *Theobroma cacao* (cocoa), *Coffea arabica* (coffee), *Vitis vinifera* (grape), *Ananas comosus* (pineapple), *Capsicum annum* (hot & sweet pepper), *Allium cepa* (onion), *Cucumis melo* (melon), *Cucumis sativus* (cucumber),
- 5 *Cucurbita maxima* (squash), *Cucurbita moschata* (squash), *Spinacea oleracea* (spinach), *Citrullus lanatus* (watermelon), *Abelmoschus esculentus* (okra), *Solanum melongena* (eggplant), *Rosa* spp. (rose), *Dianthus caryophyllus* (carnation), *Petunia* spp. (petunia), *Poinsettia pulcherrima* (poinsettia), *Lupinus albus* (lupin), *Uniola paniculata* (oats), bentgrass (*Agrostis* spp.), *Populus tremuloides* (aspen), *Pinus* spp. (pine), *Abies* spp. (fir), *Acer* spp. (maple),
- 10 *Hordeum vulgare* (barley), *Poa pratensis* (bluegrass), *Lolium* spp. (ryegrass) and *Phleum pratense* (timothy), *Panicum virgatum* (switchgrass), *Sorghum bicolor* (sorghum, sudangrass), *Miscanthus giganteus* (miscanthus), *Saccharum* sp. (energy cane), *Populus balsamifera* (poplar), *Zea mays* (corn), *Glycine max* (soybean), *Brassica napus* (canola), *Triticum aestivum* (wheat), *Gossypium hirsutum* (cotton), *Oryza sativa* (rice), *Helianthus annuus* (sunflower),
- 15 *Medicago sativa* (alfalfa), *Beta vulgaris* (sugarbeet), or *Pennisetum glaucum* (pearl millet).
- Various embodiments are directed to mutant plants, non-naturally occurring plants or transgenic plants modified to reduce isopropylmalate synthase gene expression levels thereby, producing plants – such as tobacco plants – in which the expression level of isopropylmalate synthase are reduced within plant tissues of interest as compared to a control plant. The disclosed
- 20 compositions and methods can be applied to any species of the genus *Nicotiana*, including *N. rustica* and *N. tabacum* (for example, LA B21, LN KY171, TI 1406, Basma, Galpao, Perique, Beinhart 1000-1, K326, Hicks Broadleaf and Petico). Other species include *N. acaulis*, *N. acuminata*, *N. acuminata* var. *multiflora*, *N. africana*, *N. alata*, *N. amplexicaulis*, *N. arentsii*, *N. attenuata*, *N. benavidesii*, *N. benthamiana*, *N. bigelovii*, *N. bonariensis*, *N. cavicola*, *N. clevelandii*, *N. cordifolia*, *N. corymbosa*, *N. debneyi*, *N. excelsior*, *N. forgetiana*, *N. fragrans*, *N. glauca*, *N. glutinosa*, *N. goodspeedii*, *N. gossei*, *N. hybrid*, *N. ingulba*, *N. kawakamii*, *N. knightiana*, *N. langsdorffii*, *N. linearis*, *N. longiflora*, *N. maritima*, *N. megalosiphon*, *N. miersii*, *N. noctiflora*, *N. nudicaulis*, *N. obtusifolia*, *N. occidentalis*, *N. occidentalis* subsp. *hesperis*, *N. otophora*, *N. paniculata*, *N. pauciflora*, *N. petunioides*, *N. plumbaginifolia*, *N. quadrivalvis*, *N. raimondii*, *N. repanda*, *N. rosulata*, *N. rosulata* subsp. *ingulba*, *N. rotundifolia*, *N. setchellii*, *N. simulans*, *N. solanifolia*, *N. spagazzinii*, *N. stocktonii*, *N. suaveolens*, *N. sylvestris*, *N. thyrsoiflora*, *N. tomentosa*, *N. tomentosiformis*, *N. trigonophylla*, *N. umbratica*, *N. undulata*, *N. velutina*, *N. wigandioides*, and *N. x sanderae*. In a highly preferred embodiment, the plant is a tobacco plant – such as a plant of the genus *Nicotiana* or the species *Nicotiana tabacum*.
- 35 — The use of tobacco cultivars and elite tobacco cultivars is also contemplated herein. The transgenic, non-naturally occurring or mutant plant may therefore be a tobacco variety or elite tobacco cultivar that comprises one or more transgenes, or one or more genetic mutations or a

combination thereof. The genetic mutation(s) (for example, one or more polymorphisms) can be mutations that do not exist naturally in the individual tobacco variety or tobacco cultivar (for example, elite tobacco cultivar) or can be genetic mutation(s) that do occur naturally provided that the mutation does not occur naturally in the individual tobacco variety or tobacco cultivar

5 (for example, elite tobacco cultivar). Particularly useful *Nicotiana tabacum* varieties include Burley type, dark type, flue-cured type, and Oriental type tobaccos. Non-limiting examples of varieties or cultivars are: BD 64, CC 101, CC 200, CC 27, CC 301, CC 400, CC 500, CC 600, CC 700, CC 800, CC 900, Coker 176, Coker 319, Coker 371 Gold, Coker 48, CD 263, DF911, DT 538 LC Galpao tobacco, GL 26H, GL 350, GL 600, GL 737, GL 939, GL 973, HB 04P, HB

10 04P LC, HB3307PLC, Hybrid 403LC, Hybrid 404LC, Hybrid 501 LC, K 149, K 326, K 346, K 358, K394, K 399, K 730, KDH 959, KT 200, KT204LC, KY10, KY14, KY 160, KY 17, KY 171, KY 907, KY907LC, KTY14xL8 LC, Little Crittenden, McNair 373, McNair 944, msKY 14xL8, Narrow Leaf Madole, Narrow Leaf Madole LC, NBH 98, N-126, N-777LC, N-7371LC, NC 100, NC 102, NC 2000, NC 291, NC 297, NC 299, NC 3, NC 4, NC 5, NC 6, NC7, NC 606, NC 71,

15 NC 72, NC 810, NC BH 129, NC 2002, Neal Smith Madole, OXFORD 207, PD 7302 LC, PD 7309 LC, PD 7312 LC, 'Perique' tobacco, PVH03, PVH09, PVH19, PVH50, PVH51, R 610, R 630, R 7-11, R 7-12, RG 17, RG 81, RG H51, RGH 4, RGH 51, RS 1410, Speight 168, Speight 172, Speight 179, Speight 210, Speight 220, Speight 225, Speight 227, Speight 234, Speight G-28, Speight G-70, Speight H-6, Speight H20, Speight NF3, TI 1406, TI 1269, TN 86, TN86LC,

20 TN 90, TN 97, TN97LC, TN D94, TN D950, TR (Tom Rosson) Madole, VA 309, VA359, AA 37-1, B 13P, Xanthi (Mitchell-Mor), Bel-W3, 79-615, Samsun Holmes NN, KTRDC number 2 Hybrid 49, Burley 21, KY 8959, KY 9, MD 609, PG 01, PG 04, PO1, PO2, PO3, RG 11, RG 8, VA 509, AS44, Banket A1, Basma Drama B84/31, Basma I Zichna ZP4/B, Basma Xanthi BX 2A, Batek, Besuki Jember, C104, Coker 347, Criollo Misionero,

25 Delcrest, Djebel 81, DVH 405, Galpão Comum, HB04P, Hicks Broadleaf, Kabakulak Ellassona, Kutsage E1, LA BU 21, NC 2326, NC 297, PVH 2110, Red Russian, Samsun, Saplak, Simmaba, Talgar 28, Wislica, Yayaldag, Prilep HC-72, Prilep P23, Prilep PB 156/1, Prilep P12-2/1, Yaka JK-48, Yaka JB 125/3, TI-1068, KDH-960, TI-1070, TW136, Basma, TKF 4028, L8, TKF 2002, GR141, Basma xanthi, GR149,

30 GR153, Petit Havana. Low converter subvarieties of the above, even if not specifically identified herein, are also contemplated. Cultivars and elite cultivars are preferred in certain embodiments. The species *Nicotiana tabacum* can be used in certain other embodiments. *Nicotiana tabacum* Burley type can be used.

Embodiments are also directed to compositions and methods for producing mutant, non-naturally occurring or transgenic plants that have been modified to modulate isopropylmalate synthase expression or activity which can result in plants or plant components with modulated levels of sucrose esters as compared to a control. Modulating the levels of sucrose esters in

plants may be used to generate plants in which the flavour profile of the aerosol generated therefrom is changed. The present invention may therefore provide new blending opportunities to generate desirable tobacco flavours. Thus, by way of example, the sucrose ester content of the plant may be increased by increasing the expression or activity of isopropylmalate synthase which may result in a plant capable of releasing elevated levels of flavour molecules that are more similar to oriental tobacco. Thus, by way of further example, the sucrose ester content of the plant may be decreased by decreasing the expression or activity of isopropylmalate synthase which may result in a plant capable of releasing lower levels of flavour molecules, that are more similar to flue cured tobacco.

- 10 In a further aspect there is provided a method for blending tobacco comprising the step of, replacing or adding to the one or more types of tobacco in a tobacco blend, tobacco (for example, tobacco leaves or a composition derived or derivable from tobacco leaves) derived from a transgenic, mutant or non-naturally occurring plant in which the sucrose ester content of the plant has been modulated (for example, increased or decreased) by modulating (preferably, increasing) the expression or activity of isopropylmalate synthase as described herein. In another aspect there is provided a method for blending tobacco comprising the step of replacing, reducing or omitting the Oriental type tobacco in a tobacco blend and adding tobacco (for example, tobacco leaves or a composition derived or derivable from tobacco leaves) derived from a transgenic, mutant or non-naturally occurring plant in which the sucrose ester content of the plant has been modulated (preferably, increased) by modulating (preferably, increasing) the expression or activity of isopropylmalate synthase as described herein. The replacement, reduction or addition of tobacco – such as Oriental tobacco – may correspond to about 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% or 100% of the tobacco that is replaced, reduced or added.
- 25 A further aspect relates to a method for modulating the flavour of tobacco or a tobacco product comprising, adding to the tobacco (for example, tobacco leaves or a composition derived or derivable from tobacco leaves) or the tobacco product, tobacco derived or derivable from a transgenic, mutant or non-naturally occurring plant in which the sucrose ester content of the plant has been modulated as described herein.
- 30 A further aspect relates to the use of agent that modulates the expression or activity of isopropylmalate synthase for modulating the flavour of tobacco or a tobacco product. Modulating the levels of sucrose esters in plants may also be used to generate plants in which the pest resistance of the plant is altered. The pest resistance of the plant may be increased by increasing the levels of sucrose ester production as described herein. The pest resistance of the plant may be decreased by decreasing the levels of sucrose ester production as described herein.
- 35

Compositions of the sucrose esters that are produced in the plants described herein may be extracted and optionally purified therefrom for various further uses – such as in drugs, food additives, smoking flavourants and as components of organic pesticides and the like. In one embodiment, the sucrose esters are extracted from leaves or the surface of leaves. Accordingly, products – such as drugs, food additives, smoking flavourants and organic pesticides comprising the sucrose esters - are a further aspect of this disclosure.

Sucrose esters according to the present disclosure have the general structure shown in Figure 5, wherein R1, R2, R3, R4 and R5 are either hydrogen atoms or acyl chains. The acyl chains that can be esterified in sucrose esters are set forth in Figure 6. Some examples of specific sucrose esters are set forth in Table 1 and Figure 2.

The total number of carbons in the acyl chains at R1, R2 and R4 in each of the one or more sucrose esters can be, for example, 12, 13, 14, 15, 16, 17 or 18; suitably, 14, 15, 16, 17 or 18; suitably, 16, 17 or 18; or suitably 14 or 15. Provided that this total number of carbons is present in each of the sucrose esters then the constituent acyl chains may be selected from any of those set forth in Figure 6. Suitably, the sucrose ester has an acetyl at R3 and an acetyl or hydrogen at R5. Suitably, the sucrose ester has an acetyl at R3 and a hydrogen at R5. Suitably, one or more, two or more or three of more of the acyl chains will be beta-methylvaleryl, especially when a 6 carbon acyl chain is present.

In one embodiment, the sucrose ester is a sucrose ester that has an acetyl or a hydrogen at R3 and R5 and an acyl chain with 3, 4, 5 or 6 carbons at any of positions R1, R2 and R4. Suitably, the sucrose ester has an acetyl at R3 and an acetyl or hydrogen at R5. Suitably, the sucrose ester has an acetyl at R3 and a hydrogen at R5. In another embodiment, the sucrose ester has an acetyl or a hydrogen at R3 and R5 and an acyl chain with 4, 5 or 6 carbons at any of positions R1, R2 and R4. Suitably, the sucrose ester has an acetyl at R3 and an acetyl or hydrogen at R5. Suitably, the sucrose ester has an acetyl at R3 and a hydrogen at R5. In another embodiment, the sucrose ester has an acetyl or a hydrogen at R3 and R5 and an acyl chain with 5 or 6 carbons at any of positions R1, R2 and R4. Suitably, the sucrose ester has an acetyl at R3 and an acetyl or hydrogen at R5. Suitably, the sucrose ester has an acetyl at R3 and a hydrogen at R5. In another embodiment, the sucrose ester has an acetyl or a hydrogen at R3 and R5 and an acyl chain with 6 carbons at positions R1, R2 and R4. Suitably, the sucrose ester has an acetyl at R3 and an acetyl or hydrogen at R5. Suitably, the sucrose ester has an acetyl at R3 and a hydrogen at R5.

In one embodiment, the sucrose ester has an acetyl or a hydrogen at R3 and R5 and an acyl chain with 3, 4, 5 or 6 carbons at position R1, an acyl chain with 3, 4, 5 or 6 carbons at position R2 and an acyl chain with 3, 4, 5 or 6 carbons at position R4. Suitably, the sucrose ester has an acetyl at R3 and an acetyl or hydrogen at R5. Suitably, the sucrose ester has an acetyl at R3 and a hydrogen at R5. In another embodiment, the sucrose ester has an acetyl or a

hydrogen at R3 and R5 and an acyl chain with 4, 5 or 6 carbons at position R1, an acyl chain with 4, 5 or 6 carbons at position R2 and an acyl chain with 4, 5 or 6 carbons at position R4. Suitably, the sucrose ester has an acetyl at R3 and an acetyl or hydrogen at R5. Suitably, the sucrose ester has an acetyl at R3 and a hydrogen at R5. In another embodiment, the sucrose ester has an acetyl or a hydrogen at R3 and R5 and an acyl chain with 5 or 6 carbons at position R1, an acyl chain with 5 or 6 carbons at position R2 and an acyl chain with 5 or 6 carbons at position R4. Suitably, the sucrose ester has an acetyl at R3 and an acetyl or hydrogen at R5. Suitably, the sucrose ester has an acetyl at R3 and a hydrogen at R5.

In one embodiment, the sucrose ester has an acetyl or a hydrogen at R3 and R5 and an acyl chain with 6 carbons at least at one of positions R1, R2 and R4. Suitably, the sucrose ester has an acetyl at R3 and an acetyl or hydrogen at R5. Suitably, the sucrose ester has an acetyl at R3 and a hydrogen at R5.

In one embodiment, the sucrose ester has an acetyl or a hydrogen at R3 and R5 and an acyl chain with 6 carbons at position R1, an acyl chain with 6 carbons at position R2 and an acyl chain with 6 carbons at position R4. Suitably, the sucrose ester has an acetyl at R3 and an acetyl or hydrogen at R5. Suitably, the sucrose ester has an acetyl at R3 and a hydrogen at R5. In one embodiment, the sucrose ester is a (branched chain) beta-methylvaleryl containing sucrose ester.

In one embodiment, the sucrose ester has the general structure shown in Figure 5 and wherein each of R1 to R5 is any one of the acyl chains selected from the group consisting of acetyl, butyryl, propionyl or an isomer thereof, isobutyryl, valeryl (pentanoyl), 2-methyl-butyl, isovaleryl, isopentenoyl, pentenoyl, hexanoyl or an isomer thereof 2-methylvaleryl, beta-methylvaleryl and 4-methylvaleryl. In another embodiment, the sucrose ester has the general structure shown in Figure 5 and wherein R3 is acetyl and each of R1, R2, R4 and R5 is any one of the acyl chains selected from the group consisting of acetyl, butyryl, propionyl or an isomer thereof, isobutyryl, valeryl (pentanoyl), 2-methyl-butyl, isovaleryl, isopentenoyl, pentenoyl, hexanoyl or an isomer thereof 2-methylvaleryl, beta-methylvaleryl and 4-methylvaleryl. In one embodiment, it is preferred that the acyl chain with 6 carbons is a branched chain. In another embodiment, it is preferred that the acyl chain with 6 carbons is beta-methylvaleryl.

In another embodiment, the sucrose ester has the general structure shown in Figure 5 and wherein R3 is hydrogen or acetyl, R5 is hydrogen or acetyl and each of R1, R2 and R4 is any one of the acyl chains selected from the group consisting of acetyl, butyryl, propionyl or an isomer thereof, isobutyryl, valeryl (pentanoyl) or an isomer thereof, 2-methyl-butyl, isovaleryl, isopentenoyl, pentenoyl, hexanoyl or an isomer thereof 2-methylvaleryl, beta-methylvaleryl and 4-methylvaleryl. Individual sucrose esters are disclosed in Table 1 and Figure 2.

In another embodiment, the sucrose ester has the general structure shown in Figure 5 and wherein R3 is acetyl, R5 is hydrogen or acetyl and each of R1, R2 and R4 is any one of the acyl

chains selected from the group consisting of acetyl, butyryl, propionyl or an isomer thereof, isobutyryl, valeryl (pentanoyl) or an isomer thereof, 2-methyl-butyryl, isovaleryl, isopentenyl, pentenyl, hexanoyl or an isomer thereof 2-methylvaleryl, beta-methylvaleryl and 4-methylvaleryl.

5 In another embodiment, the sucrose ester has the general structure shown in Figure 5 and wherein R3 is acetyl, R5 is hydrogen or acetyl; R1 is butyryl, propionyl or an isomer thereof, valeryl (pentanoyl) or an isomer thereof, or hexanoyl or an isomer thereof; R2 is propionyl or an isomer thereof, valeryl (pentanoyl) or an isomer thereof, or hexanoyl or an isomer thereof; and R4 is valeryl (pentanoyl) or an isomer thereof or hexanoyl or an isomer thereof.

10 In another embodiment, the sucrose ester has the general structure shown in Figure 5 and wherein R3 is acetyl, R5 is hydrogen or acetyl; R1 is propionyl or an isomer thereof, valeryl (pentanoyl) or an isomer thereof, or hexanoyl or an isomer thereof; R2 is valeryl (pentanoyl) or an isomer thereof, or hexanoyl or an isomer thereof; and R4 is hexanoyl or an isomer thereof.

In one embodiment, the sucrose ester that is modulated has the general structure shown in Figure 5 wherein R3=acetyl, R1=propionyl or an isomer thereof, R2= propionyl or an isomer thereof, R4=hexanoyl or an isomer thereof and R5 is hydrogen or acetyl; or a sucrose ester as shown in Figure 5 wherein R3=acetyl, R1= propionyl or an isomer thereof, R2=valeryl or an isomer thereof, R4= valeryl or an isomer thereof and R5 is hydrogen or acetyl; or a sucrose ester as shown in Figure 5 wherein R3=acetyl, R1=butyryl, R2=valeryl or an isomer thereof, R4=hexanoyl or an isomer thereof and R5 is hydrogen or acetyl (C2C14:0).

In one embodiment, the sucrose ester has the general structure shown in Figure 5 wherein R3=acetyl, R1=propionyl or an isomer thereof, R2=valeryl or an isomer thereof, R4=hexanoyl or an isomer thereof and R5 is hydrogen or acetyl; or a sucrose ester as shown in Figure 5 wherein R3=acetyl, R1=valeryl or an isomer thereof, R2=valeryl or an isomer thereof, R4=valeryl or an isomer thereof and R5 is hydrogen or acetyl; or a sucrose ester as shown in Figure 5 wherein R3=acetyl, R1=butyryl, R2=hexanoyl or an isomer thereof, R4=hexanoyl or an isomer thereof and R5 is hydrogen or acetyl (C12C15:0).

In one embodiment, the sucrose ester has the general structure shown in Figure 5 and wherein R3 is acetyl, R1 is valeryl or an isomer thereof, R2 is valeryl or an isomer thereof, R4 is hexanoyl or an isomer thereof and R5 is either a hydrogen atom or an acetyl moiety (C2C16:0).

In one embodiment, the sucrose ester that is modulated has the general structure shown in Figure 5 and wherein R3 is acetyl, R1 is propionyl or an isomer thereof, R2 is hexanoyl or an isomer thereof, R4 is hexanoyl or an isomer thereof and R5 is either a hydrogen atom or an acetyl moiety (C2C16:0).

35 In one embodiment, the sucrose ester has the general structure shown in Figure 5 and wherein R3 is acetyl, R1 is valeryl or an isomer thereof, R2 is hexanoyl or an isomer thereof, R4 is hexanoyl or an isomer thereof and R5 is either a hydrogen atom or an acetyl moiety (C2C17:0).

In one embodiment, the sucrose ester that is modulated has the general structure shown in Figure 5 and wherein R3 is acetyl, R1 is hexanoyl or an isomer thereof, R2 is hexanoyl or an isomer thereof, R4 is hexanoyl or an isomer thereof and R5 is either a hydrogen atom or an acetyl moiety (C2C18:0).

- 5 In another embodiment the sucrose ester has a molecular weight of from about 594 Daltons to 692 Daltons. In one embodiment, the sucrose ester has a molecular weight of from about 608 Daltons to 692 Daltons. In one embodiment, the sucrose ester has a molecular weight of from about 622 Daltons to 692 Daltons. In one embodiment, the sucrose ester has a molecular weight of from about 636 Daltons to 692 Daltons. In one embodiment, the sucrose ester has a molecular weight of from about 650 Daltons to 692 Daltons. In one embodiment, the sucrose ester has a molecular weight of from about 650 Daltons to 678 Daltons. In another embodiment the sucrose ester has a molecular weight of from about 594 Daltons to 678 Daltons. In one embodiment, the sucrose ester has a molecular weight of from about 608 Daltons to 678 Daltons. In one embodiment, the sucrose ester has a molecular weight of from about 622 Daltons to 678 Daltons. In one embodiment, the sucrose ester has a molecular weight of from about 636 Daltons to 678 Daltons. In one embodiment, the sucrose ester has a molecular weight of from about 650 Daltons to 678 Daltons. In one embodiment, the sucrose ester has a molecular weight of from about 664 Daltons to 678 Daltons.

- 10 A further aspect relates to a method for modulating the flavour of tobacco or a tobacco product comprising adding compositions of sucrose esters described herein thereto.

In another embodiment, the concentration, levels or production of the sucrose esters in a plant are increased. In another embodiment, the concentration, levels or production of the sucrose esters in a plant are decreased.

- 15 In one specific embodiment, compositions and methods for producing mutant, non-naturally occurring or transgenic plants that have been modified to increase isopropylmalate synthase expression or activity are provided which can result in plants or plant material in which the concentration, levels or production of sucrose esters are increased. Specifically, the levels or the production of the sucrose esters described herein are increased or induced.

- 20 In another specific embodiment, compositions and methods for producing mutant, non-naturally occurring or transgenic plants that have been modified to reduce or inhibit isopropylmalate synthase expression or activity are provided which can result in plants or plant material in which the concentration, levels or production of sucrose esters are reduced. Specifically, the levels of the sucrose esters described herein are decreased or abolished.

- 25 Advantageously, the mutant, non-naturally occurring or transgenic plants that are obtained according to the methods described herein are similar or substantially the same in visual appearance to the control plants. In one embodiment, the stalk height of the mutant, non-naturally occurring or transgenic plants is substantially the same as the control plants at, for

example, one, two or three or more months after field transplant or 10, 20, 30 or 36 or more days after topping. For example, the stalk height of the mutant, non-naturally occurring or transgenic plants is not less than the stalk height of the control plants. In another embodiment, the chlorophyll content of the mutant, non-naturally occurring or transgenic plants is substantially the same as the control plants. In another embodiment, the stalk height of the mutant, non-naturally occurring or transgenic plants is substantially the same as the control plants and the chlorophyll content of the mutant, non-naturally occurring or transgenic plants is substantially the same as the control plants. In other embodiments, the size or form or number or colouration of the leaves of the mutant, non-naturally occurring or transgenic plants is substantially the same as the control plants. Suitably, the plant is a tobacco plant.

