

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

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
29 December 2011 (29.12.2011)

PCT

(10) International Publication Number
WO 2011/161431 A2

(51) International Patent Classification:
C07K 14/435 (2006.01)

(21) International Application Number:
PCT/GB2011/051121

(22) International Filing Date:
16 June 2011 (16.06.2011)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
1010471.9 22 June 2010 (22.06.2010) GB
1021720.6 22 December 2010 (22.12.2010) GB

(71) Applicant (for all designated States except US): **GLAXOSMITHKLINE AUSTRALIA PTY LIMITED** [AU/AU]; 1061 Mountain Highway, Boronia, Victoria 3155 (AU).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **WINZER, Thilo Hans** [DE/GB]; Department of Biology, University of York, Heslington, York, Yorkshire YO10 5DD (GB). **WALKER, Tracy Carol** [AU/AU]; GlaxoSmithKline Australia Pty Limited, 14 Henry Street, Latrobe, Tasmania 7307 (AU). **GRAHAM, Ian Alexander** [GB/GB]; Department of Biology, University of York, Heslington, York, Yorkshire YO10 5DD (GB).

(74) Agent: **SYMBIOSIS IP**; Apollo House, Eboracum Way, Heworth Green, York, Yorkshire YO31 7RE (GB).

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

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— without international search report and to be republished upon receipt of that report (Rule 48.2(g))

(54) Title: METHYLTRANSFERASE NUCLEIC ACIDS AND POLYPEPTIDES

(57) Abstract: This disclosure relates to the isolation and sequencing of nucleic acid molecules that encode methyltransferase polypeptides from a *Papaver somniferum* cultivar; and uses in the production of noscapine and identification of poppy cultivars that include the genes that comprise said nucleic acid molecules.



WO 2011/161431 A2

Methyltransferase Nucleic Acids and Polypeptides

Introduction

5 This disclosure relates to the isolation and sequencing of nucleic acid molecules that encode methyltransferase polypeptides from a *Papaver somniferum* cultivar, [poppy plant]; transgenic cells transformed with said nucleic acid molecules, sequence variants of the genes; the use of said genes/proteins in the production of noscapine and the use of the genes as markers of poppy plants that synthesize noscapine.

10

Background

The opium poppy *P. somniferum* is the plant from which opium is extracted. The opium poppy is the only commercially exploited poppy of the family *Papaveraceae* and is the principal source of natural opiates. The opium is extracted from latex harvested from the green seed pods. A further source of opiate alkaloids is the poppy straw which is the dried mature plant. *P. somniferum* is a source of clinically useful opiate alkaloids such as morphine, codeine, thebaine, noscapine [also known as narcotine] and papaverine. The clinical application of these opiate alkaloids and their derivatives is broad having use as analgesics, cough suppressants and anti-spasmodics. Although not used as a pharmacological agent in its own right, thebaine is a particularly useful opiate which can be converted into a range of compounds such as hydrocodone, oxycodone, oxymorphone, nalbuphine, naltrexone, buprenorphine and etorphine. These intermediates also have broad pharmaceutical applications. For example, oxycodone, oxymorphone and etorphine are widely used as an analgesic for moderate to severe pain and are often combined with other analgesics such as ibuprofen. Buprenorphine is used in the treatment of heroin addiction and chronic pain. Naltrexone is used in the treatment of alcohol and opiate addiction.

30 The use of thebaine in the production of these compounds is limited since thebaine is a minor component of the opiates extracted from poppy straw only accounting for approximately 0.5-2% of the opium extracted from dry straw. Mutant varieties of *P. somniferum* have been developed that can obtain thebaine and oripavine content of at least 50% by weight of the alkaloid combination of morphine, codeine, thebaine and oripavine; see WO98/02033. Alternative means to enhance the production of thebaine include the spraying of poppy plants with growth regulatory chemicals which inhibit

35

alkaloid biosynthetic pathways to enhance the production of thebaine and other opiate alkaloids.

5 This disclosure relates to molecular analyses of gene expression in poppy cultivars that produce noscapine. Noscapine does not have significant analgesic properties but is used as a cough suppressant and is being investigated as an anti-cancer agent and in the treatment of stroke patients.

10 We have surprisingly found that certain varieties of poppy cultivars have genes that are unique to those cultivars that produce noscapine. We have cloned three genes that have homology to methyltransferases.

The first methyltransferase, PSMT1, exhibits sequence similarity to S-Adenosyl-L-Methionine:Scoulerine-9-O-Methyltransferase from *Coptis japonica* (Accession: 15 Q39522.1, 61% identical) and from *Thalictrum flavum* (Accession: AAU20770.1, 59% identical). The protein from *Coptis japonica* has been characterised and shown to catalyse the transfer of the S-methyl group of S-adenosyl-L-methionine to the 9-hydroxyl group of scoulerine to form tetrahydrocolumbamine (Sato et al. (1993) Phytochem. 32:659-664) which, in turn, serves as the precursor for the synthesis of most 20 protoberberine alkaloids.

The second methyltransferase, PSMT2, exhibits sequence similarity to S-Adenosyl-L-Methionine:Norcoclaurine-6-O-Methyltransferase from *Coptis japonica* (Accession: Q9LEL6, 42% identical) which has been shown to catalyse the transfer of the S-methyl 25 group of S-adenosyl-L-methionine to the 6-hydroxyl group of (S)-Norcoclaurine to form (S)-Coclaurine (Morishige et al. (2000) J. Biol. Chem. 275(30): 23398–23405).

The third methyltransferase, PSMT3, exhibits sequence similarity to S-Adenosyl-L-Methionine:Norcoclaurine-6-O-Methyltransferase from *Papaver somniferum* (Accession: 30 AAQ01669, 80% identical) and *Papaver bracteatum* (Accession: ACO90232, 80% identical). The protein from *Papaver somniferum* has been characterised and shown to catalyse the transfer of the S-methyl group of S-adenosyl-L-methionine to to catalyse the transfer of the S-methyl group of S-adenosyl-L-methionine to the 6-hydroxyl group of (S)-Norcoclaurine to form (S)-Coclaurine (Ounarooun et al. (2003) 36:808-819).

35

Statements of Invention

According to an aspect of the invention there is provided an isolated nucleic acid molecule that encodes a methyltransferase polypeptide wherein said nucleic acid molecule comprises or consists of a nucleotide sequence selected from the group consisting of:

- i) a nucleotide sequence as represented by the sequence in Figure 1a, 1b 1c, 3a, 3b or 3c;
- ii) a nucleotide sequence wherein said sequence is degenerate as a result of the genetic code to the nucleotide sequence defined in (i);
- iii) a nucleic acid molecule the complementary strand of which hybridizes under stringent hybridization conditions to the sequence in Figure 1a, 1b, 1c, 3a, 3b or 3c wherein said nucleic acid molecule encodes a methyltransferase polypeptide;
- iv) a nucleotide sequence that encodes a polypeptide comprising an amino acid sequence as represented in Figure 4a, 4b or 4c;
- v) a nucleotide sequence that encodes a polypeptide comprising an amino acid sequence wherein said amino acid sequence is modified by addition deletion or substitution of at least one amino acid residue as represented in iv) above and which has retained or enhanced methy transferase activity.

Hybridization of a nucleic acid molecule occurs when two complementary nucleic acid molecules undergo an amount of hydrogen bonding to each other. The stringency of hybridization can vary according to the environmental conditions surrounding the nucleic acids, the nature of the hybridization method, and the composition and length of the nucleic acid molecules used. Calculations regarding hybridization conditions required for attaining particular degrees of stringency are discussed in Sambrook et al., Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2001); and Tijssen, Laboratory Techniques in Biochemistry and Molecular Biology—Hybridization with Nucleic Acid Probes Part I, Chapter 2 (Elsevier, New York, 1993). The T_m is the temperature at which 50% of a given strand of a nucleic acid molecule is hybridized to its complementary strand. The following is an exemplary set of hybridization conditions and is not limiting:

Very High Stringency (allows sequences that share at least 90% identity to hybridize)

Hybridization: 5x SSC at 65°C for 16 hours
Wash twice: 2x SSC at room temperature (RT) for 15 minutes each
Wash twice: 0.5x SSC at 65°C for 20 minutes each

5 High Stringency (allows sequences that share at least 80% identity to hybridize)

Hybridization: 5x-6x SSC at 65°C-70°C for 16-20 hours
Wash twice: 2x SSC at RT for 5-20 minutes each
Wash twice: 1x SSC at 55°C-70°C for 30 minutes each

10 Low Stringency (allows sequences that share at least 50% identity to hybridize)

Hybridization: 6x SSC at RT to 55°C for 16-20 hours
Wash at least twice: 2x-3x SSC at RT to 55°C for 20-30 minutes each.

15 In a preferred embodiment of the invention said nucleic acid molecule comprises or consists of a nucleotide sequence as represented in Figure 1a, 1b, 1c, 3a, 3b or 3c.

According to a further aspect of the invention there is provided an isolated polypeptide selected from the group consisting of:

- i) a polypeptide comprising or consisting of an amino acid sequence as represented in Figure 4a, 4b or 4c; or
- 20 ii) a modified polypeptide comprising or consisting of a modified amino acid sequence wherein said polypeptide is modified by addition deletion or substitution of at least one amino acid residue of the sequence presented in Figure 4a, 4b or 4c and which has retained or enhanced methyltransferase activity.

25

A modified polypeptide as herein disclosed may differ in amino acid sequence by one or more substitutions, additions, deletions, truncations that may be present in any combination. Among preferred variants are those that vary from a reference polypeptide by conservative amino acid substitutions. Such substitutions are those that substitute a given amino acid by another amino acid of like characteristics. The following non-limiting
30 list of amino acids are considered conservative replacements (similar): a) alanine, serine, and threonine; b) glutamic acid and aspartic acid; c) asparagine and glutamine d) arginine and lysine; e) isoleucine, leucine, methionine and valine and f) phenylalanine, tyrosine and tryptophan. Most highly preferred are variants that retain or enhance the
35 same biological function and activity as the reference polypeptide from which it varies.

In one embodiment, the variant polypeptides have at least 43% identity, more preferably at least 62% identity, even more preferably at least 70% identity, still more preferably at least 75%, 80%, 85%, 90%, 95% identity, and most preferably at least 99% identity with the full length amino acid sequence illustrated herein.

According to a further aspect of the invention there is provided a vector comprising a nucleic acid molecule encoding a methyltransferase according to the invention wherein said nucleic acid molecule is operably linked with a nucleic acid sequence comprising a promoter sequence.

In a preferred embodiment of the invention said nucleic acid sequence comprising a promoter confers constitutive expression on said methyltransferase.

In an alternative preferred embodiment of the invention said nucleic acid sequence comprising a promoter confers regulated expression on said methyltransferase.

In a preferred embodiment of the invention said regulated expression is tissue or developmentally regulated expression.

In a further alternative embodiment of the invention said regulated expression is inducible expression.