In another aspect, there is provided a method for modulating the sucrose ester content – such as the beta-methylvaleryl sucrose ester content - in at least a part of a plant, comprising the steps of: (i) modulating the expression or activity of isopropylmalate synthase in the plant, preferably, wherein the isopropylmalate synthase comprises the polynucleotide sequence described herein or the polypeptide sequence described herein; (ii) optionally measuring the sucrose ester content in at least a part of the mutant, non-naturally occurring or transgenic plant obtained in step (i); and (iii) identifying a mutant, non-naturally occurring or transgenic plant in which the sucrose ester content therein has been modulated in comparison to a control plant. Suitably, the visual appearance of said mutant, non-naturally occurring or transgenic plant is substantially the same as the control plant. Suitably, the plant is a tobacco plant.

In another aspect, there is provided a method for increasing the sucrose ester content – such as the beta-methylvaleryl sucrose ester content - in at least a part of a plant (for example, the leaves), comprising the steps of: (i) increasing the expression or activity of isopropylmalate synthase in the plant, preferably, wherein the isopropylmalate synthase comprises the polynucleotide sequence described herein or the polypeptide sequence described herein; (ii) optionally measuring the sucrose ester content in at least a part of the mutant, non-naturally occurring or transgenic plant obtained in step (i); and (iii) identifying a mutant, non-naturally occurring or transgenic plant in which the sucrose ester content therein has been increased in comparison to a control plant.

In another aspect, there is provided a method for decreasing the sucrose ester content – such as the beta-methylvaleryl sucrose ester content - in at least a part of a plant, comprising the steps of: (i) reducing the expression or activity of isopropylmalate synthase in the plant, preferably, wherein the isopropylmalate synthase comprises the polynucleotide sequence described herein or the polypeptide sequence described herein; (ii) optionally measuring the sucrose ester content in at least a part of the mutant, non-naturally occurring or transgenic plant obtained in step (i); and (iii) identifying a mutant, non-naturally occurring or transgenic plant in which the sucrose ester content therein has been decreased in comparison to a control plant.

Advantageously, the methods described herein can also be used to generate plants with oriental traits by increasing the expression or activity of isopropylmalate synthase according to this embodiment, a non-oriental plant variety – such as a flue cured tobacco variety plant – can be engineered such that the expression or activity of isopropylmalate synthase is increased therein. This may impart oriental traits into plants which are advantageous since this may change (for example, increase or enhance) the oriental flavour thereof, thereby producing new aromatic plant lines, varieties or hybrids.

The methods described herein can also be used to generate plants in which the oriental flavour is reduced, inhibited or abolished by decreasing the expression or activity of isopropylmalate synthase. Thus, according to this embodiment, an oriental plant variety can be engineered such that the expression or activity of isopropylmalate synthase is decreased thereby producing a reduction or abolishment of oriental traits into plants can be advantageous since this may change or reduce the oriental flavour thereof, thereby producing new aromatic plant varieties or hybrids.

The increase in expression of isopropylmalate synthase as compared to the control may be from about 5 % to about 100 % or more, or an increase of at least 10 %, at least 20 %, at least 25 %, at least 30 %, at least 40 %, at least 50 %, at least 60 %, at least 70 %, at least 80 %, at least 90 %, at least 95 %, at least 98 %, 100 %, 150 %, or 200 % which includes an increase in transcriptional activity or protein expression or both.

The increase in the activity of isopropylmalate synthase as compared to a control plant may be from about 5 % to about 100 % or more, or an increase of at least 10 %, at least 20 %, at least 25 %, at least 30 %, at least 40 %, at least 50 %, at least 60 %, at least 70 %, at least 80 %, at least 90 %, at least 95 %, at least 98 %, 100 %, 150 %, or 200 % or more.

The reduction in expression of isopropylmalate synthase as compared to the control may be from about 5 % to about 100 %, or a reduction of at least 10 %, at least 20 %, at least 30 %, at least 40 %, at least 50 %, at least 60 %, at least 70 %, at least 75 %, at least 80 %, at least 90 %, at least 95 %, at least 98 %, or 100 %, which includes a reduction in transcriptional activity or protein expression or both.

The reduction in activity of isopropylmalate synthase as compared to a control plant may be from about 5 % to about 100 %, or a reduction of at least 10 %, at least 20 %, at least 30 %, at least 40 %, at least 50 %, at least 60 %, at least 70 %, at least 75 %, at least 80 %, at least 90 %, at least 95 %, at least 98 %, or 100 % or more.

The decrease in sucrose ester content as compared to a control plant may be from about 5 % to about 100 %, or a decrease of at least 10 %, at least 20 %, at least 25 %, at least 30 %, at least 40 %, at least 50 %, at least 60 %, at least 70 %, at least 75 %, at least 80 %, at least 90 %, at least 95 %, at least 98 %, or up to 100 %.

- 5 Polynucleotides and recombinant constructs described herein can be used to modulate the expression of isopropylmalate synthase described herein in a plant species of interest, suitably tobacco.

- One aspect is a seed of a mutant plant, a non-naturally occurring plant, a hybrid plant, a transgenic plant of the disclosure. Preferably, the seed is a tobacco seed. A further aspect is  
10 pollen or an ovule of a mutant plant, a non-naturally occurring plant, a hybrid plant, a transgenic plant of the disclosure. In addition, the present invention provides a mutant plant, a non-naturally occurring plant, a hybrid plant, a transgenic plant as described which further comprises a nucleic acid conferring male sterility.

- The disclosure also provides a tissue culture of regenerable cells of the mutant plant, non-  
15 naturally occurring plant, hybrid plant, or transgenic plant or a part thereof, of the present invention, which culture regenerates plants capable of expressing all the morphological and physiological characteristics of the parent. The regenerable cells include but are not limited to cells from leaves, pollen, embryos, cotyledons, hypocotyls, roots, root tips, anthers, flowers and a part thereof, ovules, shoots, stems, stalks, pith and capsules or callus or protoplasts derived  
20 therefrom.

- According to the disclosure, a plant carrying a mutant isopropylmalate synthase allele can be used in a plant breeding program to create useful lines, varieties and hybrids. In particular, the mutant isopropylmalate synthase allele may be introgressed into the commercially important varieties described above. Thus, methods for breeding plants are provided, that comprise  
25 crossing, breeding or introgressing a mutant plant, a non-naturally occurring plant or a transgenic plant as described herein with another plant – such as a plant with a different genetic identity, a different genetic background, a different variety, a different line or a different hybrid. One method comprises crossing, breeding or introgressing a mutant plant, a non-naturally occurring plant or a transgenic plant as described herein with a cultivar or elite cultivar  
30 or a plant of the genus *Nicotiana* – such as *Nicotiana tabacum* (for example, Burley type). The method may further comprise crossing the progeny plant with another plant, and optionally repeating the crossing until a progeny with the desirable genetic traits or genetic background is obtained. The method may further comprise the step of identifying a plant in which the expression or activity of isopropylmalate synthase is modulated. The method may further  
35 comprise the step of selecting a plant in which the expression or activity of isopropylmalate synthase is modulated as compared to a control plant. One purpose served by such breeding methods is to introduce a desirable genetic trait into other varieties, breeding lines, hybrids or

cultivars, particularly those that are of commercial interest. Another purpose is to facilitate stacking of genetic modifications of different genes in a single plant variety, lines, hybrids or cultivars. Intraspecific as well as interspecific matings are contemplated. The progeny plants that arise from such crosses, also referred to as breeding lines, are examples of non-naturally occurring plants of the disclosure.

In one embodiment, a method is provided for producing a non-naturally occurring plant comprising: (a) crossing a mutant or transgenic plant with a second plant to yield progeny tobacco seed; (b) growing the progeny tobacco seed, under plant growth conditions, to yield the non-naturally occurring plant. The method may further comprise the optional step of identifying a plant in which the expression or activity of isopropylmalate synthase is modulated. The method may further comprise the optional step of selecting a plant in which the expression or activity of isopropylmalate synthase is modulated as compared to a control plant. The method may further comprise: (c) crossing the previous generation of non-naturally occurring plant with itself or another plant to yield progeny tobacco seed; (d) growing the progeny tobacco seed of step (c) under plant growth conditions, to yield additional non-naturally occurring plants; and (e) repeating the crossing and growing steps of (c) and (d) multiple times to generate further generations of non-naturally occurring plants. The method may optionally comprise prior to step (a), a step of providing a parent plant which comprises a genetic identity that is characterized and that is not identical to the mutant or transgenic plant. In some embodiments, depending on the breeding program, the crossing and growing steps are repeated from 0 to 2 times, from 0 to 3 times, from 0 to 4 times, 0 to 5 times, from 0 to 6 times, from 0 to 7 times, from 0 to 8 times, from 0 to 9 times or from 0 to 10 times, in order to generate generations of non-naturally occurring plants. Backcrossing is an example of such a method wherein a progeny is crossed with one of its parents or another plant genetically similar to its parent, in order to obtain a progeny plant in the next generation that has a genetic identity which is closer to that of one of the parents. Techniques for plant breeding, particularly tobacco plant breeding, are well known and can be used in the methods described. The disclosure further provides non-naturally occurring plants produced, obtained or obtainable by these methods.

In some embodiments of methods described herein, lines resulting from breeding and screening for variant isopropylmalate synthase genes are evaluated in the field using standard field procedures. Control genotypes including the original unmutagenised parent are included and entries are arranged in the field in a randomised complete block design or other appropriate field design. For tobacco, standard agronomic practices are used, for example, the tobacco is harvested, weighed, and sampled for chemical and other common testing before and during curing. Statistical analyses of the data are performed to confirm the similarity of the selected lines to the parental line. Cytogenetic analyses of the selected plants are optionally performed to confirm the chromosome complement and chromosome pairing relationships.

In one embodiment, the plants that are obtained or obtainable by the breeding, crossing or introgressing methods described above will not produce *cis*-abienol, will not substantially produce *cis*-abienol or will not produce detectable levels of *cis*-abienol. In another embodiment, the sucrose ester compositions and the sucrose ester compositions obtained or obtainable by the methods described herein will not comprise or contain *cis*-abienol, will not comprise or contain substantial levels of *cis*-abienol or will not comprise or contain detectable levels of *cis*-abienol. Thus, the *cis*-abienol trait may be present or it may be absent.

DNA fingerprinting, single nucleotide polymorphism, microsatellite markers, or similar technologies may be used in a marker-assisted selection (MAS) breeding program to transfer or breed mutant alleles of the isopropylmalate synthase gene into other tobaccos, as described herein. For example, a breeder can create segregating populations from hybridizations of a genotype containing a mutant allele with an agronomically desirable genotype. Plants in the F<sub>2</sub> or backcross generations can be screened using a marker developed from an isopropylmalate synthase genomic sequence or a fragment thereof, using one of the techniques listed herein. Plants identified as possessing the mutant allele can be backcrossed or self-pollinated to create a second population to be screened. Depending on the expected inheritance pattern or the MAS technology used, it may be necessary to self-pollinate the selected plants before each cycle of backcrossing to aid identification of the desired individual plants. Backcrossing or other breeding procedure can be repeated until the desired phenotype of the recurrent parent is recovered.

According to the disclosure, in a breeding program, successful crosses yield F<sub>1</sub> plants that are fertile. Selected F<sub>1</sub> plants can be crossed with one of the parents, and the first backcross generation plants are self-pollinated to produce a population that is again screened for variant isopropylmalate synthase gene expression (for example, the null version of the isopropylmalate synthase gene). The process of backcrossing, self-pollination, and screening is repeated, for example, at least 4 times until the final screening produces a plant that is fertile and reasonably similar to the recurrent parent. This plant, if desired, is self-pollinated and the progeny are subsequently screened again to confirm that the plant exhibits variant isopropylmalate synthase gene expression. In some embodiments, a plant population in the F<sub>2</sub> generation is screened for variant isopropylmalate synthase gene expression, for example, a plant is identified that fails to express isopropylmalate synthase due to the absence of a isopropylmalate synthase gene according to standard methods, for example, by using a PCR method with primers based upon the nucleotide sequence information for isopropylmalate synthase described herein.

Hybrid tobacco varieties can be produced by preventing self-pollination of female parent plants (that is, seed parents) of a first variety, permitting pollen from male parent plants of a second variety to fertilize the female parent plants, and allowing F<sub>1</sub> hybrid seeds to form on the female plants. Self-pollination of female plants can be prevented by emasculating the flowers at an

early stage of flower development. Alternatively, pollen formation can be prevented on the female parent plants using a form of male sterility. For example, male sterility can be produced by cytoplasmic male sterility (CMS), or transgenic male sterility wherein a transgene inhibits microsporogenesis and/or pollen formation, or self-incompatibility. Female parent plants containing CMS are particularly useful. In embodiments in which the female parent plants are CMS, pollen is harvested from male fertile plants and applied manually to the stigmas of CMS female parent plants, and the resulting F1 seed is harvested.

Varieties and lines described herein can be used to form single-cross tobacco F1 hybrids. In such embodiments, the plants of the parent varieties can be grown as substantially homogeneous adjoining populations to facilitate natural cross-pollination from the male parent plants to the female parent plants. The F1 seed formed on the female parent plants is selectively harvested by conventional means. One also can grow the two parent plant varieties in bulk and harvest a blend of F1 hybrid seed formed on the female parent and seed formed upon the male parent as the result of self-pollination. Alternatively, three-way crosses can be carried out wherein a single-cross F1 hybrid is used as a female parent and is crossed with a different male parent. As another alternative, double-cross hybrids can be created wherein the F1 progeny of two different single-crosses are themselves crossed.

A population of mutant, non-naturally occurring or transgenic plants can be screened or selected for those members of the population that have a desired trait or phenotype. For example, a population of progeny of a single transformation event can be screened for those plants having a desired level of expression of a polypeptide or polynucleotide. Physical and biochemical methods can be used to identify expression levels. These include Southern analysis or PCR amplification for detection of a polynucleotide; Northern blots, S1 RNase protection, primer-extension, or RT-PCR amplification for detecting RNA transcripts; enzymatic assays for detecting enzyme or ribozyme activity of polypeptides and polynucleotides; and protein gel electrophoresis, Western blots, immunoprecipitation, and enzyme-linked immunoassays to detect polypeptides. Other techniques such as in situ hybridization, enzyme staining, and immunostaining also can be used to detect the presence or expression of polypeptides or polynucleotides.

Mutant, non-naturally occurring or transgenic plant cells and plants are described herein comprising one or more recombinant polynucleotides – such as one or more isolated polynucleotides described herein, one or more polynucleotide constructs, one or more double-stranded RNAs, one or more conjugates or one or more vectors/expression vectors.

Gene expression can be evaluated using methods including, for example, RT-PCR, Northern blots, RNase protection, primer extensions, Western blots, protein gel electrophoresis, immunoprecipitation, enzyme-linked immunoassays, chip assays, and mass spectrometry. It should be noted that if a polypeptide is expressed under the control of a tissue-preferential or

broadly expressing promoter, expression can be evaluated in the entire plant or in a selected tissue. Similarly, if a polypeptide is expressed at a particular time, for example, at a particular time in development or upon induction, expression can be evaluated selectively at a desired time period.

- 5 Without limitation, the plants described herein may be modified for other purposes either before or after the expression or activity has been modulated. One or more of the following genetic modifications can be present in the mutant, non-naturally occurring or transgenic plants of the disclosure. In one embodiment, one or more genes that are involved in heavy metal uptake or heavy metal transport is modified resulting in plants or parts of plants (such as leaves) having a lower heavy metal content than control plants or parts thereof without the modification(s). Non-limiting examples include genes in the family of cation diffusion facilitators (CDF), the family of Zrt-, Irt-like proteins (ZIP), the family of cation exchangers (CAX), the family of copper transporters (COPT), the family of heavy-metal P-type ATPases (HMAs, as described in WO2009074325), the family of homologs of natural resistance-associated macrophage proteins (NRAMP), and the family of ATP-binding cassette (ABC) transporters, which participate in transport of heavy metals, such as cadmium. The term heavy metal as used herein includes transition metals. In another embodiment, one or more genes that are involved in the conversion of nitrogenous metabolic intermediates is modified resulting in plants or parts of plants (such as leaves) that when heated, produces lower levels of at least one tobacco-specific nitrosamine (for example, 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone, N-nitrosornicotine, N-nitrosoanatabine, and N-nitrosoanabasine) than control plants or parts thereof. Non-limiting examples of genes that can be modified include genes encoding a nicotine demethylase, such as CYP82E4, CYP82E5 and CYP82E10 which participate in the conversion of nicotine to nor nicotine and are described in WO2006091194, WO2008070274, WO2009064771 and PCT/US2011/021088.

- Examples of other modifications include herbicide tolerance, for example, glyphosate is an active ingredient of many broad spectrum herbicides. Glyphosate resistant transgenic plants have been developed by transferring the *aroA* gene (a glyphosate EPSP synthetase from *Salmonella typhimurium* and *E.coli*). Sulphonylurea resistant plants have been produced by transforming the mutant ALS (acetolactate synthetase) gene from *Arabidopsis*. OB protein of photosystem II from mutant *Amaranthus hybridus* has been transferred in to plants to produce atrazine resistant transgenic plants; and bromoxynil resistant transgenic plants have been produced by incorporating the *bxn* gene from the bacterium *Klebsiella pneumoniae*. Another exemplary modification results in plants that are resistant to insects. *Bacillus thuringiensis* (Bt) toxins can provide an effective way of delaying the emergence of Bt-resistant pests, as recently illustrated in broccoli where pyramided *cry1Ac* and *cry1C* Bt genes controlled diamondback moths resistant to either single protein and significantly delayed the evolution of resistant

insects. Another exemplary modification results in plants that are resistant to diseases caused by pathogens (for example, viruses, bacteria, fungi). Plants expressing the *Xa21* gene (resistance to bacterial blight) with plants expressing both a Bt fusion gene and a chitinase gene (resistance to yellow stem borer and tolerance to sheath) have been engineered. Another

5 exemplary modification results in altered reproductive capability, such as male sterility. Another exemplary modification results in plants that are tolerant to abiotic stress (for example, drought, temperature, salinity), and tolerant transgenic plants have been produced by transferring acyl glycerol phosphate enzyme from *Arabidopsis*; genes coding mannitol dehydrogenase and sorbitol dehydrogenase which are involved in synthesis of mannitol and sorbitol improve

10 drought resistance. Another exemplary modification results in plants that produce proteins which have favourable immunogenic properties for use in humans. For example, plants capable of producing proteins which substantially lack alpha-1,3-linked fucose residues, beta-1,2-linked xylose residues, or both, in its N-glycan may be of use. Other exemplary modifications can result in plants with improved storage proteins and oils, plants with enhanced photosynthetic

15 efficiency, plants with prolonged shelf life, plants with enhanced carbohydrate content, and plants resistant to fungi; plants encoding an enzyme involved in the biosynthesis of alkaloids. Transgenic plants in which the expression of S-adenosyl-L-methionine (SAM) and/or cystathionine gamma-synthase (CGS) has been modulated are also contemplated.

One or more such traits may be introgressed into the mutant, non-naturally occurring or transgenic plants from another plant or tobacco cultivar or may be directly transformed into it.

20 The introgression of the trait(s) into the mutant, non-naturally occurring or transgenic plants may be achieved by any method of plant breeding known in the art, for example, pedigree breeding, backcrossing, doubled-haploid breeding, and the like (see, Wernsman, E. A, and Ruffy, R. C. 1987. Chapter Seventeen. Tobacco. Pages 669-698 In: *Cultivar Development. Crop Species*.

25 W. H. Fehr (ed.), MacMillan Publishing Co, Inc., New York, N.Y 761 pp.). Molecular biology-based techniques described above, in particular RFLP and microsatellite markers, can be used in such backcrosses to identify the progenies having the highest degree of genetic identity with the recurrent parent. This permits one to accelerate the production of tobacco varieties having at least 90%, preferably at least 95%, more preferably at least 99% genetic identity with the

30 recurrent parent, yet more preferably genetically identical to the recurrent parent, and further comprising the trait(s) introgressed from the donor parent. Such determination of genetic identity can be based on molecular markers known in the art.

The last backcross generation can be selfed to give pure breeding progeny for the nucleic acid(s) being transferred. The resulting plants generally have essentially all of the morphological

35 and physiological characteristics of the mutant, non-naturally occurring or transgenic plants of the disclosure, in addition to the transferred trait(s) (for example, one or more single gene traits). The exact backcrossing protocol will depend on the trait being altered to determine an

appropriate testing protocol. Although backcrossing methods are simplified when the trait being transferred is a dominant allele, a recessive allele may also be transferred. In this instance, it may be necessary to introduce a test of the progeny to determine if the desired trait has been successfully transferred.

- 5 A plant or plant cell can be transformed by having the recombinant polynucleotide integrated into its genome to become stably transformed. Stably transformed cells typically retain the introduced polynucleotide with each cell division. A plant or plant cell may also be transiently transformed such that the recombinant polynucleotide is not integrated into its genome. Transiently transformed cells typically lose all or some portion of the introduced recombinant polynucleotide with each cell division such that the introduced recombinant polynucleotide cannot be detected in daughter cells after a sufficient number of cell divisions.

- 10 Conjugated moieties including macromolecular compounds such as proteins (for example, antibodies), fatty acid chains, sugar residues, glycoproteins, polymers (for example, polyethylene glycol), or combinations thereof are disclosed. An oligonucleotide may be conjugated to a moiety that increases cellular uptake of the oligonucleotide. Non-limiting examples of moieties include, but are not limited to, antibodies, polypeptides, lipid moieties such as a cholesterol moiety, cholic acid, a thioether, for example, hexyl-s-tritylthiol, a thiocholesterol, an aliphatic chain, for example, dodecandiol or undecyl residues, a phospholipid, a polyamine or a polyethylene glycol chain, an adamantane acetic acid, a palmityl moiety, an octadecylamine or a hexylamino-carbonyl-oxycholesterol moiety. The moiety may be a positively charged polymer - such as a positively charged peptide that is, for example, about 1 to 50 amino acid residues in length or polyalkylene oxide such as polyethylene glycol (PEG) or polypropylene glycol. Suitably the positively charged polymer, such as a polyalkylene oxide may be attached to the oligomer via a linker such as a releasable linker.

- 25 When polypeptide expression is being measured, detecting the amount of mRNA encoding a polypeptide in the cell can be quantified by, for example, PCR or Northern blot. Where a change in the amount of a polypeptide in the sample is being measured, detecting it by use of antibodies can be used to quantify the amount of polypeptide in the cell using known techniques. Alternatively the biological activity of the enzyme can be measured using methods that are well known in the art.

- 30 In another embodiment, antibodies that are immunoreactive with the polypeptides are provided herein. The polypeptides, fragments, variants, fusion polypeptides, and the like, as set forth herein, can be employed as "immunogens" in producing antibodies immunoreactive therewith. Such antibodies may specifically bind to the polypeptide via the antigen-binding sites of the antibody. Specifically binding antibodies are those that will specifically recognize and bind with a polypeptide, homologues, and variants, but not with other molecules. In one embodiment, the

antibodies are specific for polypeptides having an amino acid sequence as set forth herein and do not cross-react with other polypeptides.

More specifically, the polypeptides, fragment, variants, fusion polypeptides, and the like contain antigenic determinants or epitopes that elicit the formation of antibodies. These antigenic determinants or epitopes can be either linear or conformational (discontinuous). Linear epitopes are composed of a single section of amino acids of the polypeptide, while conformational or discontinuous epitopes are composed of amino acids sections from different regions of the polypeptide chain that are brought into close proximity upon polypeptide folding. Epitopes can be identified by any of the methods known in the art. Additionally, epitopes from the polypeptides can be used as research reagents, in assays, and to purify specific binding antibodies from substances such as polyclonal sera or supernatants from cultured hybridomas. Such epitopes or variants thereof can be produced using techniques known in the art such as solid-phase synthesis, chemical or enzymatic cleavage of a polypeptide, or using recombinant DNA technology.

Both polyclonal and monoclonal antibodies to the polypeptides can be prepared by conventional techniques. Hybridoma cell lines that produce monoclonal antibodies specific for the polypeptides are also contemplated herein. Such hybridomas can be produced and identified by conventional techniques. For the production of antibodies, various host animals may be immunized by injection with a polypeptide, fragment, variant, or mutants thereof. Such host animals may include, but are not limited to, rabbits, mice, and rats, to name a few. Various adjuvants may be used to increase the immunological response. Depending on the host species, such adjuvants include, but are not limited to, Freund's (complete and incomplete), mineral gels such as aluminium hydroxide, surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, dinitrophenol, and potentially useful human adjuvants such as BCG (bacille Calmette-Guerin) and *Corynebacterium parvum*. The monoclonal antibodies can be recovered by conventional techniques. Such monoclonal antibodies may be of any immunoglobulin class including IgG, IgM, IgE, IgA, IgD, and any subclass thereof.