In an alternative embodiment of the invention a vector including nucleic acid according to the invention need not include a promoter or other regulatory sequence, particularly if the vector is to be used to introduce the nucleic acid into cells for recombination into the gene.

Preferably the nucleic acid in the vector is under the control of, and operably linked to, an appropriate promoter or other regulatory elements for transcription in a host cell such as a microbial, (e.g. bacterial, yeast), or plant cell. The vector may be a bi-functional expression vector which functions in multiple hosts. In the case of methyltransferase genomic DNA this may contain its own promoter or other regulatory elements and in the case of cDNA this may be under the control of an appropriate promoter or other regulatory elements for expression in the host cell.

By "promoter" is meant a nucleotide sequence upstream from the transcriptional initiation site and which contains all the regulatory regions required for transcription. Suitable promoters include constitutive, tissue-specific, inducible, developmental or other
5 promoters for expression in plant cells comprised in plants depending on design. Such promoters include viral, fungal, bacterial, animal and plant-derived promoters capable of functioning in plant cells.

Constitutive promoters include, for example CaMV 35S promoter (Odell et al. (1985)
10 Nature 313: 9810-812); rice actin (McElroy et al. (1990) Plant Cell 2: 163-171); ubiquitin (Christian et al. (1989) Plant Mol. Biol. 18: 675-689); pEMU (Last et al. (1991) Theor Appl. Genet. 81: 581-588); MAS (Velten et al. (1984) EMBO J. 3: 2723-2730); ALS promoter (U.S. Application Serial No. 08/409,297), and the like. Other constitutive promoters include those in U.S. Patent Nos. 5,608,149; 5,608,144; 5,604,121;
15 5,569,597; 5,466,785; 5,399,680, 5,268,463; and 5,608,142, each of which is incorporated by reference.

Chemical-regulated promoters can be used to modulate the expression of a gene in a plant through the application of an exogenous chemical regulator. Depending upon the
20 objective, the promoter may be a chemical-inducible promoter, where application of the chemical induced gene expression, or a chemical-repressible promoter, where application of the chemical represses gene expression. Chemical-inducible promoters are known in the art and include, but are not limited to, the maize In2-2 promoter, which is activated by benzenesulfonamide herbicide safeners, the maize GST promoter, which
25 is activated by hydrophobic electrophilic compounds that are used as pre-emergent herbicides, and the tobacco PR-1a promoter, which is activated by salicylic acid. Other chemical-regulated promoters of interest include steroid-responsive promoters (see, for example, the glucocorticoid-inducible promoter in Schena et al. (1991) Proc. Natl. Acad. Sci. USA 88: 10421-10425 and McNellis et al. (1998) Plant J. 14(2): 247-257) and
30 tetracycline-inducible and tetracycline-repressible promoters (see, for example, Gatz et al. (1991) Mol. Gen. Genet. 227: 229-237, and US Patent Nos. 5,814,618 and 5,789,156, herein incorporated by reference.

Where enhanced expression in particular tissues is desired, tissue-specific promoters
35 can be utilised. Tissue-specific promoters include those described by Yamamoto et al. (1997) Plant J. 12(2): 255-265; Kawamata et al. (1997) Plant Cell Physiol. 38(7): 792-

803; Hansen et al. (1997) Mol. Gen. Genet. 254(3): 337-343; Russell et al. (1997) Transgenic Res. 6(2): 157-168; Rinehart et al. (1996) Plant Physiol. 112(3): 1331-1341; Van Camp et al. (1996) Plant Physiol. 112(2): 525-535; Canevascni et al. (1996) Plant Physiol. 112(2): 513-524; Yamamoto et al. (1994) Plant Cell Physiol. 35(5): 773-778;
5 Lam (1994) Results Probl. Cell Differ. 20: 181-196; Orozco et al. (1993) Plant Mol. Biol. 23(6): 1129-1138; Mutsuoka et al. (1993) Proc. Natl. Acad. Sci. USA 90 (20): 9586-9590; and Guevara-Garcia et al (1993) Plant J. 4(3): 495-50.

"Operably linked" means joined as part of the same nucleic acid molecule, suitably
10 positioned and oriented for transcription to be initiated from the promoter. DNA operably linked to a promoter is "under transcriptional initiation regulation" of the promoter. In a preferred aspect, the promoter is a tissue specific promoter, an inducible promoter or a developmentally regulated promoter.

15 Particular of interest in the present context are nucleic acid constructs which operate as plant vectors. Specific procedures and vectors previously used with wide success in plants are described by Guerineau and Mullineaux (1993) (Plant transformation and expression vectors. In: Plant Molecular Biology Labfax (Croy RRD ed) Oxford, BIOS Scientific Publishers, pp 121-148. Suitable vectors may include plant viral-derived
20 vectors (see e.g. EP194809).

If desired, selectable genetic markers may be included in the construct, such as those that confer selectable phenotypes such as resistance to herbicides (e.g. kanamycin, hygromycin, phosphinotricin, chlorsulfuron, methotrexate, gentamycin, spectinomycin,
25 imidazolinones and glyphosate).

According to a further aspect of the invention there is provided a transgenic cell transformed or transfected with a nucleic acid molecule or vector according to the
invention.

30

In a preferred embodiment of the invention said cell is a plant cell.

In a preferred embodiment of the invention said plant cell is from the family *Papaveraceae*.

35

In a preferred embodiment of the invention said plant cell is a *Papaver somniferum* cell.

According to a further aspect of the invention there is provided a plant comprising a plant cell according to the invention.

In a preferred embodiment of the invention said plant is from the family *Papaveraceae*;
5 preferably *Papaver somniferum*.

In an alternative preferred embodiment of the invention said cell is a microbial cell; preferably a bacterial or fungal cell [e.g. yeast, *Saccharomyces cerevisiae*].

10 According to a further aspect of the invention there is provided a nucleic acid molecule comprising a transcription cassette wherein said cassette includes a nucleotide sequence designed with reference to Figure 1a, 1b, 1c, 3a, 3b or 3c and is adapted for expression by provision of at least one promoter operably linked to said nucleotide
15 cassette.

In a preferred embodiment of the invention said cassette is adapted such that both sense and antisense nucleic acid molecules are transcribed from said cassette wherein said sense and antisense nucleic acid molecules are adapted to anneal over at least part
20 or all of their length to form a siRNA or shRNA.

In a preferred embodiment of the invention said cassette is provided with at least two promoters adapted to transcribe both sense and antisense strands of said nucleic acid molecule.
25

In an alternative preferred embodiment of the invention said cassette comprises a nucleic acid molecule wherein said molecule comprises a first part linked to a second part wherein said first and second parts are complementary over at least part of their sequence and further wherein transcription of said nucleic acid molecule produces an
30 RNA molecule which forms a double stranded region by complementary base pairing of said first and second parts thereby forming an shRNA.

A technique to specifically ablate gene function is through the introduction of double stranded RNA, also referred to as small inhibitory/interfering RNA (siRNA) or short
35 hairpin RNA [shRNA], into a cell which results in the destruction of mRNA complementary to the sequence included in the siRNA/shRNA molecule. The siRNA

molecule comprises two complementary strands of RNA (a sense strand and an antisense strand) annealed to each other to form a double stranded RNA molecule. The siRNA molecule is typically derived from exons of the gene which is to be ablated. The mechanism of RNA interference is being elucidated. Many organisms respond to the presence of double stranded RNA by activating a cascade that leads to the formation of siRNA. The presence of double stranded RNA activates a protein complex comprising RNase III which processes the double stranded RNA into smaller fragments (siRNAs, approximately 21-29 nucleotides in length) which become part of a ribonucleoprotein complex. The siRNA acts as a guide for the RNase complex to cleave mRNA complementary to the antisense strand of the siRNA thereby resulting in destruction of the mRNA.

In a preferred embodiment of the invention said nucleic acid molecule is part of a vector adapted for expression in a plant cell.

According to a further aspect of the invention there is provided a plant cell transfected with a nucleic acid molecule or vector according to the invention wherein said cell has reduced expression of said methyltransferase.

According to an aspect of the invention there is provided a process for the methylation of an opiate alkaloid comprising:

- i) providing a transgenic plant cell according to the invention;
- ii) cultivating said plant cell to produce a transgenic plant; and optionally
- i) harvesting said transgenic plant, or part thereof.

In a preferred method of the invention said harvested plant material is dried straw and said opiate alkaloid is extracted.

According to an alternative aspect of the invention there is provided a process for the methylation of an opiate alkaloid comprising:

- i) providing a transgenic microbial cell according to the invention that expresses a methy transferase according to the invention in culture with at least one opiate alkaloid;
- ii) cultivating the microbial cell under conditions that methylate one or more opiate alkaloids; and optionally
- iii) isolating said methylated alkaloid from the microbial cell or cell culture.

In a preferred method of the invention said microbial cell is a bacterial cell or fungal/yeast cell.

If microbial cells are used as organisms in the process according to the invention they
5 are grown or cultured in the manner with which the skilled worker is familiar, depending on the host organism. As a rule, microorganisms are grown in a liquid medium comprising a carbon source, usually in the form of sugars, a nitrogen source, usually in the form of organic nitrogen sources such as yeast extract or salts such as ammonium sulfate, trace elements such as salts of iron, manganese and magnesium and, if
10 appropriate, vitamins, at temperatures of between 0°C and 100°C, preferably between 10°C and 60°C, while gassing in oxygen.

The pH of the liquid medium can either be kept constant, that is to say regulated during the culturing period, or not. The cultures can be grown batchwise, semi-batchwise or continuously. Nutrients can be provided at the beginning of the fermentation or fed in
15 semi-continuously or continuously. The methylated opiate alkaloids produced can be isolated from the organisms as described above by processes known to the skilled worker, for example by extraction, distillation, crystallization, if appropriate precipitation with salt, and/or chromatography. To this end, the organisms can advantageously be disrupted beforehand. In this process, the pH value is advantageously kept between pH
20 4 and 12, preferably between pH 6 and 9, especially preferably between pH 7 and 8.

The culture medium to be used must suitably meet the requirements of the strains in question. Descriptions of culture media for various microorganisms can be found in the textbook "Manual of Methods for General Bacteriology" of the American Society for Bacteriology (Washington D.C., USA, 1981).

25 As described above, these media which can be employed in accordance with the invention usually comprise one or more carbon sources, nitrogen sources, inorganic salts, vitamins and/or trace elements.

Preferred carbon sources are sugars, such as mono-, di- or polysaccharides. Examples of carbon sources are glucose, fructose, mannose, galactose, ribose, sorbose, ribulose,
30 lactose, maltose, sucrose, raffinose, starch or cellulose. Sugars can also be added to the media via complex compounds such as molasses or other by-products from sugar refining. The addition of mixtures of a variety of carbon sources may also be advantageous. Other possible carbon sources are oils and fats such as, for example,

soya oil, sunflower oil, peanut oil and/or coconut fat, fatty acids such as, for example, palmitic acid, stearic acid and/or linoleic acid, alcohols and/or polyalcohols such as, for example, glycerol, methanol and/or ethanol, and/or organic acids such as, for example, acetic acid and/or lactic acid.