The antibodies can also be used in assays to detect the presence of the polypeptides or fragments, either *in vitro* or *in vivo*. The antibodies also can be employed in purifying polypeptides or fragments by immunoaffinity chromatography.

Various embodiments provide mutant plants, non-naturally occurring plants or transgenic plants, as well as biomass in which the expression level of isopropylmalate synthase polynucleotide is modulated.

Parts of the such plants, particularly tobacco plants, and more particularly the leaf lamina and midrib of tobacco plants, can be incorporated into or used in making various consumable products including but not limited to aerosol forming materials, aerosol forming devices,

smoking articles, smokable articles, smokeless products, and tobacco products.. Examples of aerosol forming materials include but are not limited to tobacco compositions, tobaccos, tobacco extract, cut tobacco, cut filler, cured tobacco, expanded tobacco, homogenized tobacco, reconstituted tobacco, and pipe tobaccos. Smoking articles and smokable articles are types of aerosol forming devices. Examples of smoking articles or smokable articles include but are not limited to cigarettes, cigarillos, and cigars. Examples of smokeless products comprise chewing tobaccos, and snuffs. In certain aerosol forming devices, rather than combustion, a tobacco composition or another aerosol forming material is heated by one or more electrical heating elements to produce an aerosol. In another type of heated aerosol forming device, an aerosol is produced by the transfer of heat from a combustible fuel element or heat source to a physically separate aerosol forming material, which may be located within, around or downstream of the heat source. Smokeless tobacco products and various tobacco-containing aerosol forming materials may contain tobacco in any form, including as dried particles, shreds, granules, powders, or a slurry, deposited on, mixed in, surrounded by, or otherwise combined with other ingredients in any format, such as flakes, films, tabs, foams, or beads. As used herein, the term 'smoke' is used to describe a type of aerosol that is produced by smoking articles, such as cigarettes, or by combusting an aerosol forming material.

In one embodiment, there is also provided cured material from the mutant, transgenic and non-naturally occurring plants described herein. Processes of curing green tobacco leaves are known by those having skills in the art and include without limitation air-curing, fire-curing, flue-curing and sun-curing. The process of curing green tobacco leaves depends on the type of tobacco harvested. For example, Virginia flue (bright) tobacco is typically flue-cured, Burley and certain dark strains are usually air-cured, and pipe tobacco, chewing tobacco, and snuff are usually fire-cured.

In another embodiment, there is described tobacco products including tobacco-containing aerosol forming materials comprising leaves, preferably cured leaves, from the mutant tobacco plants, transgenic tobacco plants or non-naturally occurring tobacco plants described herein. The tobacco products described herein can be a blended tobacco product which may further comprise unmodified tobacco.

The % sucrose esters in these aerosol forming materials or tobacco compositions may be at least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, and 100%, 150%, or 200% or higher, when compared to aerosol forming materials or tobacco compositions derived from non-mutant, non-naturally occurring or non-transgenic counterpart plants.

The % sucrose esters in these aerosol forming materials or tobacco compositions may be at least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, and 100% lower, when compared to

aerosol forming materials or tobacco compositions derived from non-mutant, non-naturally occurring or non-transgenic counterpart plants.

The mutant, non-naturally occurring or transgenic plants may have other uses in, for example, agriculture. For example, mutant, non-naturally occurring or transgenic plants described herein  
5 can be used to make animal feed and human food products.

The disclosure also provides methods for producing seeds comprising cultivating the mutant plant, non-naturally occurring plant, or transgenic plant described herein, and collecting seeds from the cultivated plants. Seeds from plants described herein can be conditioned and bagged  
10 in packaging material by means known in the art to form an article of manufacture. Packaging material such as paper and cloth are well known in the art. A package of seed can have a label, for example, a tag or label secured to the packaging material, a label printed on the packaging material, or a label inserted within the package, that describes the nature of the seeds therein.

A further aspect relates to a method for producing a composition comprising sucrose esters comprising the steps of: (a) providing part of a mutant, non-naturally occurring or transgenic  
15 plant; biomass, seed or leaves; or the tobacco product as described herein; and (b) extracting the sucrose esters therefrom; and (c) optionally purifying the sucrose esters.

A further aspect relates to a method for producing beta-methylvaleric acid comprising the steps of: (a) providing part of a mutant, non-naturally occurring or transgenic plant; biomass, or leaves; or the aerosol forming material as described herein; (b) hydrolysing the material  
20 obtained in step (a) or an extract thereof; and (c) optionally collecting the beta-methylvaleric acid. Beta-methylvaleric acid can be used in the production of polyhydroxyalkanoates and the like which can give materials, including plastics, beneficial properties. The plastics that can be produced can be biodegradable and can therefore be used in the production of bioplastics. A bacterium – such as *E. coli* or *Pseudomonas* – can be engineered to express the  
25 isopropylmalate synthase described herein. Suitably, the bacterium that is used may also be engineered to also produce polyhydroxyalkanoates. Methods for the production of polyhydroxyalkanoates using microbial systems are described in, for example, US5,750,848 and US6,492,134. Thus, one aspect relates to a bacterium comprising isopropylmalate synthase and optionally one or more genes for producing a polyhydroxyalkanoate. A further aspect  
30 relates to a method for producing polyhydroxyalkanoate comprising the use of the bacterium.

In one embodiment, there is also provided cured material from the mutant, transgenic and non-naturally occurring plants described herein. For example, processes of curing green tobacco leaves are known by those having skills in the field and include without limitation air-curing, fire-curing, flue-curing and sun-curing. The process of curing green tobacco leaves depends on the  
35 type of tobacco harvested. For example, Virginia flue (bright) tobacco is typically flue-cured, Burley and certain dark strains are usually air-cured, and pipe tobacco, chewing tobacco, and snuff are usually fire-cured.

In another embodiment, there is described tobacco products including tobacco products comprising leaves, for example cured leaves, from the mutant, transgenic and non-naturally occurring plants described herein or which are produced by the methods described herein. The tobacco products described herein may further comprise unmodified tobacco.

5 In another embodiment, there is described tobacco products comprising plant material, preferably leaves – such as cured leaves, from the mutant, transgenic and non-naturally occurring plants described herein. For example, the plant material may be added to the inside or outside of the tobacco product and so upon burning a desirable aroma is released. The tobacco product according to this embodiment may even be an unmodified tobacco or a  
10 modified tobacco. The tobacco product according to this embodiment may even be derived from a mutant, transgenic or non-naturally occurring plant which has modifications in one or more genes other than the genes disclosed herein.

A further aspect relates to an (isolated) sucrose ester composition comprising, consisting or consisting essentially of one or more of the sucrose esters described herein. The sucrose ester  
15 may be a sucrose ester in which the total number of carbons in the acyl chains at R1, R2 and R4 is 14; a sucrose ester in which the total number of carbons in the acyl chains at R1, R2 and R4 is 15; a sucrose ester in which the total number of carbons in the acyl chains at R1, R2 and R4 is 16; a sucrose ester in which the total number of carbons in the acyl chains at R1, R2 and R4 is 17; and a sucrose ester in which the total number of carbons in the acyl chains at R1, R2  
20 and R4 is 18. Provided that this total number of carbons is present then the constituent acyl chains may be selected from any of those set forth in Figure 6. Suitably, the sucrose ester has an acetyl at R3 and an acetyl or hydrogen at R5. Suitably, the sucrose ester has an acetyl at R3 and a hydrogen at R5. Suitably, one or more, two or more or three or more of the acyl chains will be beta-methylvaleryl. Thus, by way of example, the sucrose ester can be a sucrose  
25 ester as shown in Figure 5 wherein R3=acetyl, R1=propionyl or an isomer thereof, R2=propionyl or an isomer thereof, R4=hexanoyl or an isomer thereof and R5 is hydrogen or acetyl; or a sucrose ester as shown in Figure 5 wherein R3=acetyl, R1=propionyl or an isomer thereof, R2=valeryl or an isomer thereof, R4=valeryl or an isomer thereof and R5 is hydrogen or acetyl; or a sucrose ester as shown in Figure 5 wherein R3=acetyl, R1=butyryl, R2=valeryl or an  
30 isomer thereof, R4=hexanoyl or an isomer thereof and R5 is hydrogen or acetyl (C2C14:0); a sucrose ester as shown in Figure 5 wherein R3=acetyl, R1=propionyl or an isomer thereof, R2=valeryl or an isomer thereof, R4=hexanoyl or an isomer thereof and R5 is hydrogen or acetyl; or a sucrose ester as shown in Figure 5 wherein R3=acetyl, R1=valeryl or an isomer thereof, R2=valeryl or an isomer thereof, R4=valeryl or an isomer thereof and R5 is hydrogen or  
35 acetyl; or a sucrose ester as shown in Figure 5 wherein R3=acetyl, R1=butyryl, R2=hexanoyl or an isomer thereof, R4=hexanoyl or an isomer thereof and R5 is hydrogen or acetyl (C2C15:0); a sucrose ester that has the general structure shown in Figure 5 and wherein R3 is acetyl, R1 is

valeryl or an isomer thereof, R2 is valeryl or an isomer thereof, R4 is hexanoyl or an isomer thereof and R5 is either a hydrogen atom or an acetyl moiety (C2C16:0); or wherein the sucrose ester has the general structure shown in Figure 5 and wherein R3 is acetyl, R1 is propionyl or an isomer thereof, R2 is hexanoyl or an isomer thereof, R4 is hexanoyl or an isomer thereof and R5 is either a hydrogen atom or an acetyl moiety (C2C16:0); a sucrose ester as shown in Figure 5 wherein R3 is acetyl, R1 is valeryl or an isomer thereof, R2 is hexanoyl or an isomer thereof, R4 is hexanoyl or an isomer thereof and R5 is either a hydrogen atom or an acetyl moiety (C2C17:0); and a sucrose ester as shown in Figure 5 wherein R3 is acetyl, R1 is hexanoyl or an isomer thereof, R2 is hexanoyl or an isomer thereof, R4 is hexanoyl or an isomer thereof and R5 is either a hydrogen atom or an acetyl moiety (C2C18:0).

Alternatively, the sucrose ester may be a sucrose ester in which the total number of carbons in the acyl chains at R1, R2 and R4 is 14; a sucrose ester in which the total number of carbons in the acyl chains at R1, R2 and R4 is 15; a sucrose ester in which the total number of carbons in the acyl chains at R1, R2 and R4 is 16; a sucrose ester in which the total number of carbons in the acyl chains at R1, R2 and R4 is 17; and a sucrose ester in which the total number of carbons in the acyl chains at R1, R2 and R4 is 18 is not present, is not substantially present or is not present at detectable levels (C2C18:0). Provided that this total number of carbons is present then the constituent acyl chains may be selected from any of those set forth in Figure 6.

Suitably, the sucrose ester has an acetyl at R3 and an acetyl or hydrogen at R5. Suitably, the sucrose ester has an acetyl at R3 and a hydrogen at R5. Suitably, one or more, two or more or three or more of the acyl chains will be beta-methylvaleryl. Thus, by way of example, the sucrose ester can be a sucrose ester as shown in Figure 5 wherein R3=acetyl, R1=propionyl or an isomer thereof, R2=propionyl or an isomer thereof, R4=hexanoyl or an isomer thereof and R5 is hydrogen or acetyl; or a sucrose ester as shown in Figure 5 wherein R3=acetyl, R1=propionyl or an isomer thereof, R2=valeryl or an isomer thereof, R4=valeryl or an isomer thereof and R5 is hydrogen or acetyl; or a sucrose ester as shown in Figure 5 wherein R3=acetyl, R1=butyryl, R2=valeryl or an isomer thereof, R4=hexanoyl or an isomer thereof and R5 is hydrogen or acetyl (C2C14:0); a sucrose ester as shown in Figure 5 wherein R3=acetyl, R1=propionyl or an isomer thereof, R2=valeryl or an isomer thereof, R4=hexanoyl or an isomer thereof and R5 is hydrogen or acetyl; or a sucrose ester as shown in Figure 5 wherein R3=acetyl, R1=valeryl or an isomer thereof, R2=valeryl or an isomer thereof, R4=valeryl or an isomer thereof and R5 is hydrogen or acetyl; or a sucrose ester as shown in Figure 5 wherein R3=acetyl, R1=butyryl, R2=hexanoyl or an isomer thereof, R4=hexanoyl or an isomer thereof and R5 is hydrogen or acetyl (C2C15:0); a sucrose ester that has the general structure shown in Figure 5 and wherein R3 is acetyl, R1 is valeryl or an isomer thereof, R2 is valeryl or an isomer thereof, R4 is hexanoyl or an isomer thereof and R5 is either a hydrogen atom or an acetyl moiety (C2C16:0); or wherein the sucrose ester has the general structure shown in Figure 5 and

wherein R3 is acetyl, R1 is propionyl or an isomer thereof, R2 is hexanoyl or an isomer thereof, R4 is hexanoyl or an isomer thereof and R5 is either a hydrogen atom or an acetyl moiety (C2C16:0); a sucrose ester as shown in Figure 5 wherein R3 is acetyl, R1 is valeryl or an isomer thereof, R2 is hexanoyl or an isomer thereof, R4 is hexanoyl or an isomer thereof and R5 is either a hydrogen atom or an acetyl moiety (C2C17:0); and wherein a sucrose ester as shown in Figure 5 wherein R3 is acetyl, R1 is hexanoyl or an isomer thereof, R2 is hexanoyl or an isomer thereof, R4 is hexanoyl or an isomer thereof and R5 is either a hydrogen atom or an acetyl moiety is not present, is not substantially present or is not present at detectable levels (C2C18:0).

Alternatively, the sucrose ester may be a sucrose ester in which the total number of carbons in the acyl chains at R1, R2 and R4 is 14; a sucrose ester in which the total number of carbons in the acyl chains at R1, R2 and R4 is 15; a sucrose ester in which the total number of carbons in the acyl chains at R1, R2 and R4 is 16; a sucrose ester in which the total number of carbons in the acyl chains at R1, R2 and R4 is 17 is not present, is not substantially present or is not present at detectable levels; and a sucrose ester in which the total number of carbons in the acyl chains at R1, R2 and R4 is 18 is not present, is not substantially present or is not present at detectable levels. Provided that this total number of carbons is present then the constituent acyl chains may be selected from any of those set forth in Figure 6. Suitably, the sucrose ester has an acetyl at R3 and an acetyl or hydrogen at R5. Suitably, the sucrose ester has an acetyl at R3 and a hydrogen at R5. Suitably, one or more, two or more or three or more of the acyl chains will be beta-methylvaleryl. Thus, by way of example, the sucrose ester can be a sucrose ester as shown in Figure 5 wherein R3=acetyl, R1=propionyl or an isomer thereof, R2=propionyl or an isomer thereof, R4=hexanoyl or an isomer thereof and R5 is hydrogen or acetyl; or a sucrose ester as shown in Figure 5 wherein R3=acetyl, R1=propionyl or an isomer thereof, R2=valeryl or an isomer thereof, R4=valeryl or an isomer thereof and R5 is hydrogen or acetyl; or a sucrose ester as shown in Figure 5 wherein R3=acetyl, R1=butyryl, R2=valeryl or an isomer thereof, R4=hexanoyl or an isomer thereof and R5 is hydrogen or acetyl (C2C14:0); a sucrose ester as shown in Figure 5 wherein R3=acetyl, R1=propionyl or an isomer thereof, R2=valeryl or an isomer thereof, R4=hexanoyl or an isomer thereof and R5 is hydrogen or acetyl; or a sucrose ester as shown in Figure 5 wherein R3=acetyl, R1=valeryl or an isomer thereof, R2=valeryl or an isomer thereof, R4=valeryl or an isomer thereof and R5 is hydrogen or acetyl; or a sucrose ester as shown in Figure 5 wherein R3=acetyl, R1=butyryl, R2=hexanoyl or an isomer thereof, R4=hexanoyl or an isomer thereof and R5 is hydrogen or acetyl (C2C15:0); a sucrose ester that has the general structure shown in Figure 5 and wherein R3 is acetyl, R1 is valeryl or an isomer thereof, R2 is valeryl or an isomer thereof, R4 is hexanoyl or an isomer thereof and R5 is either a hydrogen atom or an acetyl moiety (C2C16:0); or wherein the sucrose ester has the general structure shown in Figure 5 and wherein R3 is acetyl, R1 is propionyl or

- an isomer thereof, R2 is hexanoyl or an isomer thereof, R4 is hexanoyl or an isomer thereof and R5 is either a hydrogen atom or an acetyl moiety (C2C16:0); a sucrose ester as shown in Figure 5 wherein R3 is acetyl, R1 is valeryl or an isomer thereof, R2 is hexanoyl or an isomer thereof, R4 is hexanoyl or an isomer thereof and R5 is either a hydrogen atom or an acetyl moiety is not present, is not substantially present or is not present at detectable levels (C2C17:0); and wherein a sucrose ester as shown in Figure 5 wherein R3 is acetyl, R1 is hexanoyl or an isomer thereof, R2 is hexanoyl or an isomer thereof, R4 is hexanoyl or an isomer thereof and R5 is either a hydrogen atom or an acetyl moiety is not present, is not substantially present or is not present at detectable levels (C2C18:0).
- 10 Alternatively, the sucrose ester may be a sucrose ester in which the total number of carbons in the acyl chains at R1, R2 and R4 is 14; a sucrose ester in which the total number of carbons in the acyl chains at R1, R2 and R4 is 15; a sucrose ester in which the total number of carbons in the acyl chains at R1, R2 and R4 is 16 is not present, is not substantially present or is not present at detectable levels; a sucrose ester in which the total number of carbons in the acyl chains at R1, R2 and R4 is 17 is not present, is not substantially present or is not present at detectable levels; and a sucrose ester in which the total number of carbons in the acyl chains at R1, R2 and R4 is 18 is not present, is not substantially present or is not present at detectable levels. Provided that this total number of carbons is present then the constituent acyl chains may be selected from any of those set forth in Figure 6. Suitably, the sucrose ester has an acetyl at R3 and an acetyl or hydrogen at R5. Suitably, the sucrose ester has an acetyl at R3 and a hydrogen at R5. Suitably, one or more, two or more or three or more of the acyl chains will be beta-methylvaleryl. Thus, by way of example, the sucrose ester can be a sucrose ester as shown in Figure 5 wherein R3=acetyl, R1=propionyl or an isomer thereof, R2= propionyl or an isomer thereof, R4=hexanoyl or an isomer thereof and R5 is hydrogen or acetyl; or a sucrose ester as shown in Figure 5 wherein R3=acetyl, R1= propionyl or an isomer thereof, R2=valeryl or an isomer thereof, R4=valeryl or an isomer thereof and R5 is hydrogen or acetyl; or a sucrose ester as shown in Figure 5 wherein R3=acetyl, R1=butyryl, R2=valeryl or an isomer thereof, R4=hexanoyl or an isomer thereof and R5 is hydrogen or acetyl (C2C14:0); a sucrose ester as shown in Figure 5 wherein R3=acetyl, R1=propionyl or an isomer thereof, R2=valeryl or an isomer thereof, R4=hexanoyl or an isomer thereof and R5 is hydrogen or acetyl; or a sucrose ester as shown in Figure 5 wherein R3=acetyl, R1=valeryl or an isomer thereof, R2=valeryl or an isomer thereof, R4=valeryl or an isomer thereof and R5 is hydrogen or acetyl; or a sucrose ester as shown in Figure 5 wherein R3=acetyl, R1=butyryl, R2=hexanoyl or an isomer thereof, R4=hexanoyl or an isomer thereof and R5 is hydrogen or acetyl (C2C15:0); a sucrose ester that has the general structure shown in Figure 5 and wherein R3 is acetyl, R1 is valeryl or an isomer thereof, R2 is valeryl or an isomer thereof, R4 is hexanoyl or an isomer thereof and R5 is either a hydrogen atom or an acetyl moiety is not present, is not substantially present or is not present

at detectable levels (C2C16:0); or wherein the sucrose ester has the general structure shown in Figure 5 and wherein R3 is acetyl, R1 is propionyl or an isomer thereof, R2 is hexanoyl or an isomer thereof, R4 is hexanoyl or an isomer thereof and R5 is either a hydrogen atom or an acetyl moiety is not present, is not substantially present or is not present at detectable levels  
5 (C2C16:0); a sucrose ester as shown in Figure 5 wherein R3 is acetyl, R1 is valeryl or an isomer thereof, R2 is hexanoyl or an isomer thereof, R4 is hexanoyl or an isomer thereof and R5 is either a hydrogen atom or an acetyl moiety is not present, is not substantially present or is not present at detectable levels (C2C17:0); and wherein a sucrose ester as shown in Figure 5 wherein R3 is acetyl, R1 is hexanoyl or an isomer thereof, R2 is hexanoyl or an isomer thereof,  
10 R4 is hexanoyl or an isomer thereof and R5 is either a hydrogen atom or an acetyl moiety is not present, is not substantially present or is not present at detectable levels (C2C18:0).

A further aspect relates to a mutant, non-naturally occurring or transgenic plant (suitably, a tobacco plant), wherein expression of isopropylmalate synthase as described herein or the activity of the protein encoded thereby is increased as compared to a control plant in which the expression of isopropylmalate synthase or the activity of the protein encoded thereby has not  
15 been increased and wherein said plant produces beta-methylvaleryl containing sucrose esters therein at levels higher than said control plant. A further aspect relates to a mutant, non-naturally occurring or transgenic plant (suitably, a tobacco plant), wherein expression of isopropylmalate synthase as described herein or the activity of the protein encoded thereby is decreased as compared to a control plant in which the expression of isopropylmalate synthase or the activity of the protein encoded thereby has not been decreased and wherein production of beta-methylvaleryl containing sucrose esters is lower than said control plant.  
20

A further aspect relates to a mutant, non-naturally occurring or transgenic flue cured tobacco plant, wherein expression of isopropylmalate synthase as described herein or the activity of the protein encoded thereby is increased as compared to a control plant in which the expression of isopropylmalate synthase or the activity of the protein encoded thereby has not been increased and wherein production of beta-methylvaleryl containing sucrose esters is higher than said control plant.  
25

A further aspect relates to a mutant, non-naturally occurring or transgenic oriental tobacco plant, wherein expression of isopropylmalate synthase as described herein or the activity of the protein encoded thereby is decreased as compared to a control plant in which the expression of isopropylmalate synthase or the activity of the protein encoded thereby has not been decreased and wherein production of beta-methylvaleryl containing sucrose esters is lower than said control plant.  
30

The mutant, non-naturally occurring or transgenic tobacco plants described herein can be harvested and the tobacco derived therefrom can be treated or cured using practices known for the chosen tobacco type. The leaf surface compounds can be collected, using for example  
35

water or organic solvent washes and subsequently dried out. The resulting powder can then be used to enhance the flavour properties of smoking material. The tobacco material remaining after the washes could then be used as material for expanded tobacco for example.

In a further aspect, there is provided a method for modulating the pest resistance in at least a part of a plant, comprising the steps of: (i) modulating the expression or activity of isopropylmalate synthase in the plant, preferably, wherein the isopropylmalate synthase comprises the polynucleotide sequence or the polypeptide sequence described herein; (ii) optionally measuring the quantity of beta-methylvaleryl sucrose esters in at least a part of the mutant, non-naturally occurring or transgenic plant obtained in step (i); (iii) identifying a mutant, non-naturally occurring or transgenic plant in which the quantity of beta-methylvaleryl containing sucrose esters therein has changed in comparison to a control plant in which the expression or activity of isopropylmalate synthase has not been modulated and, preferably, wherein the visual appearance of said mutant, non-naturally occurring or transgenic plant is substantially the same as the control plant; and (iv) obtaining a plant in which the pest resistance thereof has been modulated.

Suitably, there is provided a method for increasing the pest resistance in at least a part of a plant, comprising the steps of: (i) increasing the expression or activity of isopropylmalate synthase in the plant, preferably, wherein the isopropylmalate synthase comprises the polynucleotide sequence or the polypeptide sequence described herein; (ii) optionally measuring the quantity of beta-methylvaleryl containing sucrose esters in at least a part of the mutant, non-naturally occurring or transgenic plant obtained in step (i); (iii) identifying a mutant, non-naturally occurring or transgenic plant in which the quantity of beta-methylvaleryl containing sucrose esters therein has increased in comparison to a control plant in which the expression or activity of isopropylmalate synthase has not been increased and, preferably, wherein the visual appearance of said mutant, non-naturally occurring or transgenic plant is substantially the same as the control plant; and (iv) obtaining a plant in which the pest resistance thereof has been increased.

A plant or plant material obtained or obtainable by said method is also provided.