- 5 Nitrogen sources are usually organic or inorganic nitrogen compounds or materials comprising these compounds. Examples of nitrogen sources comprise ammonia in liquid or gaseous form or ammonium salts such as ammonium sulfate, ammonium chloride, ammonium phosphate, ammonium carbonate or ammonium nitrate, nitrates, urea, amino acids or complex nitrogen sources such as cornsteep liquor, soya meal, soya protein, 10 yeast extract, meat extract and others. The nitrogen sources can be used individually or as a mixture.

Inorganic salt compounds which may be present in the media comprise the chloride, phosphorus and sulfate salts of calcium, magnesium, sodium, cobalt, molybdenum, potassium, manganese, zinc, copper and iron.

- 15 Inorganic sulfur-containing compounds such as, for example, sulfates, sulfites, dithionites, tetrathionates, thiosulfates, sulfides, or else organic sulfur compounds such as mercaptans and thiols may be used as sources of sulfur for the production of sulfur-containing fine chemicals, in particular of methionine.

- Phosphoric acid, potassium dihydrogenphosphate or dipotassium hydrogenphosphate or 20 the corresponding sodium-containing salts may be used as sources of phosphorus.

Chelating agents may be added to the medium in order to keep the metal ions in solution. Particularly suitable chelating agents comprise dihydroxyphenols such as catechol or protocatechuate and organic acids such as citric acid.

- 25 The fermentation media used according to the invention for culturing microorganisms usually also comprise other growth factors such as vitamins or growth promoters, which include, for example, biotin, riboflavin, thiamine, folic acid, nicotinic acid, panthothenate and pyridoxine. Growth factors and salts are frequently derived from complex media components such as yeast extract, molasses, cornsteep liquor and the like. It is moreover possible to add suitable precursors to the culture medium. The exact 30 composition of the media compounds heavily depends on the particular experiment and is decided upon individually for each specific case. Information on the optimization of media can be found in the textbook "Applied Microbiol. Physiology, A Practical Approach" (Editors P.M. Rhodes, P.F. Stanbury, IRL Press (1997) pp. 53-73, ISBN 0 19

963577 3). Growth media can also be obtained from commercial suppliers, for example Standard 1 (Merck) or BHI (brain heart infusion, DIFCO) and the like.

All media components are sterilized, either by heat (20 min at 1.5 bar and 121 °C) or by filter sterilization. The components may be sterilized either together or, if required,
5 separately. All media components may be present at the start of the cultivation or added continuously or batchwise, as desired.

The culture temperature is normally between 15 °C and 45 °C, preferably at from 25 °C to 40 °C, and may be kept constant or may be altered during the experiment. The pH of the medium should be in the range from 5 to 8.5, preferably around 7.0. The pH for
10 cultivation can be controlled during cultivation by adding basic compounds such as sodium hydroxide, potassium hydroxide, ammonia and aqueous ammonia or acidic compounds such as phosphoric acid or sulfuric acid. Foaming can be controlled by employing antifoams such as, for example, fatty acid polyglycol esters. To maintain the stability of plasmids it is possible to add to the medium suitable substances having a
15 selective effect, for example antibiotics. Aerobic conditions are maintained by introducing oxygen or oxygen-containing gas mixtures such as, for example, ambient air into the culture. The temperature of the culture is normally 20 °C to 45 °C and preferably 25 °C to 40 °C. The culture is continued until formation of the desired product is at a maximum. This aim is normally achieved within 10 to 160 hours.

20 The fermentation broth can then be processed further. The biomass may, according to requirement, be removed completely or partially from the fermentation broth by separation methods such as, for example, centrifugation, filtration, decanting or a combination of these methods or be left completely in said broth. It is advantageous to process the biomass after its separation.

25 However, the fermentation broth can also be thickened or concentrated without separating the cells, using known methods such as, for example, with the aid of a rotary evaporator, thin-film evaporator, falling-film evaporator, by reverse osmosis or by nanofiltration. Finally, this concentrated fermentation broth can be processed to obtain the opiate alkaloids present therein.

30 According to a further aspect of the invention there is provided the use of a gene encoded by a nucleic acid molecule as represented by the nucleic acid sequence in Figure 3a, 3b or 3c, or a nucleic acid molecule that hybridizes under stringent hybridization conditions to the nucleotide sequence in Figure 3a, 3b or 3c and encodes a

polypeptide with methyltransferase activity as a means to identify the presence or absence of a gene that encodes said methyltransferase in a *Papaveraceae* plant.

According to a further aspect of the invention there is provided a method to determine the presence or absence of a gene according to the invention in a *Papaveraceae* variety comprising:

- i) obtaining a sample from a *Papaveraceae* plant;
- ii) extracting genomic DNA from the plant; and
- iii) analyzing the genomic DNA for the presence of a gene comprising or consisting of a nucleotide sequence as represented in Figure 3a, 3b or 3c.

Methods to analyze genomic DNA are well known in the art. For example, polymerase chain reaction methods using sequence specific oligonucleotide primers to amplify specific regions of the gene according to the invention. The extraction, isolation and restriction analysis using sequence specific restriction endonucleases followed by separation and Southern blotting to analyze genomic structure have been established for over thirty years. The analysis may be directed to intron or exon structure or upstream or downstream regions of the gene; e.g. promoter regions.

According to a further aspect of the invention there is provided the use of a gene encoded by a nucleic acid molecule as represented by the nucleic acid sequence in Figure 3a, 3b or 3c, or a nucleic acid molecule that hybridizes under stringent hybridization conditions to the nucleotide sequence in Figure 3a, 3b or 3c and encodes a polypeptide with methyltransferase activity as a means to identify a locus wherein said locus is associated with altered expression or activity of said methyltransferase.

25

Mutagenesis as a means to induce phenotypic changes in organisms is well known in the art and includes but is not limited to the use of mutagenic agents such as chemical mutagens [e.g. base analogues, deaminating agents, DNA intercalating agents, alkylating agents, transposons, bromine, sodium azide] and physical mutagens [e.g. ionizing radiation, psoralen exposure combined with UV irradiation].

30

According to a further aspect of the invention there is provided a method to produce a *Papaveraceae* plant variety that has altered expression of a methyltransferase according to the invention comprising the steps of:

- 5
- i) mutagenesis of wild-type seed from a plant that does express said methyltransferase;
 - ii) cultivation of the seed in i) to produce first and subsequent generations of plants;
 - iii) obtaining seed from the first generation plant and subsequent generations of plants ;
 - iv) determining if the seed from said first and subsequent generations of plants has altered nucleotide sequence and/or altered expression of said methyltransferase;
- 10
- v) obtaining a sample and analysing the nucleic acid sequence of a nucleic acid molecule selected from the group consisting of:
 - a) a nucleic acid molecule comprising a nucleotide sequence as represented in Figure 3a, 3b or 3c;
 - 15 b) a nucleic acid molecule that hybridises to the nucleic acid molecule in a) under stringent hybridisation conditions and that encodes a polypeptide with methyltransferase activity; and optionally
 - vi) comparing the nucleotide sequence of the nucleic acid molecule in said sample to a nucleotide sequence of a nucleic acid molecule of the original wild-type plant.
- 20

In a preferred method of the invention said nucleic acid molecule is analysed by a method comprising the steps of:

- 25
- i) extracting nucleic acid from said mutated plants;
 - ii) amplification of a part of said nucleic acid molecule by a polymerase chain reaction;
 - iii) forming a preparation comprising the amplified nucleic acid and nucleic acid extracted from wild-type seed to form heteroduplex nucleic acid;
 - 30 iv) incubating said preparation with a single stranded nuclease that cuts at a region of heteroduplex nucleic acid to identify the mismatch in said heteroduplex; and
 - v) determining the site of the mismatch in said nucleic acid heteroduplex.

35 In a preferred method of the invention said *Papaveraceae* plant variety has enhanced methyltransferase expression and/or activity.

According to a further aspect of the invention there is provided a plant obtained by the method according to the invention.

- 5 According to an aspect of the invention there is provided a plant wherein said plant comprises a viral vector that includes all or part of a gene comprising a nucleic acid molecule according to the invention.

10 In a preferred embodiment of the invention said gene is encoded by a nucleic acid molecule comprising a nucleic acid sequence selected from the group consisting of:

- i) a nucleic acid molecule comprising a nucleotide sequence as represented in Figure 1a, 1b or 1c ;
- ii) a nucleic acid molecule comprising a nucleotide sequence that hybridises under stringent hybridisation conditions to a nucleic acid molecule in (i) and which encodes a methytransferase polypeptide;
- 15 iii) a nucleic acid molecule that encodes a variant polypeptide that varies from a polypeptide comprising the amino acid sequence as represented in Figure 4a, 4b or 4c.

20 In a preferred embodiment of the invention said nucleic acid molecule comprises or consists of a nucleotide sequence as represented in Figure 1a.

In a preferred embodiment of the invention said nucleic acid molecule comprises or
25 consists of a nucleotide sequence as represented in Figure 1b.

In a preferred embodiment of the invention said nucleic acid molecule comprises or
consists of a nucleotide sequence as represented in Figure 1c

30 In a preferred embodiment of the invention said nucleic acid molecule consists of a nucleotide sequence as represented in Figure 14.

In an alternative preferred embodiment of the invention said nucleic acid molecule
35 consists of a nucleotide sequence as represented in Figure 15.

According to a further aspect of the invention there is provided a viral vector comprising all or part of a nucleic acid molecule according to the invention.

According to an aspect of the invention there is provided the use of a viral vector
5 according to the invention in viral induced gene silencing in a plant.

In a preferred embodiment of the invention, said plant is from the family *Papaveraceae*.

Virus induced gene silencing [VIGS] is known in the art and exploits a RNA mediated
10 antiviral defence mechanism. Plants that are infected with an unmodified virus induces a mechanism that specifically targets the viral genome. However, viral vectors which are engineered to include nucleic acid molecules derived from host plant genes also induce specific inhibition of viral vector expression and additionally target host mRNA. This allows gene specific gene silencing without genetic modification of the plant genome and
15 is essentially a non-transgenic modification.

Throughout the description and claims of this specification, the words "comprise" and "contain" and variations of the words, for example "comprising" and "comprises", means "including but not limited to", and is not intended to (and does not) exclude other
20 moieties, additives, components, integers or steps.

Throughout the description and claims of this specification, the singular encompasses the plural unless the context otherwise requires. In particular, where the indefinite article is used, the specification is to be understood as contemplating plurality as well as
25 singularity, unless the context requires otherwise.

Features, integers, characteristics, compounds, chemical moieties or groups described in conjunction with a particular aspect, embodiment or example of the invention are to be understood to be applicable to any other aspect, embodiment or example described
30 herein unless incompatible therewith.

An embodiment of the invention will now be described by example only and with reference to the following figures:

35 Figure 1a is the nucleotide sequence of a cDNA that encodes *PSMT1*; Figure 1b is the nucleotide sequence of a cDNA that encodes *PSMT2*; Figure 1c is the nucleotide sequence of a cDNA that encodes *PSMT3*

Figure 2: Expression of the *P. somniferum* SAM-Scoulerine-9-O-methyltransferase gene (*PSMT1*). The 16 EST libraries were generated by pyrosequencing using cDNA libraries prepared from stems (S) and capsules (C) at two developmental stages 'early harvest' (EH, 1-3 days after petals had fallen off) and 'late-harvest' (LH, 4-6 days after petals had fallen off) from each of the four *P. somniferum* cultivars: GSK MORPHINE CVS1, GSK MORPHINE CVS2, GSK NOSAPINE CVS1 and GSK THEBAINE CVS1;

Figure 3a is the genomic nucleotide sequence of *PSMT1*; Figure 3b is the genomic nucleotide sequence of *PSMT2*; Figure 3c is the genomic nucleotide sequence of *PSMT3*;

Figure 4a is the deduced amino acid sequence of *PSMT1*; Figure 4b is the deduced amino acid sequence of *PSMT2*; Figure 4c is the deduced amino acid sequence of *PSMT3*;

Figure 5: PCR-amplification of *PsSOMT1* from genomic DNA of the four *Papaver somniferum* cultivars GSK MORPHINE CVS1, GSK MORPHINE CVS2, GSK NOSCAPINE CVS1, GSK THEBAINE CVS1;

20

Figure 6a is a tabular representation of the segregation of *PSMT1* gene in an F2 mapping population derived from a parental cross of cultivars GSK NOSCAPINE CVS1 and GSK THEBAINE CVS1 along with the co-segregation of the *PSMT1* gene and noscapine accumulation in individual F2 plants; Figure 6b is the equivalent tabular representation of the segregation of the *PSMT2* gene. The *PSMT2* genotyping assay failed (as indicated by the failure to amplify the internal positive control) on six DNA samples from F2 plants that were positive for *PSMT1* and *PSMT3* (indicated by asterisks); Figure 6c is the equivalent tabular representation of the segregation of the *PSMT3* gene;

30

Figure 7 illustrates a typical UPLC chromatogram for standard solution;

Figure 8 illustrates a typical UPLC chromatogram for a noscapine containing poppy variety;

35

Figure 9: Expression of the *P. somniferum* methyltransferase gene (*PSMT2*). The 16 EST libraries were generated by pyrosequencing using cDNA libraries prepared from stems (S) and capsules (C) at two developmental stages 'early harvest' (EH, 1-3 days after petals had fallen off) and 'late-harvest' (LH, 4-6 days after petals had fallen off)
5 from each of the four *P. somniferum* cultivars: GSK MORPHINE CVS1, GSK MORPHINE CVS2, GSK NOSCAPHINE CVS1, GSK THEBAINE CVS1;

Figure 10 shows the results of a PCR-amplification of *PSMT2* from genomic DNA of the four *Papaver somniferum* cultivars GSK MORPHINE CVS1, GSK MORPHINE CVS2,
10 GSK NOSCAPHINE CVS1, GSK THEBAINE CVS1;

Figure 11 Expression of the *P. somniferum* methyltransferase gene (*PSMT3*). The 16 EST libraries were generated by pyrosequencing using cDNA libraries prepared from stems (S) and capsules (C) at two developmental stages 'early harvest' (EH, 1-3 days after petals had fallen off) and 'late-harvest' (LH, 4-6 days after petals had fallen off)
15 from each of the four *P. somniferum* cultivars: GSK MORPHINE CVS1, GSK MORPHINE CVS2, GSK NOSCAPHINE CVS1, GSK THEBAINE CVS1;

Figure 12 shows the results of a PCR-amplification of *PSMT3* from genomic DNA of the four *Papaver somniferum* cultivars GSK MORPHINE CVS1, GSK MORPHINE CVS2,
20 GSK NOSCAPHINE CVS1, GSK THEBAINE CVS1;

Figure 13 is the 622 bases long part of the *PHYTOENE DESATURASE* gene sequence amplified from cDNA of GSK NOSCAPHINE CVS1. The 129-bases long sequence stretch
25 used to silence the *PHYTOENE DESATURASE* gene is underlined;

Figure 14 is the part of the cDNA sequence used to silence *PSMT1*;

Figure 15 is the part of the cDNA sequence used to silence *PSMT2*;
30

Figure 16 shows the normalised peak area of scoulerine in the UPLC chromatograms obtained from latex and mature capsules of plants that displayed the photo-bleaching phenotype after infection with the silencing constructs pTRV2-PDS-PSMT1, pTRV2-PDS-PSMT2 or pTRV2-PDS, respectively. The scoulerine peak area obtained from
35 uninfected plants is shown as well. 21-30 plants were analysed per construct; and

Figure 17 shows the normalised peak area of putative narcotoline in the UPLC chromatograms obtained from latex and mature capsules of plants that displayed the photo-bleaching phenotype after infection with the silencing constructs pTRV2-PDS-PSMT1, pTRV2-PDS-PSMT2 or pTRV2-PDS, respectively. The putative narcotoline peak area obtained from uninfected plants is shown as well. 21-30 plants were analysed per construct.

Materials and Methods

10 Generation of EST libraries

a) RNA isolation and cDNA synthesis

Material was harvested from stems and capsules at two developmental stages from four poppy cultivars. RNA was prepared individually from five plants per cultivar, developmental stage and organ. The harvested material was ground in liquid nitrogen using a mortar and pestle. RNA was isolated from the ground stem or capsule preparations using a CTAB (hexadecyltrimethylammonium bromide) based method as described in Chang et al. (1993) Plant Mol. Biol. Rep. 11: 113–116 with slight modifications (three extractions with chloroform:isoamylalcohol, RNA precipitation with Lithium chloride at -20°C over night). RNA was quantified spectrophotometrically before pooling equal amounts of RNA from five plants per cultivar, stage and organ. The pooled samples underwent a final purification step using an RNeasy Plus MicroKit (Qiagen, Crawley, UK) to remove any remaining genomic DNA from the preparations. RNA was typically eluted in 30-100 µl water. cDNA was prepared using a SMART cDNA Library Construction Kit (Clontech, Saint-Germainen-Laye, France) according to the manufacturer's instructions but using SuperScript II Reverse Transcriptase (Invitrogen, Paisley, UK) for first strand synthesis. The CDSIII PCR primer was modified to: 5' ATT CTA GAT CCR ACA TGT TTT TTT TTT TTT TTT TTT TVN 3' where R = A or G, V = A, C or G; N = A/T or C/G. cDNA was digested with MmeI (New England Biolabs Inc., Hitchin, UK) followed by a final purification using a QIAquick PCR Purification kit (Qiagen, Crawley, UK).

b) cDNA Pyrosequencing

The Roche 454 GS-FLX sequencing platform (Branford, CT, USA) was used to perform pyrosequencing on cDNA samples prepared from the following materials for each of the

four *P. somniferum* cultivars GSK MORPHINE CVS2, GSK MORPHINE CVS1, GSK NOSAPINE CVS1 and GSK THEBAINE CVS1:

1. Stem, 1-3 days after petal fall (early harvest)
2. Stem, 4 -6 days after petal fall (late harvest)
- 5 3. Capsule, 1-3 days after petal fall (early harvest)
4. Capsule, 4 -6 days after petal fall (late harvest)

c) Raw sequence analysis, contiguous sequence assembly and annotation

The raw sequence datasets were derived from parallel tagged sequencing on the 454 sequencing platform (Meyer et al. (2008) Nature Protocols 3: 267-278). Primer and tag sequences were first removed from all individual sequence reads. Contiguous sequence assembly was only performed on sequences longer than 40 nucleotides and containing less than 3% unknown (N) residues. These high quality EST sequences were assembled into unique contiguous sequences with the CAP3 Sequence Assembly Program (Huang and Madan (1999) Genome Res. 9: 868-877), and the resulting contigs were annotated locally using the BLAST2 program (Altschul et al. (1997) Nucleic Acids Res. 25: 3389-3402) against the non-redundant peptide database downloaded from the NCBI.

20 d) Expression profiling of the methyltransferase genes

The number of ESTs associated with the respective methyltransferase consensus sequences were counted in each of the 16 EST libraries. The values obtained were normalised on the basis of total ESTs obtained per library.

25 RT-PCR on GSK NOSAPINE CVS1 mRNA and direct amplification and sequencing of full length cDNA of *PSMT1*.

An aliquot of GSK NOSAPINE CVS1 cDNA prepared for cDNA pyrosequencing (see above) was used to amplify and sequence the full length coding sequence of *PSMT1* using the following PCR conditions:

30 Reaction mixture:

	5×HF buffer (Finnzymes)	5 µl
	dNTPs (20 mM each)	0.25 µl
	Fwd primer (10 µM)	2.5 µl
35	Rev primer (10 µM)	2.5 µl
	gDNA (10 ng/µl)	5 µl
	Phusion Hot Start (Finnzymes)	0.25 µl
	dH ₂ O	9.5 µl

Reaction volume: 25 µl

Phusion Hot Start from Finnzymes was purchased through New England Biolabs, (Bishops Cleeve, UK).

5

PCR programs:

	initial denaturation	98 °C	1 min
10	30 cycles of:	denaturation	98 °C 30 sec
		annealing temperature	61 °C 30 sec
		extension	72 °C 40 sec
	final extension	72 °C	10 min
15	incubation	4 °C	storage

primer combination: PSMT1_CLF: ATAGGATCCAACATGGCTACCAATGGCG
PSMT1_CLR: GCGCTCGAGCATTCATTGTGA

These primers served a dual purpose in that they were used to amplify *PSMT1* cDNA for sequencing and cloning (cloning not described here). The underlined bases of the cloning primers delineate the nucleotides which are sequence-specific to the 5' and 3' end of *PSMT1* cDNA, respectively. The *PSMT1* specific nucleotide sequence used to design the primers were confirmed upon sequencing the *PSMT1* gene from genomic DNA as described below.

25 PCR products were cleaned with Agencourt AMPure PCR purification kit (Beckman Coulter, High Wycombe, UK) and sequenced by Sanger sequencing.

Amplification and sequencing of the methyltransferase genes from genomic DNA

30 a) Genomic DNA preparation

Leaf samples (30-50mg) for DNA extraction were harvested from plants of GSK MORHINE CVS1, GSK MORPHINE CVS2, GSK NOSAPINE CVS1 and GSK THEBAINE CVS1 grown in the glasshouse. DNA was extracted using Qiagen BioSprint 96. Extracted DNA was quantified using Hoescht 33258 and normalized to 10 ng/µl.