A further aspect relates to a mutant, non-naturally occurring or transgenic plant, wherein expression of isopropylmalate synthase or the activity of the protein encoded thereby is modulated and at least a part of the plant has a change in the composition of beta-methylvaleryl containing sucrose esters as compared to a control plant in which the expression or the activity of isopropylmalate synthase has not been modulated and wherein the visual appearance of said plant is substantially the same as the control plant and wherein the pest resistance thereof is also modulated.

Suitably, a mutant, non-naturally occurring or transgenic plant is provided, wherein expression of isopropylmalate synthase or the activity of the protein encoded thereby is increased and at

least a part of the plant has a change in the composition of beta-methylvaleryl containing sucrose ester(s) as compared to a control plant in which the expression or the activity of isopropylmalate synthase has not been increased and wherein the visual appearance of said plant is substantially the same as the control plant and wherein the pest resistance thereof is also increased.

Compositions, methods and kits for genotyping plants for identification, selection, or breeding are encompassed and can comprise a means of detecting the presence of a isopropylmalate synthase polynucleotide in a sample. Accordingly, a composition is described comprising one of more primers for specifically amplifying at least a portion of isopropylmalate synthase polynucleotide and optionally one or more probes and optionally one or more reagents for conducting the amplification or detection. Gene specific oligonucleotide primers or probes comprising about 10 or more contiguous polynucleotides corresponding to the isopropylmalate synthase polynucleotide are disclosed. Said primers or probes may comprise or consist of about 15, 20, 25, 30, 40, 45 or 50 more contiguous polynucleotides that hybridise (for example, specifically hybridise under, for example, stringent hybridisation conditions) to the isopropylmalate synthase polynucleotide. In some embodiments, the primers or probes may comprise or consist of about 10 to 50 contiguous nucleotides, about 10 to 40 contiguous nucleotides, about 10 to 30 contiguous nucleotides or about 15 to 30 contiguous nucleotides that may be used in sequence-dependent methods of gene identification (for example, Southern hybridization) or isolation (for example, *in situ* hybridization of bacterial colonies or bacteriophage plaques) or gene detection (for example, as one or more amplification primers in nucleic acid amplification or detection). The one or more specific primers or probes can be designed and used to amplify or detect a part or all of the isopropylmalate synthase polynucleotide. By way of specific example, two primers may be used in a polymerase chain reaction protocol to amplify a nucleic acid fragment encoding isopropylmalate synthase nucleic acid – such as DNA or RNA. The polymerase chain reaction may also be performed using one primer that is derived from the isopropylmalate synthase nucleic acid sequence and a second primer that hybridises to a sequence upstream or downstream of the isopropylmalate synthase nucleic acid sequence – such as a isopropylmalate synthase promoter sequence, the 3' end of the mRNA precursor or a sequence derived from a vector and the like. Examples of thermal and isothermal techniques useful for *in vitro* amplification of polynucleotides are well known in the art. The sample may be or may be derived from a plant, a plant cell or plant material or a tobacco product made or derived from the plant, the plant cell or the plant material as described herein.

Thus, in a further aspect, there is also provided a method of detecting an isopropylmalate synthase polynucleotide in a sample comprising the step of: (a) providing a sample comprising, or suspected of comprising, a polynucleotide; (b) contacting said sample with one of more

primers or one or more probes for specifically detecting at least a portion of the isopropylmalate synthase polynucleotide; and (c) detecting the presence of an amplification product, wherein the presence of an amplification product is indicative of the presence of the isopropylmalate synthase polynucleotide in the sample. In a further aspect, there is also provided the use of one of more primers or probes for specifically detecting at least a portion of isopropylmalate synthase polynucleotide. Kits for detecting at least a portion of the isopropylmalate synthase polynucleotide are also provided which comprise one of more primers or probes for specifically detecting at least a portion of isopropylmalate synthase polynucleotide. The kit may comprise reagents for polynucleotide amplification - such as polymerase chain reaction (PCR) - or reagents for nucleic acid probe hybridization-detection technology - such as Southern Blots, Northern Blots, in-situ hybridization, or microarray. The kit may comprise reagents for antibody binding-detection technology such as Western Blots, ELISAs, SELDI mass spectrometry or test strips. The kit may comprise reagents for nucleic acid sequencing. The kit may comprise reagents and/or instructions for determining sucrose ester content. In some embodiments, a kit may comprise instructions for one or more of the methods described. The kits described may be useful for genetic identity determination, phylogenetic studies, genotyping, haplotyping, pedigree analysis or plant breeding particularly with co-dominant scoring.

The present invention also provides a method of genotyping a plant, a plant cell or plant material comprising a isopropylmalate synthase polynucleotide. Genotyping provides a means of distinguishing homologs of a chromosome pair and can be used to differentiate segregants in a plant population. Molecular marker methods can be used for phylogenetic studies, characterizing genetic relationships among crop varieties, identifying crosses or somatic hybrids, localizing chromosomal segments affecting monogenic traits, map based cloning, and the study of quantitative inheritance. The specific method of genotyping may employ any number of molecular marker analytic techniques including amplification fragment length polymorphisms (AFLPs). AFLPs are the product of allelic differences between amplification fragments caused by nucleotide sequence variability. Thus, the present invention further provides a means to follow segregation of a isopropylmalate synthase gene or nucleic acid as well as chromosomal sequences genetically linked to these genes or nucleic acids using such techniques as AFLP analysis.

The disclosure is further described in the Examples below, which are provided to describe the disclosure in further detail. These examples, which set forth a preferred mode presently contemplated for carrying out the disclosure, are intended to illustrate and not to limit the disclosure.

#### EXAMPLES

**Example 1:** Sucrose ester analysis in the *Nicotinia tabacum* varieties of flue cured Hicks

Broadleaf (HBL) and primitive Red Russian (RR)

To study sucrose ester synthesis in tobacco, two model varieties are chosen: the flue cured Hicks Broadleaf (HBL) and the primitive Red Russian (RR). Initial data available from the National Genetic Resources Program (Germplasm Resources Information Network - (GRIN),  
 5 National Germplasm Resources Laboratory, Beltsville, Maryland) catalogue suggests a different content in beta-methylvaleryl containing sucrose ester content between these two lines.

The composition of HBL and RR leaf exudates in sucrose esters is analysed and the results are presented in Figure 1. RR sucrose ester composition differs from HBL sucrose ester composition. In HBL, sucrose esters are mainly composed of C2C15:0 whereas in RR the  
 10 presence of higher molecular weight species - such as C2C16:0, C2C17:0 and C2C18:0 - are observed. As shown in Figure 2, the differences between RR and HBL are accounted for by the presence or absence of methylvaleryl esters in the molecules. Thus it is concluded that RR is a BMVSE variety and HBL a *bmvse* variety.

**Example 2:** Quantitative tract locus analysis

15 The composition of sucrose esters in leaf exudates is analyzed in 136 populations of 8 F3 plants descending respectively from 136 selfed F2 genotyped plants that were resulting from the cross of HBL and RR parents. A quantitative trait locus study is performed to correlate microsatellites markers and the beta-methylvaleryl containing sucrose ester phenotype. Data is analyzed using Quantitative trait locus cartographer. The beta-methylvaleryl containing sucrose  
 20 ester trait is linked to one single locus in the tobacco genome. This locus is located on the linkage group 15 ( chromosome A) of tobacco.

**Example 3:** Identification of the beta-methylvaleryl containing sucrose ester related gene

The transcription level of tobacco exons in trichomes is studied to select the preferential gene target among all the potential genes encoding for the functions described in Figure 3. The  
 25 expression was studied in several varieties selected for their beta-methylvaleryl containing sucrose ester production profile using microarrays.

The results identify isopropylmalate synthase related exon probes since their expression correlates with the estimated amounts of beta-methylvaleryl containing sucrose esters measured in each of the extracts of each variety ( $R>0.95$ ).

30 Four tobacco sequences are identified encoding for homologs of the tomato isopropylmalate synthases IPMSA and IPMSB. They are named *NtIPMS1A* (SEQ ID NO: 10), *NtIPMS1Bv1* (SEQ ID NO: 12), *NtIPMS1Bv2* (SEQ ID NO:14) and *NtIPMS2* (SEQ ID NO:1). According to EST databases, SEQ ID NO:1 was the choice candidate since most of its expression can be observed in trichomes.

35 A BAC library constructed from *N. tabacum* var. HBL is screened using the *NtIPMS2* cDNA as PCR bait, identifying a BAC named PISOGE. This BAC is shotgun sequenced. The assembly produces several contigs of which pisoge1 and pisoge2 contain sequences related to *NtIPMS2*.

The intron-exon structure is conserved among isopropylmalate synthases from tobacco and tomato. For *NtIPMS1Bv1*, *NtIPMS1Bv2*, *SlIPMSB* and *SlIPMSA* it is observed that the genes are separated into twelve exons. *NtIPMS2* is also expected to be composed of twelve exons but since the *SlIPMSA* first exon corresponds to two exons on *pisoge1* it is postulated that

5 *NtIPMS2* contains thirteen exons.

The reliability of the genomic sequence prediction is tested by PCR. Globally the predicted sequence from the *pisoge* assembly is verified using the purified *PISOG* BAC or genomic DNA isolated from either Hicks Broadleaf or Red Russian plants. One full length genomic structure corresponding to *NtIPMS2* is identified. In addition, two other gene like structures are observed and are likely to represent pseudogene duplicates. The structure proposed by the *PISOG* BAC is partially verified but some discrepancies between the BAC structure and the HBL genomic DNA are observed.

PCR is used as a genetic marker to distinguish between RR and HBL alleles of *NtIPMS2*. The *IPMS2* marker is used in PCR amplification on genomic DNA of all the F2 plants used in the quantitative trait locus mapping. The sucrose ester quantitative trait locus is recalculated with this additional marker. The quantitative trait locus peak identifies the location where the probability of association between the trait and the marker is maximal. The *IPMS2* marker, together with the simple sequence repeat marker PT30172 are the markers where the probability of effect on the beta-methylvaleryl containing sucrose ester trait is the highest.

20 **Example 4:** Demonstration of the activity of *NtIPMS2* as the key factor in the synthesis of beta-methylvaleryl containing sucrose esters

The express sequence tags corresponding to *NtIPMS2* are isolated from trichome EST libraries in the form of: GT09L09 (origin: Galpao [T11068] trichomes) and OT05C20 (origin Orinocco [T181] trichomes). Both molecules are identical in nucleotide sequence and thus only the OT05C20 cDNA is used for expression experiments. An RNAi construct is generated using the first exon and the first intron of *NtIPMS2* based on consensus sequence. Both constructs are transferred into a binary vector comprising either the *CPS2p* promoter (*CPS2* refers to the copalyl synthase 2 involved in cis-abienol synthesis which occurs in trichomes only) specific to trichomes or the *MMVp* viral promoter (*MMV* refers to the mirabilis mosaic virus). Stable transformations using *Agrobacterium tumefaciens* AGL1 carrying respective binary vectors are performed in the oriental tobacco variety Basma xanthi and the flue cured variety K326.

Leaves are collected from one month old regenerating plants resulting from independent transformation events. Internal standard (sucrose octaacetate) is added at the rate of 2.5 ng/mg fresh weight. The area of the signal recorded for the ion specific to each sucrose ester is measured and plotted as ratio of internal standard area. The signal response varies from molecule to molecule so the comparison is valid only between each molecule. The ratio to internal standard signal area allows the normalisation of signals on leaf fresh weight. Figure 2

describes the sucrose ester nomenclature presented here. Referring to Figure 4: (A) composition of leaf exudates of the parental variety K326 (Flue cured, n=3) and the transformants expressing NtIPMS2 under a trichome specific promoter; (B) composition of leaf exudates of the parental variety K326 and the transformants expressing NtIPMS2 under a viral promoter; (C) composition of leaf exudates of the parental variety Basma (Oriental, n=4) and the transformants expressing NtIPMS2-RNAi construct under a trichome specific promoter; (D): composition of leaf exudates of the parental variety Basma and the transformants expressing NtIPMS2-RNAi construct under a viral promoter.

The data presented are sufficient to demonstrate that NtIPMS2 is the key gene in beta methylvaleric acid containing sucrose ester production. Suppression of IPMS2 in a beta methylvaleric acid containing sucrose ester variety (oriental) reduces or abolishes production of beta methylvaleric acid containing sucrose ester (see C and D) thus proving that IPMS2 is a key component of the biosynthetic pathway. The constitutive and trichome specific expression of an RNA interference construct results in the loss of accumulation of the sucrose esters C2C16:0, C2C17:0 and C2C18:0 which are the esters which can contain beta methylvaleric acid.

The introduction of NtIPMS2 in a *bmvs* variety (flue cured) can induce the production of sucrose esters (see A and B). This proves that NtIPMS2 cDNA expression alone is sufficient to restore the production of the sucrose esters. Here the constitutive and trichome specific expression of NtIPMS2 cDNA results in the *de novo* accumulation of the sucrose esters C2C16:0, C2C17:0 and C2C18:0 which are the esters that can contain beta methylvaleric acid. Interestingly, the constitutive expression of the cDNA (B) did not result in a high accumulation of the sucrose esters whereas the trichome specific expression resulted in the accumulation of the sucrose esters to levels that are comparable to what can be observed in, for example, the beta methylvaleric acid containing sucrose ester variety.

**Example 5:** Search protocol for the selection of zinc finger nuclease target sites

This example illustrates how to search the isopropylmalate synthase gene to screen for the occurrence of unique target sites within the given gene sequence compared to a given genome database to develop tools for modifying the expression of the gene. The target sites identified by methods of the disclosure, the sequence motifs, and use of any of the sites or motifs in modifying the corresponding gene sequence in a plant, such as tobacco, are encompassed in the disclosure.

*Search algorithm.* A computer program is developed that allows one to screen an input query (target) nucleotide sequence for the occurrence of two fixed-length substring DNA motifs separated by a given spacer size using a suffix array within a DNA database. The suffix array construction and the search use the open source libdivsufsort library-2.0.0 (<http://code.google.com/p/libdivsufsort/>) which converts any input string directly into a Burrows-Wheeler transformed string. The program scans the full input (target) nucleotide sequence and

returns all the substring combinations occurring less than a selected number of times in the selected DNA database.

*Selection of target site for zinc finger nuclease-mediated mutagenesis of a query sequence.* A zinc finger DNA binding domain recognizes a three basepair nucleotide sequence. A zinc finger nuclease comprises a zinc finger protein comprising one, two, three, four, five, six or more zinc finger DNA binding domains, and the non-specific nuclease of a Type IIS restriction enzyme. Zinc finger nucleases can be used to introduce a double-stranded break into a target sequence. To introduce a double-stranded break, a pair of zinc finger nucleases, one of which binds to the plus (upper) strand of the target sequence and the other to the minus (lower) strand of the same target sequence separated by 0, 1, 2, 3, 4, 5, 6 or more nucleotides is required. By using plurals of 3 for each of the two fixed-length substring DNA motifs, the program can be used to identify two zinc finger protein target sites separated by a given spacer length.

*Program inputs:*

1. The target query DNA sequence
2. The DNA database to be searched
3. The fixed size of the first substring DNA motif
4. The fixed size of the spacer
5. The fixed size of the second substring DNA motif
6. The threshold number of occurrences of the combination of program inputs 3 and 5 separated by program input 4 in the chosen DNA database of program input 2

*Program output:* A list of nucleotide sequences with for each sequence the number of times the sequence occurs in the DNA database with a maximum of the program input 6 threshold.

Any publication cited or described herein provides relevant information disclosed prior to the filing date of the present application. Statements herein are not to be construed as an admission that the inventors are not entitled to antedate such disclosures. All publications mentioned in the above specification are herein incorporated by reference. Various modifications and variations of the disclosure will be apparent to those skilled in the art without departing from the scope and spirit of the disclosure. Although the disclosure has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in cellular, molecular and plant biology or related fields are intended to be within the scope of the following claims.

**SEQUENCES****SEQ ID NO:1 (NIPMS2 nucleotide sequence)**

ATGGCTTCTCTCTGTAAATCTATAATTTCCCTGAGCACTTCCCTTTTCATFACATTCTAAAAACCCACTTATTCA  
 CAGTGTCTTCAGTTTCAGGCTTCAACCAACAAGGCATTACAGCTATATGCTGCTCAAAATATTTCGTCGGCGCCGGAGT  
 ATAAACATGGAAAAATCTCTGACCCCTGATTATGTTGGTATTTTTCACACAGCTCTTCGCGATGGCGAACAGGCCGCT  
 GGTGCTACCATGACTAGTAAGAAAACTGGACATTGCACGTCAGTTGGCTAAGCTTGGTGTGATGTTATTGAGGC  
 CGGTTTTCCTTTTCCCTCTGAAGCTGAGTTCGAGCTTGTAAAGTTGATAGCAGGAAATTTGGTAATACGTAGACA  
 AAGAGGATACGTGCCGATGATATGTGCCCTTAGCTAGGCTAGTAAGAGGATATTGAAAGAGCTTGGGATGCTTTTA  
 AAGTATGCAAGAAACCAATGCTTCATATGTTTATTGCGACGAGTGATATACATATGAAGTACAAGTTAAAGATGAG  
 TAGAGAAGAAATTTGGAGACAGCTAGGAGTACGGTGGCTTATGCAAAAACCCCTATTGAGGATGTTCCGTTTAGCG  
 CTGAAGATGCTGCAAGATCTGATAGGAGTTCCTTTATCATATTTATGGAGAAGTTATCAAGCTGGTGCACACGTG  
 ATTGGCCTCCCTGATACAGTTGGATGCAATTTGCCAGTGAATATGCACAACTGATTTCTGATATAAAAGCCATAC  
 CCCAGGAATACAAGATGCAAACTTTCAACACACTGTCACAACTGCTTGGGCTTGCTACTGCCCACTCCTTAGCTG  
 GAATTTGCGCAGGCGCAAGACTAGTAGATGTACCATCAATGGAATTGGTGAAGAGCTGGAAATGCTTCTCTGGAG  
 GAGATTGTAATGGCCTTAAATATTCGTGGAGAGCAAGTACTAGGTGGTATCTATCTAGGATTAATACAAAGCATAT  
 ATTTCATGACGAGCAAAATGGTAGAAGAGTACAGTGGGCTTAAGCTGCAGCCACATAAGGCCATTGTTGGAGCTAATG  
 CATTTCATGAGAGTGGCATCCATCAGGATGGAGTGTAAAGAACAGAGATACATATGAGTTTGTATCTCATGAA  
 CATGTTGGGTATCGTCTGCTAATGAAACCGGTATTAGTCTGGGAAAGCTCAGTGGCCGCGCATGCATTGAAAGCCAA  
 AATGGCTGAGCTTGGATATGACTTTGATGGAAAGAACTTGATGACCTCTTTGTCGATTCAAGTCACTAGCTGAGA  
 GGAAAGAGAAATTACAGATGATGACTTCAGACGACTTGATCAGATGACGTTTCCAGCCCTCAAGTTCTCTGGCAA  
 CTTGGAGATGTACAGATTACTTGTGGAATGTTGGCCGCTCTACAGCAATGTTAAGCTTATTGACAGCGATGGTCA  
 AGAGCACACPGCCTTTCTGTTGGAACAGGACCTGTTGATGCAGCTTACAAGGACAGTTGACCTCATTTGTAAGGTAC  
 CTGTAAACACTCGTTGAATATTTCGTTAATGCAATCACAAACGCTATAAATTCACAGCTTCAACCAGAGTGTAGTT  
 CCTGGGAATGATGACTATGCATCGTTTAATACTTCAACCGGCAAACTGTTAATCGTACAGTTAGTGGAACAGGAGC  
 GCATATGGACATTTGTCGTTCAAGTGTCAAGCCTATGTTGAGGCGTTGAACAAATATTCAGTTACAAAAAACAG  
 GTCTCGTGACAAATTTGAAGGAGTGGCCTCGTAA

**SEQ ID NO:2 (NIPMS2 amino acid sequence)**

MASLSVNSIISLSTSLSLHSKNPLIHSVFSFTPTTRHSAICCSNIRRRPEYKHGKFSDDPDYVGI FDTSLRGEQAA  
 GATMTSKEKLDIARQLAKLGVDFIEAGFPFASEAEFELVKLIAQEI GNNVDKEGYVPMICALARSSKKDIERAWDAL  
 KYAKKPLMHFIATSDIHMXYKLKMSREEIVEFARSTVAYAKTLFEDVRFSAEDARSDFLYHIIIGEVIKAGATV  
 IGLPDTVGCNLPSEYQQLISDIKANTPGIQDANISTHCHNDLGLATANSLAGICAGARLVDOVTINGIGERAGNASLE  
 EIVMLKRYRGEQVLGGIYTGINTKHI FMTSKMVEEYSLKLPFKAIIVGANAFSHESGIHQDGLKNRDTYEFVSHE  
 DVGVRANENGISLGLSGRHALKAKMAELGYDFDGKELDDLFRFKSLAERKKKTTDDDLRALVSDVFPQVSWQ  
 LGDVQITCGNVGRSTANVKLIDSDGQEHAFSVGTGPVDAAVKAVDLIVKVPVTILVEYSVNAITKRINSTASTRVLV  
 RGNDDYASFNTSNGQTVNRVTGSGTGAHMDIVVSSVQAYVEALNKIFSYKKKTGLVKNKFEQSAQS

**SEQ ID NO:3 (NIPMS2 cDNA sequence used for complementation)**

ctcttttcttagggaaaagaataatggcttctctctctgtaaattctataatttccctgagcacttccctttcatta  
 cattctaaaaaacccacttattcacagtgcttctcagtttccagccttcaaccacaaggcattcagctatattgctgctc  
 aaattattcgtcggcggcggcggagtataaacatggaaaattctctgacccctgattatggttggtattttgacaccagctc  
 ttcgctgagtgccgaacacggcggctggtgctaccatgactagtaaaagaaaaactggacattgcacgtcagttggctaaag  
 cttggtggtgatgttattgagggcgggttttcccttttgccctctgaagctgagttcgagcttgtaaaagtgtatagcaca  
 ggaaattggttaataacgttagacaaaagaagatagctgcccgtgatgtgtgctttagctaggtctagttaagaaggata  
 ttgaaaagagcttggggtgctttaaagtagcaaaagaaaccaatgcttcataatgtttattggacgagtgatatacat  
 atgaagtacaagtttaagatgagttagagaagaattgtggagacagctaggagtagcgtggcttatgcaaaaaccct  
 atttgaggatggttcggttttagcgtgcaagatgctgcaagatctgataggagttccctttatcatattattggagaag  
 ttatcaaagctggtgcaacagtgattggcctccctgatacagttggatgcaatttgcccagtgaaatgcacaaactg  
 atttctgataataaagccaataccocagggaatacaagatgcaaacatttcaacacactgtcacacagatcttgggct  
 tgctactgccaaactccttagctggaatttgcgcaggcgcgaagactagtagatgttaccatcaatggaattggtgaaa  
 gagctggaaatgcttctctggaggagattgtaattggccttaaaatatcgtggagagcgaagtactaggtggtatctat  
 actgggatttaatacaagcatatattcatgacgagcaaaatggtagaagagtagcagtgggcttaagctgcagccaca  
 taaggccattgttgagctaatgcatttctcatgagagtggcatccatcaggtggagtggttaagaacacagagata  
 caatagagtttctatctcatgaagatgttgggtatcgtcgtgctaagtaaaacgggtattagctctgggaaagctcag  
 gcccgcctgcatgaaagccaaaatggctgagcttggaatagactttgatggaaaagaacttgatgacctctttc  
 tcgattcaagctcactagctgagaggaagaaaataacagatgagcttgagagcacttgatcagatgagcttt

tccagcctcaagtttctctggcaacttggagatgtacagattacttctgtgaaatgttggcgcgtctacagcaaatgtt  
 aagcttattgacagcgatgggtcaagagcacactgccttttctgttggacaggacctgttgatgcagcttacaaggc  
 agttgacctcattgttaaaggtacgtgtaacactcggtgaatattcgggttaatgcaatcacaaaaacgtataaattcca  
 cagcttcaaccagagtgtagtctgtgggaatgatgactatgcacgtttaaacttcaaacgggcaactgttaaat  
 cgtacagtttagtggaaacaggagcgcatatggacattgtcgtttcaagtgccaagcctatgttgaggcgttgaacaa  
 aatattcagttacaaaaaacagggtctcgtgaacaaatttgaaggcagtgcgcaatcgtaaaagtgtggtgtgtgc  
 tccagaaacaaaagccttaaatgctgcttagtgcctgatgatcataaatttatgctctattcatgcattactga  
 agttaattagaagtaattcttaattgtataagatgtaaacatgttcttattcatgcattactatgaaattaattaca  
 atttctgagccaaaaa