35

b) Amplification and sequencing of the methyltransferase genes from genomic DNA of GSK NOSCAPINE CVS1

Primers and primer sequences used for amplification of the respective methyltransferase genes from the extracted genomic DNA are shown in Table 1. Primers were designed based on the respective methyltransferase contigs assembled from ESTs unique to

40

cultivar GSK NOSCAPINE CVS1. The respective contigs contained the complete open reading frame of each methyltransferase as well as 5' and 3' untranslated regions. Amplification from genomic DNA was performed using the primer combinations shown in Table 2 and the PCR conditions shown below. The PCR amplifications were performed on pools of genomic DNA comprising DNA from four individuals. 5 µl of each PCR reaction was resolved on 1.2% Agarose alongside appropriate size standards.

Reaction mixture:

10	5×HF buffer (Finnzymes)	5 µl
	dNTPs (20 mM each)	0.25 µl
	Forward primer (10 µM)	2.5 µl
	Reverse primer (10 µM)	2.5 µl
	gDNA (10 ng/µl)	5 µl
15	Phusion Hot Start (Finnzymes)	0.25 µl
	dH ₂ O	9.5 µl

Reaction volume: 25 µl

20 PCR program:

The annealing temperature and extension time for the respective primer combinations are shown in Table 2 and 3.

	initial denaturation	98 °C	1 min
30 cycles of:	denaturation	98 °C	30 sec
25	annealing temperature	Table 2&3	30 sec
	extension	72 °C	Table 2 & 3
	final extension	72 °C	10 min
30	incubation	4 °C	storage

The PCR products were purified using the Agencourt AMPure purification kit (BECKMAN COULTER UK LTD, Bromley, UK). 30-50 ng of the respective purified PCR product were subjected to Sanger-sequencing using the primers indicated in Table 2 as sequencing primers.

35

The amino acid sequences of the respective methyltransferase, deduced from the open reading frame sequences confirmed by Sanger-sequencing, were compared to protein sequences deposited in the NCBI Non-15 redundant protein database using the Standard Protein BLAST programme (blastp).

40 c) Analysis of genomic DNA from GSK NOSAPINE CVS1, GSK MORPHINE CVS2, GSK MORPHINE CVS1 and GSK THEBAINE CVS1 for the presence of the methyltransferase genes

To check if the respective methyltransferase genes were present in all four cultivars, PCR amplifications were performed on genomic DNA from GSK NOSAPINE CVS1, GSK MORPHINE CVS2, GSK MORHINE CVS1 and GSK THEBAINE CVS1 . using the primer combinations shown in Table 3. Exactly the same PCR conditions as described above to amplify the full length genomic sequences from GSK NOSAPINE CVS1 were used . The genomic DNA was derived from four individual plants per cultivar. 5 µl of each PCR reaction was resolved on 1% agarose alongside appropriate size standards.

Analysis of genomic DNA and alkaloid profiling from poppy straw of an F2 population

a) DNA extraction from F2 plants

40-50mg of leaf tissue was harvested, in duplicate, from all poppy plants within the GSK NOSCAPINE CVS1 X GSK THEBAINE CVS1 F2 mapping population at the 'small rosette' growth stage (~10 leaves present on each plant).

Leaf tissue (40 – 50 mg wet weight) was collected into 1.2 ml sample tubes in 8 x 12 format (Part Number 1760-00, Scientific Specialties Inc, 130 Thurman St, Lodi, CA 95240 USA), closed with strip caps (Part Number 1702-00, Scientific Specialties Inc) and shipped to the AGRF (Australian Genome Research Facility) Adelaide on Techni-Ice dry Ice packs by overnight courier.

On receipt, strip caps were removed and a 3 mm tungsten carbide bead was added to each tube (Part Number 69997, Qiagen GmbH, Hilden, Germany). Samples were placed at -80°C (Freezer model; Sanyo MDF-U73V) for a minimum of two hours prior to freeze-drying for 18 hr (Christ Model Alpha 2-4 LSC).

Following freeze drying, tubes were sealed with fresh strip caps (as above), and samples were powdered by bead-milling (Model "Tissue Lyser", Part Number 85300; Qiagen) at 3,000 RPM for 2 x 60 sec cycles separated by plate inversion. DNA extraction was performed using the "Nucleospin Plant II" system (Macherey-Nagel, GmbH & Co. KG Neumann-Neander-Straße 6-8, 52355 Düren, Germany).

Cell lysis was performed using the supplied Buffer Set PL2/3. The manufacturer's protocol for centrifugal extraction was followed (Centrifuge model 4-K 15; Sigma Laborzentrifugen GmbH, 37520 Osterode am Harz, Germany).

The recovered DNA (12/96 samples, one sample per plate column) was checked for quality and quantity by ultra violet spectroscopy (Model Nanodrop-8000; NanoDrop products, 3411 Silverside Rd, Bancroft Building; Wilmington, DE 19810, USA) at 230, 260 and 280 nM.

5

b) Genotyping of F2 DNA samples for presence or absence of the methyltransferases

DNA samples from a total of 276 F2 plants were genotyped for the presence or absence of *PSMT1*, *PSMT2* and *PSMT3*, respectively, by amplifying a short fragment of each of the genes. 5'-VIC-labeled forward primers were used for the amplification of each of the methyltransferase fragments, enabling the separation and analysis of the resulting fluorescently labeled PCR fragments on the 96-capillary electrophoresis 3730x/ DNA Analyzer system (Applied Biosystems) according to the manufacturer's instructions. In addition to the respective methyltransferase fragments, an internal positive control was amplified in each PCR assay in order to distinguish lack of amplification due to absence of the respective methyltransferase genes in a DNA sample from lack of amplification due to PCR assay failure.

15

The following primers were used (primer sequences are shown in Table1):

PSMT1: PSMT1_F3/ PSMT1_R1; amplified fragment size: 129 bp

20

PSMT2: PSMT2_F4/ PSMT2_R3; amplified fragment size: 284bp

PSMT3: PSMT3_F4/ PSMT3_R2; amplified fragment size: 309 bp

The *PSMT1*- and *PSMT3*-fragments were amplified with the following PCR conditions:

Reaction mixture:

25

	5×GoTaq Buffer (Promega)	2 µl
	dNTPs (2.5 mM mix)	0.5 µl
	MgCl ₂ (25 mM)	0.6 µl
	Forward primer (10 µM)	0.5 µl
30	Reverse primer (10 µM)	0.5 µl
	gDNA (5 ng/µl)	2 µl
	GoTaq (Promega)	0.2 µl
	dH ₂ O	3.7 µl

35

Reaction volume: 10 µl

PCR program:

40	30 cycles of:	initial denaturation	94 °C	1 min
		denaturation	94 °C	30 sec
		annealing temperature	62 °C	30 sec
		extension	72 °C	20-30 sec
		final extension	72 °C	5 min

incubation

4 °C

storage

The *PSMT2*- fragment was amplified with the following PCR conditions:

5 Reaction mixture:

	5×Type-it multiplex PCR mix (Qiagen)	5 µl
	Forward primer (10 µM)	0.5 µl
	Reverse primer (10 µM)	0.5 µl
10	gDNA (5 ng/µl)	2 µl
	dH ₂ O	2 µl

Reaction volume: 10 µl

15 PCR program:

	initial denaturation	95 °C	15 min
30 cycles of:	denaturation	95 °C	15 sec
	annealing temperature	60 °C	30 sec
20	extension	72 °C	30 sec
	final extension	72 °C	5 min
	incubation	4 °C	storage

c) Poppy straw analysis from F2 plants

25 Poppy capsules were harvested by hand from the mapping population once capsules had dried to approximately 10% moisture on the plant. The seed was manually separated from the capsule, and capsule straw material (Poppy Straw) was then shipped to the GSK extraction facility in Port Fairy, Australia.

30 The poppy straw samples were then ground in a Retsch Model MM04 ball mill into a fine powder. Two gram samples of ground poppy straw were then weighed accurately (2 ± 0.003 g) and extracted in 50mL of a 10% acetic acid solution. The extraction suspension was shaken on an orbital shaker at 200rpm for a minimum of 10 minutes then filtered to provide a clear filtrate. The final filtrate was passed through a 0.22 µm filter prior to analysis.

35 The solutions were analysed using a Waters Acquity UPLC system fitted with a Waters Acquity BEH C18 column, 2.1mm x 100mm with 1.7 micron packing. The mobile phase used a gradient profile with eluent A consisting of 0.1% Tetrafluoroacetic acid in deionised water and eluent B consisting of 100% Acetonitrile. The mobile phase gradient conditions used are as listed in Table 2, the gradient curve number as determined using

40 a Waters Empower chromatography software package. The flow rate was 0.6mL per

minute and the column maintained at 45°C. The injection volume was 1 µL injection volume and the alkaloids were detected using a UV detector at 285 nm.

The loss on drying (LOD) of the straw was determined by drying in an oven at 105 degrees centigrade for 3 hours.

- Gradient Flow Program:

10

TIME (minutes)	% Eluent A	% Eluent B	Flow (mL/min)	Curve No
0.00	95.0	5.0	0.60	INITIAL
0.80	90.0	10.0	0.60	6
3.40	75.0	25.0	0.60	3
3.60	95.0	5.0	0.60	6
4.00	95.0	5.0	0.60	11

15 Alkaloid concentrations for morphine, codeine, thebaine, oripavine and noscapine were determined by comparison with standard solutions and the results calculated on a dry weight basis.

Typical retention times are as follows:

Compound	Retention Time (minutes)
Morphine	1.14
Pseudo morphine	1.26
Codeine	1.69
Oripavine	1.80
10-Hydroxythebaine	2.32
Thebaine	2.53
Noscapine	3.16

20

Virus induced gene silencing (VIGS) of *PSMT1* and *PSMT2*

25 a) Generation of silencing constructs

A tobacco rattle virus (TRV) based virus induced gene silencing system developed and described by Liu et al. (2002) Plant J. 30(4): 415-429 was used to investigate the gene

function of *PSMT1* and *PSMT2*. DNA fragments selected for silencing of *PSMT1* and *PSMT2*, respectively, were amplified by PCR and cloned into the silencing vector pTRV2 (GenBank accession no. AF406991; Liu et al. (2002) Plant J. 30(4): 415-429). They were linked to a 129 bp-long fragment of the *P. somniferum Phytoene desaturase* gene (*PsPDS*) in order to silence the respective methyltransferases and *PsPDS* simultaneously. Plants displaying the photo-bleaching phenotype that resulted from silencing of *PsPDS* (Hileman et al. (2005) Plant J. 44(2): 334-341) were identified as plants successfully infected with the respective silencing constructs and selected for analysis.

10

Generation of the pTRV2-PDS construct: A 622 bp fragment (Figure 13) of *PsPDS* was amplified from cDNA prepared from GSK NOSCAPHINE CVS1 as described above using primers ps_pds_F and ps_pds_R4 (Table 4). The sequence of the forward primer was based on a 412 bp long contig derived from the EST-libraries which shared 99% identity at its 3' end with the partial coding sequence of the *P. somniferum phytoene desaturase* (GenBank accession no. DQ116056). The sequence of the reverse primer was designed based on the DQ116056 sequence. The PCR conditions were identical to those described above for the amplification of *PSMT1* from genomic sequence except that the annealing step was carried out at 70°C and the extension time was increased to 60 seconds.