SEQ ID NO:4 (Genomic DNA sequence of *NtIPMS2* region covering exons XI, XII and XIII as observed in Red Russian. cDNA matching regions are underlined. Bold shows conserved DNA regions between Red Russian and Hicks Broadleaf DNA sequence)

AGCAAGATGTCAGAAGCACACTTGCCCTTTCTCTGTTGGAACAGGACCTGTTGATGCAGCTTACAAGGCAGTTGACC  
TCATTGTAAAGGTTTATAAAGAACTGAAACCTCTCATATGTTTGTATCTTGTACACTAGTAGGATAGGTAATTT  
AGGTTTTATATACGCTGAAGGTATAAAAAATTTTACTCAATCAGAGCGCTTACAAGACAATTGCAGGTAACCTCTATG  
AGTAGCTTAGACTAAAAGATAAACGTAATTTCAATGTAAATAGTTTAGGAGACATGCGAATCTTTCCCTTGACAA  
AGAACTAAAAAAGTTCGCAAGCACGACCTCAATCACCCGCTCTGTGTGAATTGAACGAACCTAAAAGTACAACCA  
GCTTAACTGATTTTTTACCTATTTTTAGTGTTCAGACGGATCTGCAGCAACCATGATTTGATTCACCGGTGACCA  
AATATCTCACCGTGCAATTCGAATCCCTTTCTCCACATCTTAAATCTCTCTAAGATGCTTTTGTTCGCAAC  
TCTAGCCGTCATATCTTGGTGGAAACATGCTCTCTTAATATGTGTATTTATATGTTATAAATAGTATTTAGGAGA  
ATGGCAAGTTAACAGACTAATTACACCTATTATGAATGGATCCACCGAATAGAAATATTCACACTTTAGCTAAG  
TAAGCTGCCAAGTCAATAAGACTGATAGAGTAAAACTCTTTCATATAGCAATGAATATAACCTAAATCCTTTAGT  
TACTGCTAGCAATTTTGTAGGTATAATGCTTTGTTGGTAAAAACATCTATGTAACCTTTGTTTATAAGTAATGT  
TCITTTGTCATTATGCAGGTACCTGTACACTCGTTGAATATTCGGTTAATGCAATCACAAACGTATAAATTCACAC  
GCTTCAACCCAGAGTGTAGTTCTGTTGGGAATGATGACTATGCATCCGTTTAACTCTCAACGGGCAACCTGTTAATC  
GTACAGTTAGGTATGAAATATACATGAAGTACTGTTCCCTTTAATTTATCCTATTTCCTTAATAATATTTCTTTC  
ATGCATTAATGTGGGAACATGAAGTACAGAGAGAATTAACAGATTTTTCCTAAAGGAGAAATAATATGTGAT  
GAAGATGAATGTTGTGAGTTTATATAGTTCTCAATTTTATGGTACTACGGGTATCTAGAGCAACCTGAGAATA  
GAAAGTATATATAGAAATATACACAGCAGTTAGATAGAGTACATTTTAAATATTATATATATATGCAATGCC  
TCAACACGTTTACTCACCTCTTAGCCTAATTTTGTACATGGGATCGTTCTACATTTTGAGTAGCTGAGAGATTG  
AACTTAGGACCTTTATTTTCTTAACACTATGTTATATGCCAGTGGAAACAGGAGCGCATATGGACATTGTCGTTTC  
AAGTGTCCCAAGCCTATGTTGAGGCGTTGAACAAAATATTCAGTTACAAAAAACAGNTCTCGTGAACAAATTTGAAG  
CGAGTGCAGATCGTAAAGTGTGTTGTGTCNCTNCTANAAG

SEQ ID NO:5 (Genomic DNA sequence of *NtIPMS2* region covering exons XI and XIII as observed in Hicks Broadleaf. cDNA matching regions are underlined. Bold shows conserved DNA regions between Red Russian and Hicks Broadleaf DNA sequence.)

GTCAAGAGCACACTGCCTTTTCTGTTGGAACAGGACCTGTTGATGCAGCTTACAAGGCAGTTGACCTCATTGTAAAG  
GTTTATATGAAGTGAAGTCTCAATATGTTTGTATCTTGTACACTAGTAGGATAGGTAATTTAAGTTTATAT  
ACGCTGAAGGTATAAAAAATTTTACTCAATCAGAGCGCTTACAAGACAATTGCAGGTAACCTATGAGTAGCTTAGA  
CTAAAAGTTTAAATATTATTAATATTATGCAATGCCCTCAACACGTTTACTCACCTCTTAGCCTAATTTTGTAC  
ATGGGATCGTTCTACATTTTGTAGTAGCTGAGAGATTGAACTTAGGACCTTTATTTCTTAACACTATGTTATAT  
GCCAGTGGAAACAGGAGCCATATGGACATTGTGCTTCAAGTGTCCAAGCCTATGTTGAGGCGTTGAACAAAATATT  
CAGTTACAAAAAACAGGTCTCGTGAACAAATTTGAAGGCAGTGCAGCAATCGTAAAGTATGTTGTGCTATCTGC  
CNAAAACTCGA

SEQ ID NO:6 (*NtIPMS2* corresponding BAC contig sequence (total BAC length is 76832bp. Only 14232bp are shown)

TGATGATTTCTCATTATAATTGTAAGTGTGGTGGCAAAATGGTAAAAAGAAAACAGTTATCCATCCATATTATCA  
 TTAAAAAATGGGTTGTATAATGAACCTTTTAAAAAAGGATCAATTATGGATAAGAACCATATTATCCGCTTAGAAAA  
 TGATATACCAATAAGTTAACTTCTACACTTGTAAAGCTTCAATTTGGGGTTCCTCAAGTTTGTGAGAGAGAAAT  
 CTCCCAAGTGATCATATTCAAGAGCTTGGATAAATGAATATCCATATTATCCGCGGATATCTCGTTTAAATCC  
 GTATTAAATATGAGTGGTTTGAATAATTTATCTGTTTGTATTATCTACTTTGGACATGTCCATACCCGACCCGA  
 CCCCCTCGTTTGTACCCCTAGTTGTACATATATCTGTTTGTCTTACCTAACGGGTGAGTTAAGTACCATCACGAC  
 TTGGTGGGATTTGGCTCTTACAAAACATCAATTTAGCCCTGATTATAATTTTATTATAAAATGAGGTTCCGGTAT

AATTTATTTTGGCTGATCAAATCTTGGTGATTCAAAGTTCAACTGAAATATGCAATACTAAATTTATTTCTTATTTAA  
TAAATTTTATGATGAGCTAATATTTATAATTTGAATAAAAAAAATTCATAATACTAGTATTTTGGAGTTGAATAGA  
AATTTGTCAGTACAGTTGAATATACAACTGTAATTTTCATAAAAAATTTTGTGTTGATATAAATAATCTTTCAA  
AAAAACATTTCTACAGAACTTGACTGAAACCATAAATAAGTATCTCACTACTTTTGTGCGACAAGGTTTCTTCCCTAAA  
TGAGATCAACAATCATTCTTTAAGGATTTTATTTTATTTTATCACAAATATATTACTCAATATCTAAGTAATC  
TTTAAAAATCTGTATACTATTACTATAAAGCATACMTGCAAAACGCGTACRATAAAAACTAATTTTAAAAAATAG  
GGGGAAGGAGAAGGGATGAACAGAACCCCTCATTAATAAGCAAAAAATTTAATATCCAACAAATGAACTGTTAAGA  
TTATAAAAAAAGAAAAAGAACTTCATTGTCGAAGTAAGTCAAGATACCTAGCCGTAATAATTTCCGGCACCTT  
TTAAAAAGAAATGATAAATATACTAGGGAAGCCATAAACCTCTTGGTATTATAAATACTATAGTTAGAAATTAGGCAAAA  
GCCGCTACTAGGAGTATTAATTTATTTGGTTCTGTTTATATTTATATCATTCGTTTGTGTAATGATGAGCAACGTT  
CCATTTACATGAAGAAACACGCTCAATGTTTATTCCTCATTCGTTTCTAATTTTATTCGAGATGTGAGTAACCTC  
CCCCACCCCAACACACCCCTTTTATTTTGTGTTTAAAGTTTCGACAATAGAACGTTTCAAGATTCTTGAACCTT  
CTGTTTATCTGCGAATAATATAAATTTGAGTCAACCCTGCATCCCAATTTGATGACAGCTTTATCTGAATTAATTA  
ATCAGTATATAATTAATATCTATTTTCTATGTTTCTAATTTATGTGACACATTTCTATTCGTTAGTTTAAAAAA  
AATGATAAATTTATATATTTTCAAGATAATTTAACTTTAAATTTTATTTTATCCATTTTACCTTTAATGAGAAAC  
TTTGTATACCAATACCAATGTTATAGAAATATTTAGAACACCAAGTTTCAATATTTTATAGTCACATAAATTTAAC  
ACACAATATTTAGAGCAAAATTTCAAACTACTTCACATAAATTAAGAACAGAGGAAAAAATAATACCTCCCTCCGTTCCC  
TTTACTTTGGCAGCTTTTACTTTTACGCCCCCTTAAGAAATAATAAATGAAGTGCATAATTTACCATGATACCCAT  
ATTAATTTGATGTATATTTTATTTGGATTGAAAAATGATTTGAAATGAGTAAATAATATTCGTTGGGTATTAACAAAA  
AAATTTATCTTCTCTGTATATGCGTAAAGTGACAAGCAAAATGAAATCTATTTTATGATATACATGCCAAATAAAG  
TGAAACGCTAAATGAACGGAGGGAGTATATTTATTTACTACTAAAAACGGGAACCTGGGCTATGCTAAGTTTCTCTT  
GCTATTTCAAGTTCCGATTAATAGCCATAATATAATGTTTCTAGAACAAAGAACACGCCAAGATAAATTTTCCC  
AAAAAGTAGCACAAATGTAACAAATCTTACCTCAAAATTTAATCTAAACACATAAATTTCTTTATAATTTGTTCTT  
TAAAAACTTAAACTAAATCAAAATGCCAACCAAGTTTAAATGAGCAAGGAAGTATTTCAAGTTCTATTTCTGAAAT  
TAAATAAGATGGAAATATTTCCGTTTGTTTACTGGGTTCTGTCTAGTCTAGCTTAAGACGAAAAATGTCTGTGTTT  
AAGTTTACCAACCGTAAGGGGTGGTAATCAATTTCAATTTGATAGTATGAGCAATTAATTCACACAAAAATCTCA  
CGTGAAATAATATACCTCTAGTGTTTGTGTTGGCTACGAAAAAGAAAGTCTACATAAATTAATGCAGCGTTCCATAGTT  
GATATTAATAATATACCGTAATAAATATATAAATTTTATGTTTATTTCTTTGATAAAACAGCAACACACACA  
CCAGGGGCTACTTTTGTCTATTTCTCTCCAATAAAAAACCTTTGTTTACCAGCTTTGTAGGTGTTTCAAGTTTCAACAA  
AGCGAATATCCATTAACCTCCGATCCACTCTCTTTCTTAGGGAAGAAATATAGGCTTCTCTCTGTGTAATTTCTA  
TAATTTTCCCTGAGCACTTCCCTTTTATTACATTTCTAATAACCCACTTATTACAGTGTCTTCAAGTTTCCAGCTTCA  
ACCACAAGGCATTCAGCTATATGCTGCTCAAAATTTCTGCGCGCGCGGAGTATAACATGGAATAATCTCTGACCC  
TGATTAATGTTGGTATTTTGTACACACAGTCTTCCGCGATGGCGAACAGGCGCTGTTGCTACCATGACTAGTAAGAAAA  
AATGAGCATTTGCACGTCAGTTGGCTAAGCTTGGTGTGATGTTATGAGGCGGTTTCTTTTCCCTTCCCTTCAAGCT  
GAGTTTCAGGCTTGTAAAGTTGATAGCAGCAAGAAATTTGGTAACCTTTAATGTTTAAACGTTTCAATTTCTAATTTA  
CTATTTTGTAAACATGTCGTCAGTGTAGTTTCTATTCTTTTATGAACCAACATGCTATGCAAGATATTTTAGAT  
ATTTGAGCGCGAGTGTAGATTTGAAACTAGGACCGTTTGCCTGATACAAATTAATAATATGTAACCAATTTTATGTAC  
AAGTTTAACTGTTGATAGTAGCATATTTTACTTTTATTTAAGTATACATATTTCAACAGGTAAATACGTAAGAC  
AAAGAAGGATACGTGCCGATGATATGTGCCTTAGCTAGGTCTAGTAAGAAGGATATGAAAGAGCTTGGGATGCTTT  
AAGTATGCAAGAAACCAATGCTTCAATAGTAAGAAATATTTCAAGATCGAATTAGGTCAAAAGTGTAAATCGTAAA  
AAGTAAGAAGAGAGAATCTAAAGTACTTATTTGTTCTGCTCTTTCTTTTAAATATAAGCGGCTTGGAGATTATA  
GGAGGTTTGATATAGACCTCTACCATGTAAATTTCTTATTTATTTGCTTTTGTGTTAGCTTTTAACTTTTCCCTGTA  
CTGAAATAGATTTGAATGGCTTAAATATCGTGGAGAGCAAGTACTAGGTGGTATCTATACCTGGGATTAATACAAA  
GCATATATTTCAATGACGAGCAAAATGGTACTCTCATCTTAACTTCTAGTTAGAGATATATTTTACCAGTTTAAAGC  
TGCTTAATTTTGTGCTTTCTGTTGACACAATTTGCTATGATAGGTAGAAGAGTACAGTGGGCTTAAAGCTGCAG  
CCACATTAAGGCCATTTGCGGAGCTAATGCCATTTCTCATGAGAGTGGCATCCATCAGGTTTTCGCTTATTTTCTCT  
AAAACTTTTACTTCTACATACACTAGATTTGTTGAAGATTACGTGAATTTGGCGACCTTCTACACTAGGTTAAATGTC  
AATAACATTTCAATACCACTACCGGTTGATATAAATAAATCCCTTCTCGTTAAGATAAGATGGGTAAATGAAGAA  
TTCAAAGTTGAATTTTCAATATAGAAATGATCATCTTTTGTAAACAGACTAATAAGGAAATTTGTGTCATCT  
AATAACACATGAAGTATTGATTTAGGCCATTTGCCCTTCTCTTCTATTCGCTTAAATCAGGATTCTCTACT  
TCCATAGTATCATTCAGTTTTCATTTCTTCTTTTAAATTTTAAATGTTGTTTAGGATGAGAGTTAAAGAAC  
GAGATACATATGAATTTGTATCTCTGAAGATGTTGGGTATCGTCTGCTAATGAAACGCTATTTAGCTGCGGAAAG  
CTCAGGATGATTTGATTTTCAAGATGATCGATTTGTTATGTTGATGTTTGTATAGAGATTATTTTCAAAATCTCT  
TGGTCTTTCCCCATATTTTGGATGAGTCAAAATAATTCGACTAATATATAGTCTTACCAATTTGGTGTCAACATA  
TGTACGAATGAATATAGTATCAATCAACTATGTTCAACTATCCACTGTAAGGTATCTCCCTTTACATTAATA  
AATCAAGTGAATGAAAAAAGCAATTTTAAATTAAGTGTGCTATCTTAACCGTTTAAATTTTATAGATTA  
GATTATCATACAGTTCAAAACCTTTAAACTCGTACTATCCCCTAGAATTTCAAGTACCTCCGTTGATTTGAATCAAC  
CCCCACCTTCAATTTTCTCATTTACATTTAGGCACATTTGATGATCTTCTTCTTTAAGGAATACATTTTGTAAACAT  
GGTAGAATAATTTGTGAGTGGCGCCATGCAATGAAAGCCAAATGGCTGAGGTAAGATACATAATCTATAGTACG  
ATATATACAAATATTTAGTATATGGAATGGATCAATCATTTTATTTTGTACAGCTTGGATATGACTTTGAT  
GGAAAGAACCTGATGACCTCTTTCGTCGATTAAGTCACGCTGAGAGGAAAAAGTGATTTACGTTGCTGATTT

TCATAGTGCAACAGTCACAAATTTGTGATCGCACCTTTTCGTTCTAGTGTGTTGCCATATTTTGTGAGTTAATTTCAA  
ATATGATCACTAATTTCTACTATAACAATAAAATTTAAATTTTATTCGCATTAAGTGACTGAACATAACAATTA  
AAAAATTTTGAAAAATATATGAAAAATTACGGCCAAAGAAAACTCTTTGACTCTCGACATCCGAGCATCACCACA  
TAGTAATAAATGAGAAATTAATTTTATCAAAATTAAGTAACTGAATTTACTTATATTCATAATAAAGAAAAAT  
CACTAAACTATACGCATATAGTGGCAAAACCTATCCATATTCAGGTGACTTTAAGTACCGGTCTAGTAGGTAGAT  
TTTTGTGATTTTACTGTTATAGTGGATAAAATTAAGTCACITTTTGTGATAAATGGAGAATAGCCTGACAGACA  
TCTCACTTATTCGGCTTTGACATCCCGTCCCTCTACTCCATATAATTTCAACCAACACTTAAACTTGTATAAAC  
ATTAGTCTTAAACACCTCTGATACTGTGTGTGTTGTTCACTCGTGTGATATAAAGGACACATCAGATCCACCTCAG  
ATATATTAATGCCACATCATTATCTATTTTACAAAAATATTTTAAATAAAATATCAAACTCATCCCTTCTCTC  
TCCAAAAAACAACCACTTATCTCTCTCCACCATTCATCTCGCTCTCTCCACCGCTCACCGCCACTACCA  
AACACTGCACCTTCTCTCTCCATGAGAAAACTCAATCAAACTACTTCACTTTTATTTTGTCTTAAACTCT  
CCATTACTACATCTATTCAACCAACCATGGTAGTCTTGTGCCACCATCGCCAGCAACCACTCCAAAAACCACT  
ATATCTCTCTCTCTCTCATATTCAATGTCACTCTCTTCAAACTTAGACATGAACGAGTTAATAAATATCTATCTCT  
AAATTTCTTCTTAGCCAAATACCCAAAGCAAAACCAAAAAATTTGAAAAAGAAAGAGAAATTTTACTCAAAA  
TCTTATTTGACCCCTCCAGCAACGATAAAATTTATAAACTAACAAAAACAAGCTTCGGCGTTTCAACAACCGATG  
ACGATCCAAATTTCTGGCTTAAACACTATTTCAAGCTTCAATGTCTCTTAAATCTCCGGCGCTCTCAATGTCTTCT  
CAGTTATTTGAGAAAAACAAAAATTAATAACTTTGTTCCITGACTAGCTGCTCTGCTTAACTCTATTTAGTTTGGT  
CTTGTTCAAGTCTCTCTTTTATTTTATCATAGTCCAGTTGATTGCTAAAAATGTAAGAGAAATATTAATGTC  
TGTTTGAGAAAACTCTACTTGAGAAAAAATTTAATAATAGAGGGTGTGGAGCGTGTCAATTTCAATGGCGTTTT  
AAGAGATTTGGGGTAAGAAGAAGCGGCTGGGATAGAGAGGGGCGAGATAATGGCTAGCATTTGTTTTAATTTTTT  
TCTTTTCAATTTTCTCTTTTATTTTCTAATCAAAATTTGTATAGTTACTCGCTTTAGGAGCGGTGCAATCACT  
TATTTTGTGTTGACTCAGAAAAACAAAGGCTACATAAGTTTGGTCAATGGTTAGTAGGTTTAAACTTATGTCTTATA  
CAAGTTCAAGTGCCTGATTGAACTATGTGGAGTAGGGTGTGCGAATGTCAAAGCGAATAAGTCGAGGCAATGGC  
GAAGCCAAGAAATCAACAAGGGTATTCAAACCTTTGCCAGTGGGCTATGTAAGGGTGTCAAAGCCTATTTTTAAT  
CAATAACAGTAATATTTTACCTTATACGGAGTATAATTTCTGGCAAGGGTGTGAGTTGACCAACCTTGATTTGC  
ACGTAGCTTCGCCCTGAGTCGAGGTGTCTCTTAGGCTAAAAATAGCGTGATAAATTTATAATAAAATTTATTTTAT  
ATGAGTAAAGTTAATAAAATTTATTTATATGAGTAAAGTTAAGTAAATTAATACTTAAAGAAAAATGAAGCTT  
ACATTTGCAAGAGTCCCGTCTCGACCCACTTATCTTGATCCCTCTTCTCTCTCTTCAAAATATGCGCTCTCT  
GAATAGCTCATGACTCATTGGAGCCCATACTCTCTTTCTGAAAAATAAATTAATTTAGTTAATAAATAAAATA  
GATATACTTAGAAAAAGTAAATAAATTTAAGCAATACGTACCAATCTCTCTTTTATTTGTCCTTAGGCTGTGCGAC  
TCTAGTTGCAAGAGGCTCCCTTCTCGGATGCGGAGCTTTCTCTCTTTCTGCTCTCTTAAGAACTTTGCGG  
TAAGCCATTTGCTTTGCAAACTATTCCCAAAATCTCAAGTAACCAAGCCAGGCTCTTGTCTCTCTTCTAGCATCC  
GAGAAAGCATCCGCCAATCTCTTACTAGCTTTGTGATGAAAAATTTGAGCCACTTCCGCGCTATAGCGGTCTTTCCA  
TACACACTTGTCTGTATTTTAAAAATTTAATTTATAGATGGAACATCAATAATAAATTTAATAAATCTGTTAAAA  
AAATATTTATACCTTAAATTTGATTGAAAAATTTGCTCTTCTAGTGAGAAATGGACAATCAGTCCAAAGTCGTATAAGGGC  
CATCATAAAGCTTTCTGATGGCATTAGTGATTATCTTCTGAGTCTTATTACCCGGCTGAACCTACAATTTAATAAT  
AAAAACACATTAATATCTTAGTAAAAATGTTACAAAACAATAGACGGAATAAATGAATAACACTTACCATCACC  
CTTAGGACTATGATGATCTGCCATATCGATCATAATGCATACCTCGTCACTATTCGAGCATGTGTATCAG  
AGGCATGTGAAGACGGTGTAGGCGGGTCAAGGCTACTGCTCGTAGGCGAAGGCTGCAATAGATGGAGTCGATAAT  
GAAGATGGTTGTATCCCTGTGACTACAACATAGCTGGATGTGACCCAAAGTGGATGTGATAATGCTGGCTGTGACCC  
GAGTGGCTGTGACATGGATGGATGCCATCCATGTGGCTGTGATATGTAGGGATGTGATCCATATGGATGTGATG  
ATGGCTGCGATGCAATGGCTGTGAGGACGTGAGCTGTATGTGCAACATATAGTCTATGGCCCTGTGATGGAAGA  
CTGGTGTCTGAAGCAAGGTGTATGATTCGTGATGCTGTGGAACACTCAGTATAGCCGTGTGGTGAGGATAAGACAT  
AGGCATCTCTGAAAAAGGTGGACAAGAGGACTATTTATTTCTACCTCTCTTTCCCTTCTTACCTTTTCTCGAC  
CTTGAGAACTAGTAGGGTCAATTTTACCTTGACCTTGCCTGCCATCTGTATGATAAATGCACATTTAGATTGAAA  
CATATAAAGAACTAGCTATAAACAAGAGTGTCTTAAAAAACAACCTTTTAACTCTCATCGACGATTTGTCTCTCG  
TCCGAGAAATCTTCTGCTCTCATTTGTTTGAGCTTCATCACTTGATTCTTCTGCTCTCATTTTCTTAAATGTACTTC  
ATTTATATCACCTTCTTCCAATATGCATTCAGGATGTTCCAAATCATTTTCTAATGATGCTGCCACTATTTGGTGAA  
TATTTAGAAATATCATTTTGATATGCAATATTTAACACATCTCGACTTCCACCTACCTACATGCTTTAGTTTATTT  
ACAACCCCAACATCGACTTATTCGCCGCAATGGATAAGGAGCATAATACACTTGCCCTAACGTTATGTGCAATTAAT  
GAAAGGATCATAGCGATCATACTCCCTCGTATGATTAACTCAATTTATGTTGATTTGGTGGTGACACTTGTACCTC  
TTGTTGGATTTGGGTCAAACCACTTGCATATAAAGAGTATCAATTTCTATATGGCCAACTGTATATTTCACTGT  
AATATTTCTTTGACCACACCAATAAATCAATATCTCCAATTTGGTGGCTACCACTTTGAACCCACACCCCGCT  
GTTGTTACTATTTTATTTTATAGGCCATCTTTGTATGAACTTATAACCATCTACTACGTACTTAGACATTTGTTG  
TGACCTGAAGCCAGGTCCCAAGATATATCTTTCAAAAATGATTTACACATTTATTTGGATTTATTTACCTACATA  
GTGTTAAAAATATTCATAAGTTAGCATAACTTATAAATCAATTTTATACATTTATTTAGATATATATATATATAT  
ATATAACTAGTGATATATATATACAACTGTTTGAACAGTATCAATCTCGTATATACAACATCATGGCCAAAT  
GACCTACGAAGTACTGTGACACATCAAAATTTTATAGTGTGATCAATATGATGACAGTAAATTTTACATTT  
TGATAACTAATAAATCAATCTTACTTTGAGAAATGGTACAACCTCGGAACAATTTAGCAACATGAAGTGTAGCTG  
ACTTGTACTCCATATCACTCAACTTCTCTTCTAATCATCTTAGAACATCGGCTGGTGTGATTGAATATAGATAT  
GGTGGATATAATGATCATTCATATTCGACCTGTGCTTATTGGGCTTATCTTAGAACATGACGCTTATCTCT