20

Sau3AI digestion of the PCR-fragment yielded among others two fragments (280 bp and 129 bp in length) that carried BamHI-compatible sticky ends at both, their 5' and 3' ends. The 129 bp long fragment (underlined stretch in Figure 13) was cloned into the BamHI site of the pTRV2 vector. Because Sau3AI was used to produce BamHI-compatible sticky ends, the BamHI site at the 5-end of the *PDS*-insert was abolished in the pYL156-PDS construct. However, the BamHI recognition site at its 3'-end was kept intact due to the nature of the *PDS*-insert sequence.

25

A sequence-confirmed pTRV2-PDS construct, with the 129 bp fragment in sense orientation, was subsequently used as a vector for generating the *PSMT1* and *PSMT2* silencing constructs, and served as a control in the VIGS experiments.

30

Generation of silencing constructs for *PS PSMT1* and *PSMT2* (pTRV2-PDS-*PSMT1* and pTRV2-PDS-*PSMT2*): The DNA fragments selected for silencing *PSSOMT1* and *PSSOMT2* were amplified from cDNA of GSK NOSCAPHINE CVS1 prepared as

35

described above with the use of the primer sequences shown in Table 5. Additional restriction sites (forward primers: XhoI and HindIII for forward primers; KpnI site for reverse primers) were added to the gene-specific primers in order to facilitate cloning. The amplification conditions were as described above for amplifying the *PDS*-fragment except that the annealing temperatures were 60.9°C for *PSMT1* and 62.8°C for *PSMT2* and the extension time was 30 seconds.

The sequence selected to silence *PSMT1* (Figure 14) was cloned into pTV00 (Ratcliff et al. (2001) Plant J. 25(2): 237–245) using HindIII and KpnI and subcloned into pTRV2-PDS using BamHI and KpnI. A sequence-confirmed pTRV2-PDS-*PSMT1* construct was used in the VIGS experiments.

The sequence selected for silencing *PSMT2* (Figure 15) was cloned directly into pTRV2-PDS. Since the PCR amplification was carried out with Phusion (Finnzymes, Finland) the resulting PCR fragment carried blunt ends and, after cutting with XhoI, was cloned into pYL156-PDS cut with XhoI and SmaI. A sequence-confirmed pTRV2-PDS-*PSMT2* construct was used in the VIGS experiments.

b) Transformation of constructs into *Agrobacterium tumefaciens*

The propagation of the silencing constructs was carried out with the *E. coli* strain DH5α and, subsequently, the respective silencing constructs, as well as pTRV1 (GenBank accession no. AF406990; Liu et al. (2002) Plant J. 30(4): 415-429) were independently transformed into electrocompetent *Agrobacterium tumefaciens* (strain GV3101).

c) Infiltration of plants

Overnight liquid cultures of *A. tumefaciens* containing each silencing construct were used to inoculate Luria-Bertani (LB) medium containing 10 mM MES, 20 μM acetosyringone and 50 μg/ml kanamycin. Cultures were maintained at 28°C for 24 hours, harvested by centrifugation at 3000g for 20 min, and resuspended in infiltration solution (10 mM MES, 200 μM acetosyringone, 10 mM MgCl₂) to an OD₆₀₀ of 2.5. *A. tumefaciens* harbouring the respective constructs (pTRV2-PDS-*PSMT1*, pTRV2-PDS-*PSMT2* or, as a control, pTRV2-PDS) were each mixed 1:1 (v/v) with *A. tumefaciens* containing pTRV1, and incubated for two hours at 22°C prior to infiltration. Two weeks old seedlings of GSK NOSCAPHINE CVS1 grown under standard greenhouse conditions (22°C, 16h photoperiod), with emerging first leaves, were infiltrated as described by Hagel and Facchini (2010) Nat. Chem. Biol. 6: 273–275.

d) Latex and capsule analysis of silenced plants

Leaf latex of infiltrated opium poppy plants displaying photo-bleaching as a visual marker for successful infection and silencing was analysed when the first flower buds emerged (~7 week old plants). Plants showing a similar degree of photo-bleaching of leaves were selected for analysis.

Latex was collected from cut petioles, with a single drop dispersed into 500 μ L 10% acetic acid. This was diluted 10x in 1% acetic acid to give an alkaloid solution in 2% acetic acid for further analysis. Capsules were harvested by hand from glasshouse-grown from the same plants used for latex analysis and single capsules were ground in a Retsch Model MM04 ball mill into a fine powder. Ten mg samples of ground poppy straw were then weighed accurately (10 ± 0.1 mg) and extracted in 0.5 mL of a 10% acetic acid solution with gentle shaking for 1h at room temperature. Samples were then clarified by centrifugation and a 50 μ L subsample diluted 10x in 1% acetic acid to give an alkaloid solution in 2% acetic acid for further analysis.

All solutions were analysed using a Waters Acquity UPLC system fitted with a Waters Acquity BEH C18 column, 2.1mm x 100mm with 1.7 micron packing. The mobile phase used a gradient profile with eluent A consisting of 10mM ammonium bicarbonate pH 10.2 and eluent B methanol. The mobile phase gradient conditions used are as listed in Table 1, with a linear gradient. The flow rate was 0.5mL per minute and the column maintained at 60°C. The injection volume was 2 μ L and eluted peaks were ionised in positive APCI mode and detected within ~3 ppm mass accuracy using a Thermo LTQ-Orbitrap. The runs were controlled by Thermo Xcalibur software.

– Gradient Flow Program:

TIME (minutes)	% Eluent A	% Eluent B	Flow (mL/min)
0.0	98.0	2.0	0.50
0.2	98.0	2.0	0.50
0.5	60.0	40	0.50
4.0	20.0	80.0	0.50
4.5	20.0	0.0	0.50

All data analysis was carried out in R. Putative alkaloid peaks were quantified by their pseudomolecular ion areas using custom scripts. Peak lists were compiled and any

peak-wise significant differences between samples were identified using 1-way ANOVA with p-values adjusted using the Bonferroni correction for the number of unique peaks in the data set. For any peak-wise comparisons with adjusted p-values < 0.05, Tukey's HSD test was used to identify peaks that were significantly different between any given sample and the control. Alkaloids were identified by comparing exact mass and retention time values to those of standards. Where standards were not available, neutral exact masses were used to generate molecular formulae hits within elemental constraints of C = 1:100, H = 1:200, O = 0:200, N = 0:3 and mass accuracy < 20ppm. The hit with the lowest ppm error within these constraints was used to assign a putative formula.

Example 1

Assembly of full length *PSMT1* cDNA sequence from ESTs and confirmation by sequencing of full length cDNA

The full length cDNA sequence for *PSMT1* (Figure 1a) was assembled from ESTs derived from the 454 sequencing platform using the CAP3 sequence assembly programme. The full length cDNA sequence was confirmed by RT-PCR and direct amplification of the full length cDNA from GSK NOSAPINE CVS1 mRNA.

Example 2

PSMT1 is exclusively expressed in the noscapine producing *Papaver somniferum* cultivar GSK NOSAPINE CVS1.

Figure 2 shows the normalized distribution of ESTs associated with the *PSMT1* consensus sequence across each of the 16 EST libraries prepared from two organs (capsules and stems) at two developmental stages (early and late harvest) from each of the four poppy cultivars, GSK MORHINE CVS1, GSK MORPHINE CVS2, GSK NOSAPINE CVS1 and GSK THEBAINE CVS1. ESTs corresponding to *PsSOMT* were exclusively found in libraries derived from the noscapine producing cultivar GSK NOSAPINE CVS1 (Figure 2). *PSMT1* expression was strongest in stem tissue shortly after flowering.

Example 3

PCR-amplification of *PSMT1* from genomic DNA of the four *Papaver somniferum* cultivars GSK NOSAPINE CVS1, GSK MORHINE CVS1, GSK MORPHINE CVS2 and
5 GSK THEBAINE CVS1.

PCR-amplifications of *PSMT1* fragments were performed on genomic DNA from the four poppy cultivars GSK NOSAPINE CVS1 (noscapine cultivar), GSK MORHINE CVS1 and GSK MORPHINE CVS2 (morphine cultivars) and GSK THEBAINE CVS1 (thebaine
10 cultivar) using the primer combinations shown in Table 2 and 3.

Figure 5 shows the PCR-amplification of *PSMT1* from genomic DNA of the four *Papaver somniferum* cultivars GSK NOSAPINE CVS1, GSK MORHINE CVS1, GSK MORPHINE CVS2 and GSK THEBAINE CVS1.

15 The amplification from genomic DNA yielded the gene sequence shown in Figure 3a.

Example 4

The putative protein encoded by *PSMT1* shows highest sequence similarity to (S)-scoulerine 9-O-methyltransferases from *Coptis japonica* and *Thalictrum flavum*

20 The closest homologues to the putative protein encoded by the *PSMT1* open reading frame (Figure 4) are (S)-scoulerine 9-O-methyltransferases from *Coptis japonica* (Accession: Q39522.1, Evalue: 1e-120, 61% identical) and from *Thalictrum flavum* (Accession: AAU20770.1, Evalue: 5e-118 59% identical).

25

Example 5

PSMT1 is only present in the genome of the noscapine producing *P. somniferum* cultivar GSK NOSAPINE CVS1

30 The transcribed region covered by the ESTs contained the complete coding sequence of *PSMT1* (including 5' and 3' untranslated regions), which was used for primer design (Table 1) to amplify the *PSMT1* gene from genomic DNA in a series of overlapping fragments for sequencing (Figure 3a). Upon testing a subset of the primer combinations (Table 3) on genomic DNA samples from all four cultivars it was discovered that the
35 *PSMT1* fragments could only be amplified from genomic DNA of the noscapine producing cultivar GSK NOSAPINE CVS1 but not from genomic DNA of the

predominantly morphine (GSK MORPHINE CVS1 and GSK MORPHINE CVS2) or thebaine (GSK THEBAINE CVS1) producing cultivars (Figure 5). The PCR amplifications were performed on pools of genomic DNA comprising DNA from four individuals per cultivar. This discovery explains why the *PSMT1* is only expressed in the GSK NOSCAPINE CVS1 cultivar and is absent from the transcriptome of the other three cultivars.

Example 6

Assembly of full length *PSMT2* cDNA sequence from ESTs and confirmation by sequencing from genomic DNA

The full length cDNA sequence for *PSMT2* (Figure 1b) was assembled from ESTs derived from the 454 sequencing platform using the CAP3 sequence assembly programme. The full length cDNA sequence was confirmed by amplification and sequencing from genomic DNA.