SEQ ID NO:7 (NtIPMS2 corresponding BAC contig sequence (total BAC length is 10082bp))

TAAACACGCTTCGTTGCTACTTAACTACATATATAGCATATTATATATAGTATATGTTAGTGTATTCATTAT  
TGTATTACAAAAGTGTAAACGAATTACATTTATATTTAGATAGTAAATATAAAAGTAAATCGAAAGTTTGACCA  
ATGTTGACGTAATACCGACCAACTTTGGTCGGTTATTTACAGAAAATATCAATTACCGACCAAAAGTTGGTCGGTA  
AATACTTTCCCTGCATTAAACCGACCAACGTTGGTCGATAATTTAAATATAAAATCTTAAATATTATAAAATATTTT  
AAATAATAAAATATTTATTAATTTTAAAAATTTGGGTCAGATTACCGACCAACGTTGGTCGGTAATCCAAAATTT  
TCTTGCTGACCAAGCTAGGTCAATCGTTGACCAATTTGCCGACCAACGTTGGTCGGTATTTAGTTC/AAAAAAAT  
GTAAATTTTTTGTTCAGTTTTTGTCCAAAGCTGGACGGTTTTTGGTCGGTTTTTTTTTACCGACCAACGTTGGT  
CGGAAATATTTGGTCAGTTTTTATCAGATTTTAGTAGTGTCAAACTTTAAACTCGTTTTATCCCTTAGAATTT  
CAGTACCTCCATTTGATTGAATCAACCCCAACCTCCATTTTCTCATTACATTAGGCGCATTTGATGACCTCTTT  
CCTTTTAAAGGAATATTTTGTAAACATGGTGAATATTTGTGACGTGGCCACCATGCAATTGAAAGCCAAAATGGCTG  
AGGTAAAGATACATAATCTATAGTACGATATACTACAAATATTATAGTATATGGACTGGATCAAAATCATTTTATTTT  
TGTTACAGCTTGATATGACTTTGATGGAAAAGAACTTGATGACCTCTTCTCGATTCAAGTCACTGGCTGAGAGG  
AAAAAGGTGATTAGTTGCTGATCTTTCATAGTGCAATGGTCAAACTTGTGATCGCACCTTTTCGTTCTAGTGT  
GTTTCATATTTTGTAGTTAATTTAAGTATGATCACTAATTTCTATTATAATAAAAAATTAATTTATTTATTTCA  
CATTAATAATGACTGAATTAACAATAAAAATATTTGAAAACATATGAAAATTTACGGCCAAAGAAAATTTCTTTG  
ACTCTCAACACATCCAAGCATCACCACATAGTAAATAAGTCAAAACATATTTTATCACATTAAGTAACTGAATTT  
TACTTACTATTATAATAAAAGAAAACCACTAACTTTACGCACTATAGTGGAAAACCTATCCATATTCAGGTGA  
CTTTAAGTACTGGICTAATAGGTAGATTTTGTGATTTTACTGTTATAGTGGATAAAATTAAGTAACTTTTTTGAT  
AAATTAGCAGAAATAGCCTGAGTGACATCTCACTAATTCGGCTTTGACATCTCGGTCCCTTACTCCACAGTTTCAAC  
CAACACCTTGAACTTGTATAAAATATTTAGTTTAAACACCTCTGATACCTGTGTGTTCCTCACTCGCGCTGATATAAA  
TGACACATCAGAGTCACCTCAGATCAGATTAATGCGCCAGTCATTATCTATTTTACAAAAAATATTTTAAATAAA  
AATATCAAACTCATCCCTTTTCTCCAAAAAAGAAAACCAACCTTAGTCTCTCTCCACCGTTCCATATCTGCTC  
TTCCACCATCACCGCCACTACCAACACTGCACCTTCTCTCTCCATGAGAAAACACTCAATCAAAACACTTTTATC  
TTTTTTTTTGTCTAAATTTCTCCATTACTACAATCTATTCAACCCCATGGAAGTTCTGTTGCCACCATCGCCAG  
CAACCACTCCAAACCCCACTATATCTCTCTCTCTCATATTCATGTCACCTTTTCAAAACCTAGACATTTGAAC  
GAGTTTAAAGAAAATTTATCCTTCAATTTCTTTCTAGCCAAACACCAAGAAAACCCCAAAATGAAAAGAAAG  
AGAGAAAATTTTACTTAATATTTGGTACCTCAAAATCTTATTAACCCCTCTAGCTACGGCAAAATTTATAAAAT  
AACAAATCAAGCTTCGACGCTTTCAACAACCGATGACGACCAATTTTGGCTTAAACACTATTTCAAGCTTCAA  
TGCTTTTAAATTTCTCCGAGCTTTTATTTGCTTTCTTCAGTTATTTGAGAAAAACAAAAATTAACAACTTTGTTCT  
CTTGACATGCTGCTGCTGAAGCTCTATTTAGTTTGAATGTTGATGCTACACCTTTTGTGTTTTATCATAGTCCAG  
TTGATTTGCTAAGATTGTAAGAGAAATTAATGTAATGTTATTTGAGAAAACCTCTACTTGAGAAAACAAAAATTTAATATT  
AGAGGGTGTGGAGCGTGCCAAATTTCAATGGCGTTTTTAAGGGATTGGGGGTAGAGAGAGCGGATGGGATGAGGG  
GCGAGATAATGGCTAGCATTTGTTTTAAATTTTTTAAATTTTTTATTTTCTCTTATATTTTCTAATCAAAATTTG  
TATAGTTTACTTGCTTTTAGGAGTGCAATCACTTATTTGCTTGACTCAACAAACAAAGGCGGTATAAGCTTTGGT  
CAATGGTCATATGGGTTTAAACATATGCTTTATACAGTTCAAGTACCTGATTGAAACTATGTGGAGTAGGGGAT  
CGTAATGAGCAAGCCGAAATAGTCGAGGTGCTCTTAGTCTAAAGTAGCCTTATTTTATAGTGAGTAAAGTTATGTC  
ATTATATTTATATGAAATTAACCTATAGGTTTTCCAGTTCTTTGGTGAAGCTTACCTAATTAATAATTTTACTG  
TGTCGGTATAATGAATATGATGTCGGCTATTTATCTTCCCTACATCTATTTTTCAGAAAATACAGATGATGACT  
TGAGAGCACTTGATCAGATGACGTTTTCCAGCCTCAAGTTTCTGGCAACTTGAGATGTACAGGTACATGTAAT  
ACTATTTCTATGTTTTTGCCTAAGTTGCTCTGACACGGCAGTTAGGTGACACGCGCTGTGCGCAAGACACTAGT  
ATGAGTGTGGATGTGAGATCCGTACCGGATCTGGTCMAATAATTTTGGGTACTTTGACCGCAGCGTACGAAAAAAT  
AGGATAAGATACAAATTTGATTTCCGAAATCAGAACCAAGCTAGGGTGAATTTGAACAAAAATAGCATACATATCT  
AGGAAATCAATCCTTTACTTTATCTACAACTCGAGAATAAAAAATAAATCCACACATTACAAGTTATACGTTAAGTAT  
CCACAAAATTTCTCGTAATTTAGATATATTTTATATTTTATTTAAATGAATPATTTTATGTCGGATCCCCGACC  
CGTATCCGTATTAGGATCAGTATCCCAAAATCTTAGAATTTACATCTCGAAGGATCCGACCTCTAGATCTGCACCTG  
TATCGGATACCCGTACCCGTATCTGAGCAACTTAGGTTTTTGTAGTCTCACTTAGTTGTGGATTGAAAAGGAATTA  
ATGTTCACTAAATCTAATGATTGCTGTTTTATGTACAGATTACTTTGGAAATGTTGGCGCTCTACAGCAATG  
TTAAGCTTATGACAGCGATGTTCAAGAGCACACTGCTTTTTCTGTTGGAACAGGACCTGTGATGACGCTTACAAG  
GCGTTGACCTCATTTGTAAGGTTTATAATGAACTGAAACCTCATAATGTTTGTATCTTGTCACTAGTAGGA  
TAGGTAATTTAAGTTTTATATACGCTGAAGGTATAAAAAATTTACTCAATCAGAGCGCTTACAAGACAATTTGAGG  
TAACTCTATGAGTAGCTTAGACTAAAAGTTTTAATTTATTTAATTATATTTATGATGCTCAACACGTTTACTCACC  
TCTTAGCTAATTTTGTACATGGGATCGTTCTACATTTTGTAGTGTGAGGATTGAACCTTAGGACCTTTATTT  
TCCTAACACTATGTTATATTTGCCAGTGGAAACAGGAGCGCATATGGACATTTGTGTTCAAGTCTCAAGCCTATGT  
TGAGGCGTTGAACAAAATATTCAGTTTCAAAAAACAGGTCCTGTAACAAATTTGAAGGCACTGCGCAATCGTAA  
AGTGATGGTGTGCTCCAGAAAACAAAGCTTAAATGCTGCTTAGTGCTGATGATCATAAATTTATGCTCTATTCA  
TGATGCAATTTGAGTTAATTAAGAAAGTAACTTAATTGATTAAGATGTAACATGTTTCTATTCTAGTCACTACT  
ATGAAATTAATTAATAATTTCTGAGCCAATTTTTTATTTGTTCTTAAATTTATACAAATCTCTTTATTTGTATTT  
GACTAACAGATTTTCGGGAATTAGGGCAAGTAGTCCAAATCAGAATCACAATTAAGCTTGGAGAGTAAATCAAAATCT  
GCCTCTTGAACAAACCTTACTCAAGTACCTTTAACATATTTTAAACATGAAAAGGAGCAGCCATATGAGAAGC  
AGCCATATGAGAAGCAGCCATGTGCGGTCTTACATTTGCTTGAAGCTGCAGGTATTTCTGATGTCACAATTTGATC  
TGACGTGCACCTTCACTTCTACGATCGCAAAATTCATGTCAGTCCGCACTTTTGCAAGGCTTCAGTCTTGTCTCA

80

GTCTCAAAATTACATAATTAACCTATTCCAACCTTTCAGAATAACAATCCAGACCCGACAGCCTCAATGTCAACTCCCG  
 GTAAAACCTATGAACCTTCCAAACCTTTAAATTTCCAACATTCGCCAAATAGAGCCAAATAAACCTAGGAACCTCCA  
 AATCTAAATCCGGACATACGCCTAAGTCTAAAATCACAATACAATCTATTGGAACCATCAAAATACCATTCAGAA  
 TCATTTTATATAAAAGTCAAAATACGCTCAACTCCTATAACTTAAGCTTGCATGCCGTCAGGTCGACTCTAGAGGAT  
 CCCCAGGTACCGAGCTCGAATTGCGCCTATAGTGAGTCGTATTACAATTCAGTGGCCGTCGTTTACAAACGTCGTGA  
 CTGGGAAAACCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCGAGCTGGCGTAATAGCGAAG  
 AGGCCGCGCAGGATCGCCCTTCCCAACAGTTGCGCAGCTGAATGGCGAATGGCGCTGATGCGGTATTTCTCCTTA  
 CGCATCTGTGTGGTATTTACACCCGCATATGGTGCACCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAG  
 CCCCCGACACCCGCCAACCCCGCTGACGCGAACCCCTTGCAGGCCGATCGAATATAACTTCGTATAATGT

SEQ ID NO:8 (Promoter sequence used to drive trichome specific expression. This promoter was isolated from *Nicotinia tabacum* and is located 1.5kb upstream of the gene encoding for the copalyl-synthase involved in the synthesis of *cis-abienol*.)

agcttctgcgaatctcccaacattatcaccttctgcccgcctctgaaagcaagtagctaaagcattaattaagaatgac  
 cacacaaaaataaataagtagtagtaaaagtagccgaactacataataatagacccatcatataggtttcttagccaata  
 cttttccaattaaagattaggttttcccttttaaaaaatttgcacaattcttagagagatctaatagtgcaaaacaca  
 gaaatataatccaaactacctttctctctccttaaacattttatctaactaacacggcactagttgaaacaaatca  
 ggggatatttgaaggttaataaatgactgggtgatttttaaacgttagatatgttgaaataaattcaatttgaaaa  
 aacgactaataaataaagctggaacgctacgtattcaacactaaagaaataaagtgtctatttgacaatatatgaa  
 gtcaagaaataaagactggcattatataattttcaagtaggtttaggctatggcaaaataactaataagcaagcact  
 taattttgcggtacaaaataaagttttgtattgttaataaagaagatatcaactgaacaaaataagagagatttgcc  
 tatagtttaattgcatcgctcgctcctttgtgagcatttcaataggcttatgatcacacataaatttggtgtgaatt  
 gcttttagaaaaattacataatttgaaatttgaggtcttaatatgtgttcaatccagaggagtaggcacottagctcga  
 gcgatatcgctgacacccgcttcgacaacatatttcataaataatgtatgtataaataatgaagaaaaataaaatatt  
 gagtataaataataaagatgacattgcaactctcttaaatattgataccgcttcaaaaactctcttgccgcacgcca  
 ttgatccaatctatcatagttcttgaataatgtatttccagatcttcaatttaaaatattttaaataaagtcacatct  
 ttaacatctagattttctcgtttttactttttgtttattatatacaattttggacatagcactaagtcggtataact  
 aattgtgactttgtcaagtttaagaaatacaaatgcaatgctgctgaacacacagagcaactcgtttccagtaaaaat  
 ctaaaagtttactactttcaaaaaactaataaaagtttttagagtgccgtttgacaattatttcatgcaaaagattcga  
 gaacaatcaacacagaattaggctgaagtgctaaagagaatttaataatttgcccttcataagaggcaatatataaa  
 taagcctgcgcatttgcaacaactcaaacatttccaatattgcctcaaaagtcagtagtgcccttccctctcctaaa  
 cgttcattgtctttatctcccttcccccaattctcattggaagaataaaacaaaataaattagaaaatg

SEQ ID NO:9 (Synthetic construct used to trigger the suppression of NIPMS2 gene expression via the RNAi mechanism. Composed of a fragment of NIPMS2 first intron plus the first exon, the fragment of first intron is repeated invertedly after the exon. Translation and splicing of this construct theoretically results in a double stranded RNA in hairpin structure.)

gctaagcttctgcgaccatggaaattctctgacccgtgattatgttggtatttttgacacagctcttcgcgaaggcgaa  
oaggcgcgtggtgctaccatgactagtaaaagaaaaactggacattgcaagtcagttggcgaagcttggtgttgatgt  
tatttgagccgggttttcccttttgcctctgaagctgagttcgagcttgaagttgtagcaccaggaattggtaacc  
tttaattgtttaaccgttccacatttctaataatttacttatttgaacatgctgctcagctgttagtttccattctttta  
tgaaccaaactgcatgcaaaagataatttttgatatttggacggcgagtgagatttgaactaggacccgtttgcctg  
atacaataataaaatagtaaccattttatgtacaagtttaactgttgatagtagcataatttttacttttattta  
agtatactataattccaacaggtaaGTTAACCaatttccctgtgctatcaactttacaagctgaactcagcttcagag  
gcacaaaggaaaaacggccctcaataacatcaacacaaagcttagcgaactgacgtgcaatgtccagttttttcttact  
agtcattgttagcaccagcgccctgttgcacatgcggaagactggtgtcaaaaataccaacataatcagggtcagaga  
 atttgccgcgccaattgacc

Bold and underlined: exon 1,21-229

Bold: splicing border

Italic: splicing site

Italic and underlined: intron 1

SEQ ID NO:10 (NIPMS1A nucleotide sequence)

ATGTCCTCTCTCTGTTCAAACCTCTGCAACTTCTCTTAGTTGTAATAACCTCCTTCCAAATCCAAAAATCCTCTTCTTCA  
 CACCATCTTTAACTTCTTTCCTTCAATCAAAACACACTCTTGTTCCTCATATACAGTTATCCGGTGTCTCAATTCAAA  
 AGCGAAGCTGAATATATACCGAGTAAATCTCCGACCAAAATACGTACGCATTTTCGACACGACTCTCCGGGATGGA  
 GAGCAATCCCCGGGCGCTACGATGACTACGAAAGAAAAATAGATGTTGCCGTAAATAGCGAAACTCGGAGTTGA  
 CATAATCGAAGCTGGGTTTCCAGCTTCACTCTGAAGCTGATTTTGAAGCTGTGAGATTAAATAGCAGAGGAAATGGTA  
 ATAATAGCGACGGTGATTATGTGCCGGTGATTGTGGATTAGCGAGGTGCAATAAGAGGGATATTGATAAAGCGCTGG  
 GAAGCTGTGAAGTGTGCGAAGAGACCTAGGCTTCATACGTTTATAGCGACGAGTGAGATACATATGAAGTATAAGTT  
 GAAGATGAGTAAAGAAGAAGTAGTGGAGAAAGCGCGGAGTATGGTTGCTTATGCGAGGAGTTTGGGATGTGAGGATG  
 TTGAATTTAGCCCTGAAGATGCTGGAAGGTCTGAGCGGTGAGTTCCTTTACCATATCCTTGAGAAAGTTATCAAAGCT  
 GGTGCAACAAACCCTTAACATACCTGATACCTGTGGATACACTGTGCCACTGAATTTGGACAATTAATTGCTgacat  
 aaaagccaataccccaggaattgaaaatgtgatcatttctacacactgccagaatgatcttggacttctactgcca  
 acacttttagctggagcttgtgcaggggcaagacaagtagaagtgaccatcaatggcattggtgaaagagctggaaat  
 gcttctctggaggaggttgaatggccttaaaatgtcgtggagagcaagtagactaggtggcctgtatcacaggaattaa  
 tacacaacataatactcatgtcaagcaagatggttagaggagtacacgggctcatgtgcagccacacaaagccattg  
 ttggagctaatgcttttgcctcatgaaagtggcatccatcagGATGGAATGTTAAACACAAAGATACATATGAGATT  
 ATATCTCCTGAAGATATTGGGCTTAGTCGTGCTAATGAAGCGGTATTGTCTTGGGAAGCTCAGTGGCGCCATGC  
 ATTTAAATCCAAATGCTTGAGCTTGATATGACATTGAGGAAAAGAACTGGAGGACCTCTTCTGGCGATTTAAGT  
 CGGTGGCTGAGAGAAAAAGAAATACAGATGATGACATAAATAGCACTGATGTCAGATGAAGTTTCCAGCCTCAA  
 GTTGTGTGGCAACTTGCAGATGTACAGATTGCCTGTGGAAGTCTTGGCCTCTCTACAGCAACTGTTAAGCTTATTGA  
 CAGTGTGCTCAACAGCATGTTGCTTGTCTGTGGAAACCGGACAGTGTGATGACGCTTATAAGGCAGTGTGACCTCA  
 TTGTAAGGTACCTATAACACTCCTCGAGTATTCCATGAATGCAGTCACAGAAGGTATAGATGCCATAGCCTCAACC  
 AGAGTATTAAATCCGCGGGAGGATGACCATGCTATAACCAATGGTTCAATTTGACCGGACTCATCACCCTATATTAG  
 TGGAACTGGAGCTGATATGGAGCTTGTCTCTAGTGTCCGAGCCTATATTGGTGCAATGAACAAAATGTTGAGTT  
 TCGGGAAGCTGCTTTCGAGGTACAAGAAGCCTGAAGGTAGTGTGGTAGTATAA

## SEQ ID NO: 11 (NIPMS1A amino acid sequence)

MSSLCSNATSLSNNSFQSKNPLLHTIFNFFPSIKPHSCFPYTVIRCSIQKRPEYIPSKISDNHYVRIEDTTLRDLG  
 EQSPGATMTTKKLDVARKLAKLVGDIIEAGFPASSEADFEAVRLIAEEIGNNSDGDYVPIVCLARCNKRDIKAW  
 EAVKCAKPRVHTFIATSEIHMKYKLKMSKEEVVEKARSMVAYARSLGCEDEVFSPEDAGRSEREFLYHILGEVKA  
 GATTLNIPTVGYTVPTEFGQLIADIKANTPGIENVIISTHCNDLGLSTANTLAGACAGARQVEVFTINGIGERAGN  
 ASLEEVEVMALKCRGQVLGLGLYTGINTQHILMSSKMVEEYTGHLVQPHKAIVGANAFHESGIIHQDGLMKHKDYYEI  
 ISPEDIGLSRANEAGIVLGLSGRHALKSKMLELGYDIEGKELEDLEWRFKSVAEKKKKTDDDDIILMSDEVFOPO  
 VVWQLADVQIACGSLGLSTATVKLIDSDGQEHVACSVGTGPVDAAYKAVDLIVKVPITLLEYSMNAVTEGIDATAST  
 RVLIRGEDDHATTNGSIGPTHHRIFSGTGADMDVVISSVRAYIGALNKLMSFGKLVSRYYKKKEGSSVVV

## SEQ ID NO: 12 (NIPMS1Bv1 nucleotide sequence)

ATGGCGTCTATCACCATAAACCATTCATTTTCCCGTAACCTTAACATCTCATTCATCCATCCCAAAATCCTCTCATTTCA  
 AACCCAGCTCTCTTCAATTTCAAACCATCAATCTCCAAATGTTCCCTTATTATCCACTGCGCAATCCGCGTCGAC  
 CCGAATATACCCCGAGCCACATTCCCGACCCGAACACTACATTCGCATCTTCGACACCACTCTCCGCGACGGCGAACAA  
 TCCCCAGGCGCCACAATGACCACAAAAGAAAACTCGACGTTGCGCGTCAGTTAGCTAAGCTTGGTGTTCACATAAT  
 TGAAGCCGGTTTTCTCTCTTCTTCTGAAGCTGATCTCGAAGCTGTGAATTAATAGCGAAGGAAGTTGGAATGGTG  
 TGAATGAAGAGGGACATGTTCCGGTAATTTGTGGACTTGCAGGTGTAATAAGAGGGATATTGATAAGGCTTGGGAG  
 GCTGTGAAGTATGCGAAAAACCGAGGATTATACGTTTATTCGCACTAGTGAGATACATATGAAGTTTAAGTTGAA  
 GATGAGTAGAGATGAAGTTGTGGAGAAAGCTAGGAGTATGGTTGCTTATGCTAGGAGTATTGGTTGTGAGGATGTG  
 AATTTAGCCCAAGAGTCTGGAAGATCCGATCCAGAGTCTCTCTATCATATCCTTGAGAGGTCATCAAGCTGGG  
 GCAACACACCTTAACATCCCTGATCTGTTGGATACACTGTCCAGCGAATTTGGAAAAATTGATTGCTGATATAA  
 GGCAATATCCCGAGGAATGGAGATGTGATCATCTCAACACACTGCCAGAACGATCTTGGGCTTTCTACTGCCAACA  
 CCTTAGCTGGAGCATGCGCAGGTGCAAGACAAGTAGAAGTGACCATCAACCGAATCCGTTGAAAGAGCTGGAATGCT  
 TCTTTGGAGGAGGTTGTAATGGCTTAAATGTCTGTGGAGAGCAAGTACTAGGTGGCCTGTATACAGGAATTAATAC  
 ACAACATATACTCATGTCAAGCAAGATGGTAGAGGAGTACACCGGCTTCATGTGACGCCACACAAGGCCATTTGTG  
 GAGCTAATGCGTTTGTCTATGAAAGTGGCATCCATCAGGATGGAATGTTAAACACAAAGATACATATGAGATTATA  
 TCTCCTGAAGATATTGGGCTTAACCGAGTTAATGAATCTGGCATCGTCTTGGGAACTCAGTGGGCGTCATGCTTT  
 GCAAGCCAAAATGCTCGAGCTTGATACGATATTGAGGGAAAAGAACTTGAGGACCTCTTTTGGCGATTCAAATCTG  
 TGGCCCGAGAGAAAAAGAAATTAACAGATGATCACCTCATAGCATTAATGTCAGATGAAGTTTCCAGCCTCAATTT  
 GTTTGGCAACTGAAAAATGTACAGTTACATGTGGAAGTCTTGGCCTTTCTACGGCAACTGTTAAGCTCATTGACGC  
 TGATGCTCAAGAGCATGTTTCTGTCTGTGGAACGGGGCCAGTTGATGCGGCTTATAAGGCAGTGTATCTCATTTG  
 TAAAGGTACCTGTAGCACTCCTTGAATATTCTTGAATGCAGTCACGGAAGGTATAGATGCCATAGCTTCAACCCAGA  
 GTTTTAATTCGGGGAGAAATGGCCATACATCAACCATGCTTAACTGGAGAGACTGTACACCGTTCTTTTAGTTG

AACCGGAGCAGATATGGATATTGTTATCTCCAGTGTCGAGCCTATATTGGTGCATTGAATAAGATGTTGAGTTTCA  
GAAAGCTGGTATCGAAACACAGCAACCTGAAGGCAGTGAGTCGTATAG

SEQ ID NO:13 (NiPMS1Bv1 amino acid sequence)

MASITINHSFSRNPNI SFHPQNPLIQTQALFNFKPSISKCSPIIHCAIRRRPEYTPSHIPDPNYIRIFDTTLRDGEQ  
SPGATMTTKEKLDVARQLAKLGVDII EAGFPASSEADLEAVKLI AKEVNGVNEEGHVPVICGLARCNKRDIKAW  
AVKYAKKPRIHTFIATSEIHMFKLKMRSRDEVVEKARSMVAYARSIGCEDVEFSPEAGRSDFEFLYHILGEVIKAG  
ATTLNIPDTVGYTVPSEFGKLIADIKANTPGIGDVIISTHCQNDLGLSTANTLAGACAGARQVEVTINGIGERAGNA  
SLEEVVMALKCRGEQVLGGLYTGINTQHILMSSKMVEEYTGHLVQPHKAI VGANAFHESGIHQDGM LKHKDTY EII  
SPEDIGLNRVNESGIVLGKLSGRHALQAKMLELGYDIEGKELEDLFWRFKSVAEKKKKITDDDLIALMSDEVFPQPF  
VWQLENVQVTCGSLGLSTATVKLIDADGQEHVSCSVGTGPVDAAYKAVDLIVKVPVALLEYSLNAVTEGIDAIASR  
VLIRGENGHTSTHALTGETVHRSFSGTGADMDIVISSVRAYIGALNKM LSFRLVSKHSKPEGS AVV

SEQ ID NO:14 (NiPMS1Bv2 nucleotide sequence)

ATGGCGTCTATCACCGCAAACCATACATTTTCCCGTAACCTAACATCTCATTGCATCCCCAAAATCCTCTCATTCA  
AACCCAAAGCTCTCTCAACTTCAAAATCATCAATCCCCAAATGTTCCCTATTTATCTGCTGCGCAATCCGCGTGCAC  
CCGACTATACCCCGAGCCACATTCCCGACCCGAAATACATCCGCATCTTCGACACCCTCTCCGCGACGGCGAACAA  
TCTCCAGGCGCCACAATGACCACAAAAGAAAACCTCGACGTTGCCGCTCAGTTAGCTTAGCTTGGTGTGACATAAT  
TGAAGCCGCTTTTCTGCTTCTTCTGAAGCTGATCTCGAAGCTGTGAATTAATAGCGAAGGAAGTTGGAATGGTG  
TGATGAAGAGGGACATGTTCCGGTAATCTGTGGACTTCCGAGGTGTAATAAGAGGGATATTGATAAGGCTTGGGAG  
GCTGTGAAGTATGCGAAAAAACCGAGGATTTCATACGTTTATTGCGACTAGTGAGATACATATGAAGTTTAAGTTGAA  
GATGAGTAGAGATGAAGTTGTGGAGAAAGCTAGAAGTATGGTTGCTTATGCTAGGAGTATTGGTTGTGAGGATGTTG  
AATTTAGCCCTGAAGATGCTGGAGATCTGATCCTGAGTTCTCTATCATATCCTTGGAGAGGTGATCAAAGCTGGG  
GCAACAACCCCTTAACATCCCTGATATGTTGGATACACTGTTCCAGTGAATTTGGAATTTGATCGCTGATATAAA  
GGCAATACCCAGGAATTGGAGATGTGATCATCTCAACGCACTGCCAGAACGATCTTGGGCTTTCTACTGCCAACA  
CCTTAGCTGGAGCATGTGCAGGTGCAAGACAAGTAGAAGTGACCATCAATGGAATCGGTGAAAGAGCTGGAATGCT  
TCTTTGAGGAGGTTGTAATGGCTTAAATGTCGTGGAGAGCAAGTACTAGGTGGCTGTATACAGGAATTAATAC  
ACAACATATATCATGTCAAGCAAGATGGTAGAGGAGTACACCGGGCTTCTATGTGCAGCCACACAGGCCATTGTTG  
GAGCTAATGCTTTGCTCATGAAAGTGCCATCCATCAGGATGGAATGTTAAACACACAAGATACATATGAGATTATA  
TCTCTGAAGATATTGGGCTTAACCGAGCTAATGAATCTGGTATCGTCTCGGAAACTCAGTGGGCGTCATGCTTT  
GCAAGCCAAATGCTCGAGCTTGGATACGATATTGAGGGAAAGAACTTGAAGGACCTCTTTGGCGATTCAAATCTG  
TGGCTGAGAAGAAAAAGAAATTACAGATGATGACCTGATAGCATTGATGTCAGATGAAGTTTCCAGCCTCAATTT  
GTTTGGCACTCGAAATGTACAGGTACATGTGGAAGTCTTGGCCTTTCTACGGCAACTGTTAAGCTCATTGACGC  
TGATGCTCAAGAGCATGTTTCTTGTCTGTGGAACGGGCGCAGTTGATGCGGCTTACAAGGCAGTTGATCTCATTC  
TAAAGGTACCTGTAGCACTACTTGAATATTCTTGAATGCAGTCACGGAAGGTATAGATGCCATAGCTTCAACCAGA  
GTTTTAATTCGTGGGAGAAATGGACATACATCAACCCATGCTTTAAGTGGAGAGACTGTACACCGTTGTTAGTGG  
AACCGGAGCAGATATGGATATTGTTATCTCTAGTGTCCGAGCCTATATTGGAGCATTAATAAGATGCTGAGTTTCA  
GAAAGCTGGTATCGAAACACAGCAGACCTGAAGGCAGTGAGTCGTATAG

SEQ ID NO:15 (NiPMS1Bv2 amino acid sequence)

MASITANHTFSRNPNISLHPQNPLIQTQALFNFKSSIPKCSPIICCAIRRRPDYTPSHIPDPKYIRIFDTTLRDGEQ  
SPGATMTTKEKLDVARQLAKLGVDII EAGFPASSEADLEAVKLI AKEVNGVVEEGHVPVICGLARCNKRDIKAW  
AVKYAKKPRIHTFIATSEIHMFKLKMRSRDEVVEKARSMVAYARSIGCEDVEFSPEAGRSDFEFLYHILGEVIKAG  
ATTLNIPDTVGYTVPSEFGKLIADIKANTPGIGDVIISTHCQNDLGLSTANTLAGACAGARQVEVTINGIGERAGNA  
SLEEVVMALKCRGEQVLGGLYTGINTQHILMSSKMVEEYTGHLVQPHKAI VGANAFHESGIHQDGM LKHKDTY EII  
SPEDIGLNRANESGIVLGKLSGRHALQAKMLELGYDIEGKELEDLFWRFKSVAEKKKKITDDDLIALMSDEVFPQPF  
VWQLENVQVTCGSLGLSTATVKLIDADGQEHVSCSVGTGPVDAAYKAVDLIVKVPVALLEYSLNAVTEGIDAIASR  
VLIRGENGHTSTHALTGETVHRSFSGTGADMDIVISSVRAYIGALNKM LSFRLVSKHSRPEGS AVV

TABLE 1

Structure of sucrose esters. The general structure of the sucrose ester is shown in Figure 5 and the various combinations of R1, R2, R3, R4 and R5 groups are described.

| R5 | R3 | R1  | R2   | R4   | Name  |
|----|----|-----|------|------|-------|
| -  | C2 | C4i | C4i  | C4i  | C2C12 |
| -  | C2 | C4i | C4i  | C5ai | C2C13 |
| -  | C2 | C4i | C4i  | C5i  | C2C13 |
| -  | C2 | C4i | C4i  | C6ai | C2C14 |
| -  | C2 | C4i | C4i  | C6i  | C2C14 |
| -  | C2 | C4i | C5ai | C4i  | C2C13 |
| -  | C2 | C4i | C5ai | C5ai | C2C14 |
| -  | C2 | C4i | C5ai | C5i  | C2C14 |
| -  | C2 | C4i | C5ai | C6ai | C2C15 |
| -  | C2 | C4i | C5ai | C6i  | C2C15 |
| -  | C2 | C4i | C5i  | C4i  | C2C13 |
| -  | C2 | C4i | C5i  | C5ai | C2C14 |
| -  | C2 | C4i | C5i  | C5i  | C2C14 |
| -  | C2 | C4i | C5i  | C6ai | C2C15 |
| -  | C2 | C4i | C5i  | C6i  | C2C15 |
| -  | C2 | C4i | C6ai | C4i  | C2C14 |
| -  | C2 | C4i | C6ai | C5ai | C2C15 |
| -  | C2 | C4i | C6ai | C5i  | C2C15 |
| -  | C2 | C4i | C6ai | C6ai | C2C16 |
| -  | C2 | C4i | C6ai | C6i  | C2C16 |
| -  | C2 | C4i | C6i  | C4i  | C2C14 |
| -  | C2 | C4i | C6i  | C5ai | C2C15 |
| -  | C2 | C4i | C6i  | C5i  | C2C15 |
| -  | C2 | C4i | C6i  | C6ai | C2C16 |
| -  | C2 | C4i | C6i  | C6i  | C2C16 |
| -  | -  | C4i | C4i  | C4i  | C12   |
| -  | -  | C4i | C4i  | C5ai | C13   |
| -  | -  | C4i | C4i  | C5i  | C13   |
| -  | -  | C4i | C4i  | C6ai | C14   |
| -  | -  | C4i | C4i  | C6i  | C14   |
| -  | -  | C4i | C5ai | C4i  | C13   |
| -  | -  | C4i | C5ai | C5ai | C14   |
| -  | -  | C4i | C5ai | C5i  | C14   |
| -  | -  | C4i | C5ai | C6ai | C15   |
| -  | -  | C4i | C5ai | C6i  | C15   |
| -  | -  | C4i | C5i  | C4i  | C13   |
| -  | -  | C4i | C5i  | C5ai | C14   |
| -  | -  | C4i | C5i  | C5i  | C14   |
| -  | -  | C4i | C5i  | C6ai | C15   |
| -  | -  | C4i | C5i  | C6i  | C15   |
| -  | -  | C4i | C6ai | C4i  | C14   |
| -  | -  | C4i | C6ai | C5ai | C15   |
| -  | -  | C4i | C6ai | C5i  | C15   |
| -  | -  | C4i | C6ai | C6ai | C16   |
| -  | -  | C4i | C6ai | C6i  | C16   |
| -  | -  | C4i | C6i  | C4i  | C14   |
| -  | -  | C4i | C6i  | C5ai | C15   |
| -  | -  | C4i | C6i  | C5i  | C15   |
| -  | -  | C4i | C6i  | C6ai | C16   |
| -  | -  | C4i | C6i  | C6i  | C16   |
| C2 | -  | C4i | C4i  | C4i  | C12C2 |
| C2 | -  | C4i | C4i  | C5ai | C13C2 |
| C2 | -  | C4i | C4i  | C5i  | C13C2 |
| C2 | -  | C4i | C4i  | C6ai | C14C2 |
| C2 | -  | C4i | C4i  | C6i  | C14C2 |
| C2 | -  | C4i | C5ai | C4i  | C13C2 |
| C2 | -  | C4i | C5ai | C5ai | C14C2 |
| C2 | -  | C4i | C5ai | C5i  | C14C2 |
| C2 | -  | C4i | C5ai | C6ai | C15C2 |
| C2 | -  | C4i | C5ai | C6i  | C15C2 |
| C2 | -  | C4i | C5i  | C4i  | C13C2 |
| C2 | -  | C4i | C5i  | C5ai | C14C2 |
| C2 | -  | C4i | C5i  | C5i  | C14C2 |
| C2 | -  | C4i | C5i  | C6ai | C15C2 |
| C2 | -  | C4i | C5i  | C6i  | C15C2 |
| C2 | -  | C4i | C6ai | C4i  | C14C2 |

|    |    |      |      |      |         |
|----|----|------|------|------|---------|
| C2 | -  | C4i  | C6ai | C5ai | C15C2   |
| C2 | -  | C4i  | C6ai | C5i  | C15C2   |
| C2 | -  | C4i  | C6ai | C6ai | C16C2   |
| C2 | -  | C4i  | C6ai | C6i  | C16C2   |
| C2 | -  | C4i  | C6i  | C4i  | C14C2   |
| C2 | -  | C4i  | C6i  | C5ai | C15C2   |
| C2 | -  | C4i  | C6i  | C5i  | C15C2   |
| C2 | -  | C4i  | C6i  | C6ai | C16C2   |
| C2 | -  | C4i  | C6i  | C6i  | C16C2   |
| C2 | C2 | C4i  | C4i  | C4i  | C2C12C2 |
| C2 | C2 | C4i  | C4i  | C5ai | C2C13C2 |
| C2 | C2 | C4i  | C4i  | C5i  | C2C13C2 |
| C2 | C2 | C4i  | C4i  | C6ai | C2C14C2 |
| C2 | C2 | C4i  | C4i  | C6i  | C2C14C2 |
| C2 | C2 | C4i  | C5ai | C4i  | C2C13C2 |
| C2 | C2 | C4i  | C5ai | C5ai | C2C14C2 |
| C2 | C2 | C4i  | C5ai | C5i  | C2C14C2 |
| C2 | C2 | C4i  | C5ai | C6ai | C2C15C2 |
| C2 | C2 | C4i  | C5ai | C6i  | C2C15C2 |
| C2 | C2 | C4i  | C5i  | C4i  | C2C13C2 |
| C2 | C2 | C4i  | C5i  | C5ai | C2C14C2 |
| C2 | C2 | C4i  | C5i  | C5i  | C2C14C2 |
| C2 | C2 | C4i  | C5i  | C6ai | C2C15C2 |
| C2 | C2 | C4i  | C5i  | C6i  | C2C15C2 |
| C2 | C2 | C4i  | C6ai | C4i  | C2C14C2 |
| C2 | C2 | C4i  | C6ai | C5ai | C2C15C2 |
| C2 | C2 | C4i  | C6ai | C5i  | C2C15C2 |
| C2 | C2 | C4i  | C6ai | C6ai | C2C16C2 |
| C2 | C2 | C4i  | C6ai | C6i  | C2C16C2 |
| C2 | C2 | C4i  | C6i  | C4i  | C2C14C2 |
| C2 | C2 | C4i  | C6i  | C5ai | C2C15C2 |
| C2 | C2 | C4i  | C6i  | C5i  | C2C15C2 |
| C2 | C2 | C4i  | C6i  | C6ai | C2C16C2 |
| C2 | C2 | C4i  | C6i  | C6i  | C2C16C2 |
| -  | C2 | C5ai | C4i  | C4i  | C2C13   |
| -  | C2 | C5ai | C4i  | C5ai | C2C14   |
| -  | C2 | C5ai | C4i  | C5i  | C2C14   |
| -  | C2 | C5ai | C4i  | C6ai | C2C15   |
| -  | C2 | C5ai | C4i  | C6i  | C2C15   |
| -  | C2 | C5ai | C5ai | C4i  | C2C14   |
| -  | C2 | C5ai | C5ai | C5ai | C2C15   |
| -  | C2 | C5ai | C5ai | C5i  | C2C15   |
| -  | C2 | C5ai | C5ai | C6ai | C2C16   |
| -  | C2 | C5ai | C5ai | C6i  | C2C16   |
| -  | C2 | C5ai | C5i  | C4i  | C2C14   |
| -  | C2 | C5ai | C5i  | C5ai | C2C15   |
| -  | C2 | C5ai | C5i  | C5i  | C2C15   |
| -  | C2 | C5ai | C5i  | C6ai | C2C16   |
| -  | C2 | C5ai | C5i  | C6i  | C2C16   |
| -  | C2 | C5ai | C6ai | C4i  | C2C15   |
| -  | C2 | C5ai | C6ai | C5ai | C2C16   |
| -  | C2 | C5ai | C6ai | C5i  | C2C16   |
| -  | C2 | C5ai | C6ai | C6ai | C2C17   |
| -  | C2 | C5ai | C6ai | C6i  | C2C17   |
| -  | C2 | C5ai | C6i  | C4i  | C2C15   |
| -  | C2 | C5ai | C6i  | C5ai | C2C16   |
| -  | C2 | C5ai | C6i  | C5i  | C2C16   |
| -  | C2 | C5ai | C6i  | C6ai | C2C17   |
| -  | C2 | C5ai | C6i  | C6i  | C2C17   |
| -  | -  | C5ai | C4i  | C4i  | C13     |
| -  | -  | C5ai | C4i  | C5ai | C14     |
| -  | -  | C5ai | C4i  | C5i  | C14     |
| -  | -  | C5ai | C4i  | C6ai | C15     |
| -  | -  | C5ai | C4i  | C6i  | C15     |
| -  | -  | C5ai | C5ai | C4i  | C14     |
| -  | -  | C5ai | C5ai | C5ai | C15     |
| -  | -  | C5ai | C5ai | C5i  | C15     |
| -  | -  | C5ai | C5ai | C6ai | C16     |
| -  | -  | C5ai | C5ai | C6i  | C16     |
| -  | -  | C5ai | C5i  | C4i  | C14     |
| -  | -  | C5ai | C5i  | C5ai | C15     |
| -  | -  | C5ai | C5i  | C5i  | C15     |
| -  | -  | C5ai | C5i  | C6ai | C16     |
| -  | -  | C5ai | C5i  | C6i  | C16     |
| -  | -  | C5ai | C6ai | C4i  | C15     |
| -  | -  | C5ai | C6ai | C5ai | C16     |

|    |    |      |      |      |         |
|----|----|------|------|------|---------|
| -  | -  | C5ai | C6ai | C5i  | C16     |
| -  | -  | C5ai | C6ai | C6ai | C17     |
| -  | -  | C5ai | C6ai | C6i  | C17     |
| -  | -  | C5ai | C6i  | C4i  | C15     |
| -  | -  | C5ai | C6i  | C5ai | C16     |
| -  | -  | C5ai | C6i  | C5i  | C16     |
| -  | -  | C5ai | C6i  | C6ai | C17     |
| -  | -  | C5ai | C6i  | C6i  | C17     |
| C2 | -  | C5ai | C4i  | C4i  | C13C2   |
| C2 | -  | C5ai | C4i  | C5ai | C14C2   |
| C2 | -  | C5ai | C4i  | C5i  | C14C2   |
| C2 | -  | C5ai | C4i  | C6ai | C15C2   |
| C2 | -  | C5ai | C4i  | C6i  | C15C2   |
| C2 | -  | C5ai | C5ai | C4i  | C14C2   |
| C2 | -  | C5ai | C5ai | C5ai | C15C2   |
| C2 | -  | C5ai | C5ai | C5i  | C15C2   |
| C2 | -  | C5ai | C5ai | C6ai | C16C2   |
| C2 | -  | C5ai | C5ai | C6i  | C16C2   |
| C2 | -  | C5ai | C5i  | C4i  | C14C2   |
| C2 | -  | C5ai | C5i  | C5ai | C15C2   |
| C2 | -  | C5ai | C5i  | C5i  | C15C2   |
| C2 | -  | C5ai | C5i  | C6ai | C16C2   |
| C2 | -  | C5ai | C5i  | C6i  | C16C2   |
| C2 | -  | C5ai | C6ai | C4i  | C15C2   |
| C2 | -  | C5ai | C6ai | C5ai | C16C2   |
| C2 | -  | C5ai | C6ai | C5i  | C16C2   |
| C2 | -  | C5ai | C6ai | C6ai | C17C2   |
| C2 | -  | C5ai | C6ai | C6i  | C17C2   |
| C2 | -  | C5ai | C6i  | C4i  | C15C2   |
| C2 | -  | C5ai | C6i  | C5ai | C16C2   |
| C2 | -  | C5ai | C6i  | C5i  | C16C2   |
| C2 | -  | C5ai | C6i  | C6ai | C17C2   |
| C2 | -  | C5ai | C6i  | C6i  | C17C2   |
| C2 | C2 | C5ai | C4i  | C4i  | C2C13C2 |
| C2 | C2 | C5ai | C4i  | C5ai | C2C14C2 |
| C2 | C2 | C5ai | C4i  | C5i  | C2C14C2 |
| C2 | C2 | C5ai | C4i  | C6ai | C2C15C2 |
| C2 | C2 | C5ai | C4i  | C6i  | C2C15C2 |
| C2 | C2 | C5ai | C5ai | C4i  | C2C14C2 |
| C2 | C2 | C5ai | C5ai | C5ai | C2C15C2 |
| C2 | C2 | C5ai | C5ai | C5i  | C2C15C2 |
| C2 | C2 | C5ai | C5ai | C6ai | C2C16C2 |
| C2 | C2 | C5ai | C5ai | C6i  | C2C16C2 |
| C2 | C2 | C5ai | C5i  | C4i  | C2C14C2 |
| C2 | C2 | C5ai | C5i  | C5ai | C2C15C2 |
| C2 | C2 | C5ai | C5i  | C5i  | C2C15C2 |
| C2 | C2 | C5ai | C5i  | C6ai | C2C16C2 |
| C2 | C2 | C5ai | C5i  | C6i  | C2C16C2 |
| C2 | C2 | C5ai | C6ai | C4i  | C2C15C2 |
| C2 | C2 | C5ai | C6ai | C5ai | C2C16C2 |
| C2 | C2 | C5ai | C6ai | C6ai | C2C17C2 |
| C2 | C2 | C5ai | C6ai | C6i  | C2C17C2 |
| C2 | C2 | C5ai | C6i  | C6i  | C2C17C2 |
| -  | C2 | C5i  | C4i  | C4i  | C2C13   |
| -  | C2 | C5i  | C4i  | C5ai | C2C14   |
| -  | C2 | C5i  | C4i  | C5i  | C2C14   |
| -  | C2 | C5i  | C4i  | C6ai | C2C15   |
| -  | C2 | C5i  | C4i  | C6i  | C2C15   |
| -  | C2 | C5i  | C5ai | C4i  | C2C14   |
| -  | C2 | C5i  | C5ai | C5ai | C2C15   |
| -  | C2 | C5i  | C5ai | C5i  | C2C15   |
| -  | C2 | C5i  | C5ai | C6ai | C2C16   |
| -  | C2 | C5i  | C5ai | C6i  | C2C16   |
| -  | C2 | C5i  | C5i  | C4i  | C2C14   |
| -  | C2 | C5i  | C5i  | C5ai | C2C15   |
| -  | C2 | C5i  | C5i  | C5i  | C2C15   |
| -  | C2 | C5i  | C5i  | C6ai | C2C16   |
| -  | C2 | C5i  | C5i  | C6i  | C2C16   |
| -  | C2 | C5i  | C6ai | C4i  | C2C15   |
| -  | C2 | C5i  | C6ai | C5ai | C2C16   |
| -  | C2 | C5i  | C6ai | C5i  | C2C16   |