Example 7

***PSMT2* is exclusively expressed in the noscapine producing *Papaver somniferum* cultivar GSK NOSCAPINE CVS1**

Figure 9 shows the normalized distribution of ESTs associated with the *PSMT2* consensus sequence across each of the 16 EST libraries prepared from two organs (capsules and stems) at two developmental stages (early and late harvest) from each of the four poppy cultivars, GSK MORPHINE CVS1, GSK MORPHINE CVS2, GSK NOSCAPINE CVS1 and GSK THEBAINE CVS1. ESTs corresponding to *PSMT2* were exclusively found in libraries derived from the noscapine producing cultivar GSK NOSCAPINE CVS1. *PSMT2* expression was strongest in stem tissue shortly after flowering.

Example 8

PCR-amplification of *PSMT2* from genomic DNA of the four *Papaver somniferum* cultivars GSK MORPHINE CVS1, GSK MORPHINE CVS2, GSK NOSCAPINE CVS1 and GSK THEBAINE CVS1.

PCR-amplifications of *PSMT2* fragments were performed on genomic DNA from the four poppy cultivars GSK MORPHINE CVS1, GSK MORPHINE CVS2, GSK

NOSCAPINE CVS1 and GSK THEBAINE CVS1 using the primer combinations shown in Table 2 and 3.

Figure 10 shows the PCR-amplification of *PSMT2* from genomic DNA of the four
5 *Papaver somniferum* cultivars GSK MORPHINE CVS1, GSK MORPHINE CVS2, GSK
NOSCAPINE CVS1 and GSK THEBAINE CVS1;

The amplification from genomic DNA yielded the gene sequence shown in Figure 3b.

10 Example 9

The putative protein encoded by *PSMT2* shows highest sequence similarity to
norcoclaurine 6-O-methyltransferases from *Coptis japonica*, *Papaver bracteatum*
and *Papaver somniferum*

15 The closest homologues to the putative protein encoded by the *PSMT2* open reading
frame (Figure 1b) are (R,S)-norcoclaurine 6-O-methyltransferase from *Coptis japonica*
(accession: Q9LEL6, identities: 146/355 (42%)), putative norcoclaurine 6-O-
methyltransferase from *Papaver bracteatum* (accession: ACO90232, identities: 151/346
(44%)) and S-adenosyl-L-methionine:norcoclaurine 6-O-methyltransferase from *Papaver*
20 *somniferum* (accession: AAP45315, identities = 148/343 (44%)). The sequence
comparisons were carried out using NCBI's 'blastp' algorithm (method: compositional
matrix adjust).

25 Example 10

***PSMT2* is only present in the genome of the noscapine producing *P. somniferum*
cultivar GSK NOSCAPINE CVS1.**

The transcribed region covered by the ESTs contained the complete coding sequence of
30 *PSMT2* (including 5' and 3' untranslated regions), which was used for primer design
(Table 1) to amplify the *PSMT2* gene from genomic DNA in a series of overlapping
fragments for sequencing (Figure 3b). Upon testing a subset of the primer combinations
on genomic DNA samples from all four cultivars it was discovered that the *PsSOMT2*
fragments could only be amplified from genomic DNA of the noscapine producing
35 cultivar GSK NOSCAPINE CVS1 but not from genomic DNA of the predominantly
morphine (GSK MORPHINE CVS1, GSK MORPHINE) or thebaine (GSK THEBAINE

CVS1) producing cultivars (Figure 10). The PCR amplifications were performed on pools of genomic DNA comprising DNA from four individuals per cultivar using primer combinations shown in Table 3. This discovery explains why the *PSMT2* is only expressed in the GSK NOSCAPHINE CVS1 cultivar and is absent from the transcriptome of the other three cultivars.

Example 11

Assembly of full length *PSMT3* cDNA sequence from ESTs and confirmation by sequencing from genomic DNA

10

The full length cDNA sequence for *PSMT3* (Figure 1c) was assembled from ESTs derived from the 454 sequencing platform using the CAP3 sequence assembly programme. The full length cDNA sequence was confirmed by amplification and sequencing from genomic DNA.

15

Example 12

***PSMT3* is exclusively expressed in the noscapine producing *Papaver somniferum* cultivar GSK NOSCAPHINE CVS1.**

Figure 11 shows the normalized distribution of ESTs associated with the *PSMT3* consensus sequence across each of the 16 EST libraries prepared from two organs (capsules and stems) at two developmental stages (early and late harvest) from each of the four poppy cultivars, GSK MORPHINE CVS1, GSK MORPHINE CVS2, GSK NOSCAPHINE CVS1 and GSK THEBAINE CVS1. ESTs corresponding to *PsSOMT3* were exclusively found in libraries derived from the noscapine producing cultivar GSK NOSCAPHINE CVS1 (Figure 11). *PSMT3* expression was strongest in stem tissue shortly after flowering.

Example 13

PCR-amplification of *PSMT3* from genomic DNA of the four *Papaver somniferum* cultivars GSK MORPHINE CVS1, GSK MORPHINE CVS2, GSK NOSCAPHINE CVS1 and GSK THEBAINE CVS1.

PCR-amplifications of *PSMT3* fragments were performed on genomic DNA from the four poppy cultivars GSK MORPHINE CVS1, GSK MORPHINE CVS2, GSK NOSCAPHINE

CVS1 and GSK THEBAINE CVS1 using the primer combinations shown in Table 2 and 3.

Figure 12 shows the PCR-amplification of *PSMT3* from genomic DNA of the four
5 *Papaver somniferum* cultivars GSK MORPHINE CVS1, GSK MORPHINE CVS2, GSK
NOSCAPINE CVS1 and GSK THEBAINE CVS1;

The amplification from genomic DNA yielded the gene sequence shown in Figure 3c.

10 Example 14

The putative protein encoded by *PSMT3* shows highest sequence similarity to
norcoclaurine 6-O-methyltransferases from *Papaver somniferum* and *Papaver*
bracteatum

15 The closest homologues to the putative protein encoded by the *PSMT3* open reading
frame (Figure 4) are (R,S)-norcoclaurine 6-O-methyltransferases from *Papaver*
somniferum (accession: AAQ01669, identities: 275/346 (80%), *Papaver somniferum*
(accession: AAP45315, identities = 275/346 (80%), and *Papaver bracteatum* (accession:
ACO90232, identities: 274/346 (80%)). The sequence comparisons were carried out
20 using NCBI's 'blastp' algorithm (method: compositional matrix adjust).

Example 15

PSMT3* is only present in the genome of the noscapine producing *P. somniferum
25 **cultivar GSK NOSCAPINE CVS1.**

The transcribed region covered by the ESTs contained the complete coding sequence of
PSMT3 (including 5' and 3' untranslated regions), which was used for primer design
(Table 1) to amplify the *PSMT3* gene from genomic DNA in a series of overlapping
30 fragments for sequencing (Figure 3c). Upon testing a subset of the primer combinations
on genomic DNA samples from all four cultivars it was discovered that the *PSMT3*
fragments could only be amplified from genomic DNA of the noscapine producing
cultivar GSK NOSCAPINE CVS1 but not from genomic DNA of the predominantly
morphine (GSK MORPHINE CVS1, GSK MORPHINE) or thebaine (GSK THEBAINE
35 CVS1) producing cultivars (Figure 12). The PCR amplifications were performed on pools
of genomic DNA comprising DNA from four individuals per cultivar using primer

combinations shown in Table 3. This discovery explains why the *PSMT3* is only expressed in the GSK NOSCAPHINE CVS1 cultivar and is absent from the transcriptome of the other three cultivars.

5 **Example 16**

Scoulerine and putative narcotoline accumulate in *PSMT1*- and *PSMT2*-silenced plants, respectively. Virus induced gene silencing led to the accumulation of scoulerine in both latex and mature capsules of *PSMT1*-silenced plants but not of *PSMT2*-silenced plants, PDS-silenced control plants or uninfected plants of GSK NOSCAPHINE CVS1 (Figure
10 16). Scoulerine has been shown to be an intermediate in the biosynthesis of noscapine (Battersby et al. (1968) J. Chem. Soc. (C): 2163–2172).

Putative narcotoline (4-desmethyl-noscapine) accumulated in latex and capsules of *PSMT2*-silenced plants but not of *PSMT1*-silenced plants, PDS-silenced control plants or
15 uninfected plants of GSK NOSCAPHINE CVS1 (Figure 17).

Example 17

PSMT1, *PSMT2* and *PSMT3* are linked and involved in noscapine biosynthesis
20

All three methyltransferase genes are present in all F2 plants that do contain noscapine. All F2 plants that do not contain the methyltransferase genes do not make noscapine (Figure 6a-c).

25

30

35

Claims

1. An isolated polypeptide selected from the group consisting of:
- 5 i) a polypeptide comprising or consisting of an amino acid sequence as represented in Figure 4a, 4b or 4c; or
- ii) a modified polypeptide comprising or consisting of a modified amino acid sequence wherein said polypeptide is modified by addition, deletion or substitution of at least one amino acid residue
- 10 of the sequence presented in Figure 4a, 4b or 4c and which has retained or enhanced methyltransferase activity.
2. An isolated nucleic acid molecule that encodes a methyltransferase polypeptide wherein said nucleic acid molecule comprises or consists of a nucleotide sequence
- 15 selected from the group consisting of:
- i) a nucleotide sequence as represented by the sequence in Figure 1a, 1b, 1c, 3a, 3b or 3c;
- ii) a nucleotide sequence wherein said sequence is degenerate as a result of the genetic code to the nucleotide sequence defined in
- 20 (i);
- iii) a nucleic acid molecule the complementary strand of which hybridizes under stringent hybridization conditions to the sequence in Figure 1a, 1b, 1c, 3a, 3b or 3c wherein said nucleic acid molecule encodes a methyltransferase polypeptide;
- 25 iv) a nucleotide sequence that encodes a polypeptide comprising an amino acid sequence as represented in Figure 4a, 4b or 4c;
- v) a nucleotide sequence that encodes a polypeptide comprising an amino acid sequence wherein said amino acid sequence is modified by addition deletion or substitution of at least one amino acid residue as represented in iv) above and which has retained or
- 30 enhanced methy transferase activity.
3. The isolated nucleic acid molecule according to claim 2 wherein said nucleic acid molecule comprises or consists of a nucleotide sequence as represented in Figure 1a,
- 35 1b, 1c, 3a, 3b or 3c.