|    |    |     |      |      |         |
|----|----|-----|------|------|---------|
| -  | C2 | C5i | C6ai | C6ai | C2C17   |
| -  | C2 | C5i | C6ai | C6i  | C2C17   |
| -  | C2 | C5i | C6i  | C4i  | C2C15   |
| -  | C2 | C5i | C6i  | C5ai | C2C16   |
| -  | C2 | C5i | C6i  | C5i  | C2C16   |
| -  | C2 | C5i | C6i  | C6ai | C2C17   |
| -  | C2 | C5i | C6i  | C6i  | C2C17   |
| -  | -  | C5i | C4i  | C4i  | C13     |
| -  | -  | C5i | C4i  | C5ai | C14     |
| -  | -  | C5i | C4i  | C5i  | C14     |
| -  | -  | C5i | C4i  | C6ai | C15     |
| -  | -  | C5i | C4i  | C6i  | C15     |
| -  | -  | C5i | C5ai | C4i  | C14     |
| -  | -  | C5i | C5ai | C5ai | C15     |
| -  | -  | C5i | C5ai | C5i  | C15     |
| -  | -  | C5i | C5ai | C6ai | C16     |
| -  | -  | C5i | C5ai | C6i  | C16     |
| -  | -  | C5i | C5i  | C4i  | C14     |
| -  | -  | C5i | C5i  | C5ai | C15     |
| -  | -  | C5i | C5i  | C5i  | C15     |
| -  | -  | C5i | C5i  | C6ai | C16     |
| -  | -  | C5i | C5i  | C6i  | C16     |
| -  | -  | C5i | C6ai | C4i  | C15     |
| -  | -  | C5i | C6ai | C5ai | C16     |
| -  | -  | C5i | C6ai | C5i  | C16     |
| -  | -  | C5i | C6ai | C6ai | C17     |
| -  | -  | C5i | C6ai | C6i  | C17     |
| -  | -  | C5i | C6i  | C4i  | C15     |
| -  | -  | C5i | C6i  | C5ai | C16     |
| -  | -  | C5i | C6i  | C5i  | C16     |
| -  | -  | C5i | C6i  | C6ai | C17     |
| -  | -  | C5i | C6i  | C6i  | C17     |
| C2 | -  | C5i | C4i  | C4i  | C13C2   |
| C2 | -  | C5i | C4i  | C5ai | C14C2   |
| C2 | -  | C5i | C4i  | C5i  | C14C2   |
| C2 | -  | C5i | C4i  | C6ai | C15C2   |
| C2 | -  | C5i | C4i  | C6i  | C15C2   |
| C2 | -  | C5i | C5ai | C4i  | C14C2   |
| C2 | -  | C5i | C5ai | C5ai | C15C2   |
| C2 | -  | C5i | C5ai | C5i  | C15C2   |
| C2 | -  | C5i | C5ai | C6ai | C16C2   |
| C2 | -  | C5i | C5ai | C6i  | C16C2   |
| C2 | -  | C5i | C5i  | C4i  | C14C2   |
| C2 | -  | C5i | C5i  | C5ai | C15C2   |
| C2 | -  | C5i | C5i  | C5i  | C15C2   |
| C2 | -  | C5i | C5i  | C6ai | C16C2   |
| C2 | -  | C5i | C5i  | C6i  | C16C2   |
| C2 | -  | C5i | C6ai | C4i  | C15C2   |
| C2 | -  | C5i | C6ai | C5ai | C16C2   |
| C2 | -  | C5i | C6ai | C5i  | C16C2   |
| C2 | -  | C5i | C6ai | C6ai | C17C2   |
| C2 | -  | C5i | C6ai | C6i  | C17C2   |
| C2 | -  | C5i | C6i  | C4i  | C15C2   |
| C2 | -  | C5i | C6i  | C5ai | C16C2   |
| C2 | -  | C5i | C6i  | C5i  | C16C2   |
| C2 | -  | C5i | C6i  | C6ai | C17C2   |
| C2 | -  | C5i | C6i  | C6i  | C17C2   |
| C2 | C2 | C5i | C4i  | C4i  | C2C13C2 |
| C2 | C2 | C5i | C4i  | C5ai | C2C14C2 |
| C2 | C2 | C5i | C4i  | C5i  | C2C14C2 |
| C2 | C2 | C5i | C4i  | C6ai | C2C15C2 |
| C2 | C2 | C5i | C4i  | C6i  | C2C15C2 |
| C2 | C2 | C5i | C5ai | C4i  | C2C14C2 |
| C2 | C2 | C5i | C5ai | C5ai | C2C15C2 |
| C2 | C2 | C5i | C5ai | C5i  | C2C15C2 |
| C2 | C2 | C5i | C5ai | C6ai | C2C16C2 |
| C2 | C2 | C5i | C5ai | C6i  | C2C16C2 |
| C2 | C2 | C5i | C5i  | C4i  | C2C14C2 |
| C2 | C2 | C5i | C5i  | C5ai | C2C15C2 |
| C2 | C2 | C5i | C5i  | C5i  | C2C15C2 |
| C2 | C2 | C5i | C5i  | C6ai | C2C16C2 |
| C2 | C2 | C5i | C6ai | C4i  | C2C15C2 |
| C2 | C2 | C5i | C6ai | C5ai | C2C16C2 |
| C2 | C2 | C5i | C6ai | C5i  | C2C16C2 |
| C2 | C2 | C5i | C6ai | C6ai | C2C17C2 |

|    |    |      |      |      |         |
|----|----|------|------|------|---------|
| C2 |    | C5i  | C6ai | C6i  | C2C17C2 |
| C2 | C2 | C5i  | C6i  | C4i  | C2C15C2 |
| C2 | C2 | C5i  | C6i  | C5ai | C2C16C2 |
| C2 | C2 | C5i  | C6i  | C5i  | C2C16C2 |
| C2 | C2 | C5i  | C6i  | C6ai | C2C17C2 |
| C2 | C2 | C5i  | C6i  | C6i  | C2C17C2 |
| -  | C2 | C6ai | C4i  | C4i  | C2C14   |
| -  | C2 | C6ai | C4i  | C5ai | C2C15   |
| -  | C2 | C6ai | C4i  | C5i  | C2C15   |
| -  | C2 | C6ai | C4i  | C6ai | C2C16   |
| -  | C2 | C6ai | C4i  | C6i  | C2C16   |
| -  | C2 | C6ai | C5ai | C4i  | C2C15   |
| -  | C2 | C6ai | C5ai | C5ai | C2C16   |
| -  | C2 | C6ai | C5ai | C5i  | C2C16   |
| -  | C2 | C6ai | C5ai | C6ai | C2C17   |
| -  | C2 | C6ai | C5ai | C6i  | C2C17   |
| -  | C2 | C6ai | C5i  | C4i  | C2C15   |
| -  | C2 | C6ai | C5i  | C5ai | C2C16   |
| -  | C2 | C6ai | C5i  | C5i  | C2C16   |
| -  | C2 | C6ai | C5i  | C6ai | C2C17   |
| -  | C2 | C6ai | C5i  | C6i  | C2C17   |
| -  | C2 | C6ai | C6ai | C4i  | C2C16   |
| -  | C2 | C6ai | C6ai | C5ai | C2C17   |
| -  | C2 | C6ai | C6ai | C5i  | C2C17   |
| -  | C2 | C6ai | C6ai | C6ai | C2C18   |
| -  | C2 | C6ai | C6ai | C6i  | C2C18   |
| -  | C2 | C6ai | C6i  | C4i  | C2C16   |
| -  | C2 | C6ai | C6i  | C5ai | C2C17   |
| -  | C2 | C6ai | C6i  | C5i  | C2C17   |
| -  | C2 | C6ai | C6i  | C6ai | C2C18   |
| -  | C2 | C6ai | C6i  | C6i  | C2C18   |
| -  | -  | C6ai | C4i  | C4i  | C14     |
| -  | -  | C6ai | C4i  | C5ai | C15     |
| -  | -  | C6ai | C4i  | C5i  | C15     |
| -  | -  | C6ai | C4i  | C6ai | C16     |
| -  | -  | C6ai | C4i  | C6i  | C16     |
| -  | -  | C6ai | C5ai | C4i  | C15     |
| -  | -  | C6ai | C5ai | C5ai | C16     |
| -  | -  | C6ai | C5ai | C5i  | C16     |
| -  | -  | C6ai | C5ai | C6ai | C17     |
| -  | -  | C6ai | C5ai | C6i  | C17     |
| -  | -  | C6ai | C5i  | C4i  | C15     |
| -  | -  | C6ai | C5i  | C5ai | C16     |
| -  | -  | C6ai | C5i  | C5i  | C16     |
| -  | -  | C6ai | C5i  | C6ai | C17     |
| -  | -  | C6ai | C5i  | C6i  | C17     |
| -  | -  | C6ai | C6ai | C4i  | C16     |
| -  | -  | C6ai | C6ai | C5ai | C17     |
| -  | -  | C6ai | C6ai | C5i  | C17     |
| -  | -  | C6ai | C6ai | C6ai | C18     |
| -  | -  | C6ai | C6ai | C6i  | C18     |
| -  | -  | C6ai | C6i  | C4i  | C16     |
| -  | -  | C6ai | C6i  | C5ai | C17     |
| -  | -  | C6ai | C6i  | C5i  | C17     |
| -  | -  | C6ai | C6i  | C6ai | C18     |
| -  | -  | C6ai | C6i  | C6i  | C18     |
| C2 | -  | C6ai | C4i  | C4i  | C14C2   |
| C2 | -  | C6ai | C4i  | C5ai | C15C2   |
| C2 | -  | C6ai | C4i  | C5i  | C15C2   |
| C2 | -  | C6ai | C4i  | C6ai | C16C2   |
| C2 | -  | C6ai | C4i  | C6i  | C16C2   |
| C2 | -  | C6ai | C5ai | C4i  | C15C2   |
| C2 | -  | C6ai | C5ai | C5ai | C16C2   |
| C2 | -  | C6ai | C5ai | C5i  | C16C2   |
| C2 | -  | C6ai | C5ai | C6ai | C17C2   |
| C2 | -  | C6ai | C5ai | C6i  | C17C2   |
| C2 | -  | C6ai | C5i  | C4i  | C15C2   |
| C2 | -  | C6ai | C5i  | C5ai | C16C2   |
| C2 | -  | C6ai | C5i  | C6ai | C17C2   |
| C2 | -  | C6ai | C5i  | C6i  | C17C2   |
| C2 | -  | C6ai | C6ai | C4i  | C16C2   |
| C2 | -  | C6ai | C6ai | C5ai | C17C2   |
| C2 | -  | C6ai | C6ai | C5i  | C17C2   |
| C2 | -  | C6ai | C6ai | C6ai | C18C2   |
| C2 | -  | C6ai | C6ai | C6i  | C18C2   |

|    |    |      |      |      |         |
|----|----|------|------|------|---------|
| C2 | -  | C6ai | C6i  | C4i  | C16C2   |
| C2 | -  | C6ai | C6i  | C5ai | C17C2   |
| C2 | -  | C6ai | C6i  | C5i  | C17C2   |
| C2 | -  | C6ai | C6i  | C6ai | C18C2   |
| C2 | -  | C6ai | C6i  | C6i  | C18C2   |
| C2 | C2 | C6ai | C4i  | C4i  | C2C14C2 |
| C2 | C2 | C6ai | C4i  | C5ai | C2C16C2 |
| C2 | C2 | C6ai | C4i  | C5i  | C2C16C2 |
| C2 | C2 | C6ai | C4i  | C6ai | C2C16C2 |
| C2 | C2 | C6ai | C4i  | C6i  | C2C16C2 |
| C2 | C2 | C6ai | C5ai | C4i  | C2C16C2 |
| C2 | C2 | C6ai | C5ai | C5ai | C2C16C2 |
| C2 | C2 | C6ai | C5ai | C5i  | C2C16C2 |
| C2 | C2 | C6ai | C5ai | C6ai | C2C17C2 |
| C2 | C2 | C6ai | C5ai | C6i  | C2C17C2 |
| C2 | C2 | C6ai | C5i  | C4i  | C2C16C2 |
| C2 | C2 | C6ai | C5i  | C5ai | C2C16C2 |
| C2 | C2 | C6ai | C5i  | C5i  | C2C16C2 |
| C2 | C2 | C6ai | C5i  | C6ai | C2C17C2 |
| C2 | C2 | C6ai | C5i  | C6i  | C2C17C2 |
| C2 | C2 | C6ai | C6ai | C4i  | C2C16C2 |
| C2 | C2 | C6ai | C6ai | C5ai | C2C17C2 |
| C2 | C2 | C6ai | C6ai | C5i  | C2C17C2 |
| C2 | C2 | C6ai | C6ai | C6ai | C2C18C2 |
| C2 | C2 | C6ai | C6ai | C6i  | C2C18C2 |
| C2 | C2 | C6ai | C6i  | C4i  | C2C16C2 |
| C2 | C2 | C6ai | C6i  | C5ai | C2C17C2 |
| C2 | C2 | C6ai | C6i  | C5i  | C2C17C2 |
| C2 | C2 | C6ai | C6i  | C6ai | C2C18C2 |
| C2 | C2 | C6ai | C6i  | C6i  | C2C18C2 |
| -  | C2 | C6i  | C4i  | C4i  | C2C14   |
| -  | C2 | C6i  | C4i  | C5ai | C2C15   |
| -  | C2 | C6i  | C4i  | C5i  | C2C15   |
| -  | C2 | C6i  | C4i  | C6ai | C2C16   |
| -  | C2 | C6i  | C4i  | C6i  | C2C16   |
| -  | C2 | C6i  | C5ai | C4i  | C2C15   |
| -  | C2 | C6i  | C5ai | C5ai | C2C16   |
| -  | C2 | C6i  | C5ai | C5i  | C2C16   |
| -  | C2 | C6i  | C5ai | C6ai | C2C17   |
| -  | C2 | C6i  | C5ai | C6i  | C2C17   |
| -  | C2 | C6i  | C5i  | C4i  | C2C15   |
| -  | C2 | C6i  | C5i  | C5ai | C2C16   |
| -  | C2 | C6i  | C5i  | C5i  | C2C16   |
| -  | C2 | C6i  | C5i  | C6ai | C2C17   |
| -  | C2 | C6i  | C5i  | C6i  | C2C17   |
| -  | C2 | C6i  | C6ai | C4i  | C2C16   |
| -  | C2 | C6i  | C6ai | C5ai | C2C17   |
| -  | C2 | C6i  | C6ai | C5i  | C2C17   |
| -  | C2 | C6i  | C6ai | C6ai | C2C18   |
| -  | C2 | C6i  | C6ai | C6i  | C2C18   |
| -  | C2 | C6i  | C6i  | C4i  | C2C16   |
| -  | C2 | C6i  | C6i  | C5ai | C2C17   |
| -  | C2 | C6i  | C6i  | C5i  | C2C17   |
| -  | C2 | C6i  | C6i  | C6ai | C2C18   |
| -  | C2 | C6i  | C6i  | C6i  | C2C18   |
| -  | -  | C6i  | C4i  | C4i  | C14     |
| -  | -  | C6i  | C4i  | C5ai | C15     |
| -  | -  | C6i  | C4i  | C5i  | C15     |
| -  | -  | C6i  | C4i  | C6ai | C16     |
| -  | -  | C6i  | C4i  | C6i  | C16     |
| -  | -  | C6i  | C5ai | C4i  | C15     |
| -  | -  | C6i  | C5ai | C5ai | C16     |
| -  | -  | C6i  | C5ai | C5i  | C16     |
| -  | -  | C6i  | C5ai | C6ai | C17     |
| -  | -  | C6i  | C5ai | C6i  | C17     |
| -  | -  | C6i  | C5i  | C4i  | C15     |
| -  | -  | C6i  | C5i  | C5ai | C16     |
| -  | -  | C6i  | C5i  | C5i  | C16     |
| -  | -  | C6i  | C5i  | C6ai | C17     |
| -  | -  | C6i  | C5i  | C6i  | C17     |
| -  | -  | C6i  | C6ai | C4i  | C16     |
| -  | -  | C6i  | C6ai | C5ai | C17     |
| -  | -  | C6i  | C6ai | C5i  | C17     |
| -  | -  | C6i  | C6ai | C6ai | C18     |
| -  | -  | C6i  | C6ai | C6i  | C18     |
| -  | -  | C6i  | C6i  | C4i  | C16     |

|    |    |     |      |      |         |
|----|----|-----|------|------|---------|
| -  | -  | C6i | C6i  | C5ai | C17     |
| -  | -  | C6i | C6i  | C6i  | C17     |
| -  | -  | C6i | C6i  | C6ai | C18     |
| -  | -  | C6i | C6i  | C6i  | C18     |
| C2 | -  | C6i | C4i  | C4i  | C14C2   |
| C2 | -  | C6i | C4i  | C5ai | C15C2   |
| C2 | -  | C6i | C4i  | C5i  | C15C2   |
| C2 | -  | C6i | C4i  | C6ai | C16C2   |
| C2 | -  | C6i | C4i  | C6i  | C16C2   |
| C2 | -  | C6i | C5ai | C4i  | C15C2   |
| C2 | -  | C6i | C5ai | C5ai | C16C2   |
| C2 | -  | C6i | C5ai | C5i  | C16C2   |
| C2 | -  | C6i | C5ai | C6ai | C17C2   |
| C2 | -  | C6i | C5ai | C6i  | C17C2   |
| C2 | -  | C6i | C5i  | C4i  | C15C2   |
| C2 | -  | C6i | C5i  | C5ai | C16C2   |
| C2 | -  | C6i | C5i  | C5i  | C16C2   |
| C2 | -  | C6i | C5i  | C6ai | C17C2   |
| C2 | -  | C6i | C5i  | C6i  | C17C2   |
| C2 | -  | C6i | C6ai | C4i  | C16C2   |
| C2 | -  | C6i | C6ai | C5ai | C17C2   |
| C2 | -  | C6i | C6ai | C5i  | C17C2   |
| C2 | -  | C6i | C6ai | C6ai | C18C2   |
| C2 | -  | C6i | C6ai | C6i  | C18C2   |
| C2 | -  | C6i | C6i  | C4i  | C16C2   |
| C2 | -  | C6i | C6i  | C5ai | C17C2   |
| C2 | -  | C6i | C6i  | C5i  | C17C2   |
| C2 | -  | C6i | C6i  | C6ai | C18C2   |
| C2 | -  | C6i | C6i  | C6i  | C18C2   |
| C2 | C2 | C6i | C4i  | C4i  | C2C14C2 |
| C2 | C2 | C6i | C4i  | C5ai | C2C15C2 |
| C2 | C2 | C6i | C4i  | C5i  | C2C15C2 |
| C2 | C2 | C6i | C4i  | C6ai | C2C16C2 |
| C2 | C2 | C6i | C4i  | C6i  | C2C16C2 |
| C2 | C2 | C6i | C5ai | C4i  | C2C15C2 |
| C2 | C2 | C6i | C5ai | C5ai | C2C16C2 |
| C2 | C2 | C6i | C5ai | C5i  | C2C16C2 |
| C2 | C2 | C6i | C5ai | C6ai | C2C17C2 |
| C2 | C2 | C6i | C5ai | C6i  | C2C17C2 |
| C2 | C2 | C6i | C5i  | C4i  | C2C15C2 |
| C2 | C2 | C6i | C5i  | C5ai | C2C16C2 |
| C2 | C2 | C6i | C5i  | C5i  | C2C16C2 |
| C2 | C2 | C6i | C5i  | C6ai | C2C17C2 |
| C2 | C2 | C6i | C5i  | C6i  | C2C17C2 |
| C2 | C2 | C6i | C6ai | C4i  | C2C16C2 |
| C2 | C2 | C6i | C6ai | C5ai | C2C17C2 |
| C2 | C2 | C6i | C6ai | C5i  | C2C17C2 |
| C2 | C2 | C6i | C6ai | C6ai | C2C18C2 |
| C2 | C2 | C6i | C6ai | C6i  | C2C18C2 |
| C2 | C2 | C6i | C6i  | C4i  | C2C16C2 |
| C2 | C2 | C6i | C6i  | C5ai | C2C17C2 |
| C2 | C2 | C6i | C6i  | C5i  | C2C17C2 |
| C2 | C2 | C6i | C6i  | C6ai | C2C18C2 |
| C2 | C2 | C6i | C6i  | C6i  | C2C18C2 |

C2 denotes acetyl; C3 denotes butyryl; C4 denotes propionyl; C4i denotes isobutyryl; C5 denotes valeryl (pentanoyl); C5ai denotes 2-methyl-butyryl; C5i denotes isovaleryl; C5:1i denotes isopentenoyl; C5:1 denotes pentenoyl; C6 denotes hexanoyl; C6(2) denotes 2-methylvaleryl; C6ai denotes beta-methylvaleryl; C6i denotes 4-methylvaleryl; - is typically hydrogen.

**CLAIMS**

1. A mutant, non-naturally occurring or transgenic tobacco plant, wherein the plant comprises a plant cell comprising:

- (i) at least one polynucleotide comprising, consisting or consisting essentially of a sequence encoding an isopropylmalate synthase and having at least 80% sequence identity to SEQ ID NO:1 or SEQ ID NO:10 or SEQ ID NO: 12 or SEQ ID NO:14; or
- (ii) a polypeptide encoded by the polynucleotide(s); or
- (iii) a polypeptide having at least 80% sequence identity to SEQ ID NO:2 or SEQ ID NO:11 or SEQ ID NO:13 or SEQ ID NO:15; or
- (iv) a construct, vector or expression vector comprising the polynucleotide sequence(s), optionally in which the construct, vector or expression vector additionally comprises a promoter comprising, consisting or consisting essentially of the sequence set forth in SEQ ID NO:8 or a variant thereof with at least 60% identity thereto or a trichome promoter or a naturally occurring isopropylmalate synthase promoter; and wherein the visual appearance of said tobacco plant is substantially the same as the control tobacco plant; and wherein the chlorophyll content and/or the colouration of the leaves of the mutant, non-naturally occurring or transgenic plant is substantially the same as the control plant at three months after transplant or 36 days after topping.

2. A method for modulating the quantity or type of sucrose esters in at least a part of a tobacco plant, comprising the steps of:

- (i) modulating the expression or activity of isopropylmalate synthase in the plant, in which the isopropylmalate synthase comprises a polynucleotide comprising, consisting or consisting essentially of a sequence encoding isopropylmalate synthase and having at least 80% sequence identity to SEQ ID NO:1 or SEQ ID NO:10 or SEQ ID NO: 12 or SEQ ID NO:14; or is encoded by a polypeptide encoded by the polynucleotide(s); or in which the isopropylmalate synthase is encoded by a polypeptide having at least 80% sequence identity to SEQ ID NO:2 or SEQ ID NO:11 or SEQ ID NO:13 or SEQ ID NO:15;

- (ii) optionally measuring the quantity or determining the type of one or more sucrose esters in at least a part of the mutant, non-naturally occurring or transgenic plant obtained in step (i); and
- (iii) identifying a mutant, non-naturally occurring or transgenic plant in which the quantity or type of the one or more sucrose esters therein has changed in comparison to

a control plant in which the expression or activity of isopropyl malate synthase has not been modulated and, in which the visual appearance of the mutant, non-naturally occurring or transgenic plant is substantially the same as the control plant, and wherein the chlorophyll content and/or the colouration of the leaves of the mutant, non-naturally occurring or transgenic plant is substantially the same as the control plant at three months after transplant or 36 days after topping.

3. A mutant, non-naturally occurring or transgenic tobacco plant according to Claim 1, in which expression of isopropylmalate synthase or the activity of the protein encoded thereby is modulated and at least a part of the tobacco plant has a change in the composition of one or more sucrose esters as compared to a control tobacco plant in which the expression or the activity of isopropylmalate synthase has not been modulated.

4. Plant material including biomass, seed or leaves comprising cells or tissue from a plant according to Claim 1 or Claim 3.

5. A tobacco product comprising a part of a plant according to Claim 1 or Claim 3 or plant material according to Claim 4.

6. A method for producing a composition comprising one or more sucrose esters comprising the steps of:

- (i) providing at least part of a mutant, non-naturally occurring or transgenic tobacco plant according to Claim 1 or Claim 3, plant material according to Claim 4 or a tobacco product according to Claim 5;
- (ii) extracting the sucrose ester(s) therefrom; and
- (iii) optionally, isolating or purifying the extracted sucrose ester(s).

7. A method according to Claim 2 or 6 or the mutant, non-naturally occurring or transgenic plant according to Claims 1 or 3, in which one or more of the sucrose esters has the structure as shown in Figure 5 and in which R3 is acetyl or hydrogen, preferably acetyl; one or more of R1, R2 and R4 comprise an acyl chain with 6 carbons, preferably beta-methylvaleryl; and R5 is an acetyl or hydrogen, preferably, hydrogen.

8. A method for modulating the flavour of tobacco or a tobacco product comprising: (i) adding to tobacco or a tobacco product, a part of a plant, preferably leaves, from a mutant,

non-naturally occurring or transgenic plant according to any of Claims 1 or 3 or plant material according to Claim 4; or (ii) adding to tobacco or a tobacco product a composition comprising sucrose esters obtained or obtainable by a method according to Claims 6 or 7.

9. A method for producing beta-methylvaleric acid comprising the steps of:

- (i) providing at least part of a mutant, non-naturally occurring or transgenic plant according to any of Claims 1 or 3, plant material according to Claim 4 or a tobacco product according to Claim 5;
- (ii) hydrolysing the material or an extract thereof provided in step (i); and
- (iii) optionally isolating or purifying the beta-methylvaleric acid.