4. A vector comprising a nucleic acid molecule encoding a methyltransferase according to claim 2 or 3 wherein said nucleic acid molecule is operably linked with a nucleic acid sequence comprising a promoter sequence.
- 5
5. The vector according to claim 4 wherein said nucleic acid sequence comprising a promoter confers constitutive expression on said methyltransferase.
6. The vector according to claim 4 wherein said nucleic acid sequence comprising a promoter confers regulated expression on said methyltransferase.
- 10
7. The vector according to claim 6 wherein said regulated expression is tissue or developmentally regulated expression.
- 15
8. The vector according to claim 6 wherein said regulated expression is inducible expression.
9. The transgenic cell transformed or transfected with a nucleic acid molecule or vector according to any one of claims 2 to 8.
- 20
10. A transgenic cell according to claim 9 wherein said cell is a plant cell.
11. The transgenic cell according to claim 10 wherein said plant cell is from the family *Papaveraceae*.
- 25
12. The transgenic cell according to claim 11 wherein said plant cell is a *Papaver somniferum* cell.
13. A plant comprising a plant cell according to any one of claims 10 to 12.
- 30
14. The plant according to claim 13 wherein said plant is from the family *Papaveraceae*.
15. The plant according to claim 14 wherein said plant is *Papaver somniferum*.
- 35
16. The transgenic cell according to claim 9 wherein said cell is a microbial cell.

17. The transgenic cell according to claim 16 wherein said microbial cell is a bacterial cell.

5 18. The transgenic cell according to claim 16 wherein said microbial cell is a fungal cell.

19. A nucleic acid molecule comprising a transcription cassette wherein said cassette includes a nucleotide sequence designed with reference to Figure 1a, 1b, 1c, 3a, 3b or
10 3c and is adapted for expression by provision of at least one promoter operably linked to said nucleotide sequence such that both sense and antisense molecules are transcribed from said cassette.

20. The nucleic acid molecule according to claim 19 wherein said cassette is adapted
15 such that both sense and antisense nucleic acid molecules are transcribed from said cassette wherein said sense and antisense nucleic acid molecules are adapted to anneal over at least part or all of their length to form a siRNA or shRNA.

21. The nucleic acid molecule according to claim 19 or 20 wherein said cassette is
20 provided with at least two promoters adapted to transcribe both sense and antisense strands of said nucleic acid molecule.

22. The nucleic acid molecule according to any of claims 19-21 wherein said cassette comprises a nucleic acid molecule wherein said molecule comprises a first part
25 linked to a second part wherein said first and second parts are complementary over at least part of their sequence and further wherein transcription of said nucleic acid molecule produces an RNA molecule which forms a double stranded region by complementary base pairing of said first and second parts thereby forming an shRNA.

30 23. The nucleic acid molecule according to any of claims 19-22 wherein said nucleic acid molecule is part of a vector adapted for expression in a plant cell.

24. The plant cell transfected with a nucleic acid molecule or vector according to any of claims 19-23 wherein said cell has reduced expression of said methyltransferase.

25. A process for the methylation of an opiate alkaloid comprising:
- 5 i) providing a transgenic plant cell according to any one of claims 10 to 12;
- ii) cultivating said plant cell to produce a transgenic plant; and optionally
- ii) harvesting said transgenic plant, or part thereof.
26. The process according to claim 25 wherein said harvested plant material is dried straw and said opiate alkaloid is extracted.
- 10
27. A process for the methylation of an opiate alkaloid comprising:
- i) providing a transgenic microbial cell according to any one of claims 16 to 18 that expresses a methy transferase according to claim 1 in culture with at least one opiate alkaloid;
- 15 ii) cultivating the microbial cell under conditions that methylate one or more opiate alkaloids; and optionally
- iii) isolating said methylated alkaloid from the microbial cell or cell culture.
- 20 28 The process according to claim 27 wherein said microbial cell is a bacterial cell or fungal/yeast cell.
29. The use of a gene encoded by a nucleic acid molecule as represented by the nucleic acid sequence in Figure 3a, 3b or 3c, or a nucleic acid molecule that hybridizes under stringent hybridization conditions to the nucleotide sequence in Figure 3a, 3b or
- 25 3c and encodes a polypeptide with methyltransferase activity as a means to identify the presence or absence of a gene that encodes said methyltransferase in a *Papaveraceae* plant.
30. A method to determine the presence or absence of a gene according to
- 30 the invention in a *Papaveraceae* variety comprising:
- i) obtaining a sample from a *Papaveraceae* plant;
- ii) extracting genomic DNA from the plant; and
- iii) analyzing the genomic DNA for the presence of a gene comprising or consisting of a nucleotide sequence as represented in Figure 3a, 3b or 3c.

31. The use of a gene encoded by a nucleic acid molecule as represented by the nucleic acid sequence in Figure 3a, 3b or 3c, or a nucleic acid molecule that hybridizes under stringent hybridization conditions to the nucleotide sequence in Figure 3a, 3b or 3c and encodes a polypeptide with methyltransferase activity as a means to identify a locus wherein said locus is associated with altered expression or activity of said methyltransferase.

32. A method to produce a *Papaveraceae* plant variety that has altered expression of a methyltransferase according to the invention comprising the steps of:

- i) mutagenesis of wild-type seed from a plant that does express said methyltransferase;
- ii) cultivation of the seed in i) to produce first and subsequent generations of plants;
- iii) obtaining seed from the first generation plant and subsequent generations of plants ;
- iv) determining if the seed from said first and subsequent generations of plants has altered expression of said methyltransferase;
- v) obtaining a sample and analysing the nucleic acid sequence of a nucleic acid molecule selected from the group consisting of:
 - a) a nucleic acid molecule comprising a nucleotide sequence as represented in Figure 3a, 3b or 3c;
 - b) a nucleic acid molecule that hybridises to the nucleic acid molecule in a) under stringent hybridisation conditions and that encodes a polypeptide with methyltransferase activity; and optionally
- vi) comparing the nucleotide sequence of the nucleic acid molecule in said sample to a nucleotide sequence of a nucleic acid molecule of the original wild-type plant.

33. A method according to claim 32 wherein said nucleic acid molecule is analysed by a method comprising the steps of:

- i) extracting nucleic acid from said mutated plants;
- ii) amplification of a part of said nucleic acid molecule by a polymerase chain reaction;

- iii) forming a preparation comprising the amplified nucleic acid and nucleic acid extracted from wild-type seed to form heteroduplex nucleic acid;
iv) incubating said preparation with a single stranded nuclease that cuts at a region of heteroduplex nucleic acid to identify the mismatch in said heteroduplex; and
v) determining the site of the mismatch in said nucleic acid heteroduplex.

34. The method according to claim 33 wherein said *Papaveraceae* plant variety has enhanced methyltransferase expression and/or activity.

10

35. A plant obtained by the method according to any one of claims 32 to 34.

36. A plant wherein said plant comprises a viral vector that includes a nucleic acid molecule wherein said nucleic acid molecule comprise at least part of a gene that encodes a methyltransferase polypeptide

15

37. The plant according to claim 36 wherein said gene is encoded by a nucleic acid molecule comprising a nucleic acid sequence selected from the group consisting of:

20 i) a nucleic acid molecule comprising a nucleotide sequence as represented in Figure 1a, 1b or 1c ;

ii) a nucleic acid molecule comprising a nucleotide sequence that hybridises under stringent hybridisation conditions to a nucleic acid molecule in (i) and which encodes a methyltransferase polypeptide;

25 iii) a nucleic acid molecule that encodes a variant polypeptide that varies from a polypeptide comprising the amino acid sequence as represented in Figure 4a, 4b or 4c.

38. The plant according to claim 37 wherein said nucleic acid molecule comprises a nucleotide sequence as represented in Figure 1a.

30

39. The plant according to claim 37 wherein said nucleic acid molecule comprises a nucleotide sequence as represented in Figure 1b.

35

40. The plant according to claim 37 wherein said nucleic acid molecule comprises a nucleotide sequence as represented in Figure 1c

41. The plant according to claim any one of claims 36 to 38 wherein said nucleic acid molecule comprises or consists of a nucleotide sequence as represented in Figure 14.

5

42. The plant according to claim any one of claims 36, 37 or 39 wherein said nucleic acid molecule comprises or consists of a nucleotide sequence as represented in Figure 15.

10 43. A viral vector comprising all or part of a nucleic acid molecule according to claim 2 or 3.

44. The use of a viral vector according to claim 43 in viral induced gene silencing in a plant.

15

45. Use according to claim 44 wherein said plant is from the family *Papaveraceae*.

20

25

30

35

40

45

Figure 1a

1/14

ATGGCTACCAATGGCGAAATTTTCAATACCTATGGTCATAATCATCAATCAGCCACAGTCACTAAAATCACTGCT
TCTAATGAAAGCAGCAATGGTGTCTGTTATCTTTCAGAAACGGCTAACTTGGGGAAGTTAATATGCATTCCAATG
GCTAAGAGCTGCGATGGAGCTAAATGTGTTCCAATTATCTCAAAGTTCGGAAGTACGCAAAAAGTTTCGGCT
TCTGAAATTGCCTCTAAAATGCCAAACGCGAAGAATAATCCTGAAGCAGCTATGTATTTGGATAGAATTCCTCGA
CTGCTCGGGGCAAGTTCTATTCTTTCTGTTTCTACTACAAAAAATCAATCAACAGAGGAGGAGATGATGTAGTA
GTACATGAGAAGCTTTATGGGTAAACAAATTCGTCGTGTTGTTTGGTCCCTCGACAAGAAGACGGGGTGTCATTA
GTCGAAGAATTGCTATTACATCTGACAAGGTTGTTGTGGATAGTTTTTTTCAAACCTGAAATGTGTGGTGGAAGAA
AAAGACAGTGTGCCATTTGAGGTTGCTCATGGTGCTAAAACTTTGAGTATGCTGCTACAGAACCAAGAATGAAT
CAAGTATTTAACGATGGAATGGCAGTTTTCTCTATTGTTGTTTTTGGAGCTGTTTTTAGAGTTTACGATGGATTT
CTTGATATGAAAGAATTGTTAGATGTTGGTGGTGGTATTGGTACTTCGGTTAGTAAGATTGTTGCTAAATACCCCT
TTGATTTCGCGGTGTCAACTTCGACTTGCCCTCATGTTATTTCTGTTGCCCTCAATACCCAGGTGTAGAGCATGTT
GCAGGAGATATGTTTCGAGGAAGTCCCAAAGGGTCAAAACATGTTGCTAAAAATGGGTACTGCACGATTGGGGTGAT
GAACGATGTGTGAAGCTGTTAAAGAATTGTTGGAACCTATTACCTGTGGGTGGAAAAGTTTTGATAATCGAGTTT
GTTCTCCCGAATGAACCTGGTAACAATGCTGAATCATTCAATGCGTTGATTCCCGATTACTCCTGATGGCTCTG
AATCCAGGCGGTAAAGAGCGAACGATTTCCGAATACGATGATTTTAGGCAAAGCAGCTGGATTCAAAAACTATA
CCTATCCCTATCTCCAATGGTCTTCATGTCATTGAGTTTCACAAATGA

Figure 1b

ATGGAAATTCATTTAGAAAGCCAAGAACAAGAAATGAAATATCAATCTCAAATCTGGAACCAAATATGTGGCACT
GTTGATACCTCTGTTCTGAGATGTGCAATTCAATTAGGTATATTTGATGCCATTCACTAACTCTGGCAAACCAATG
ATTACCTTAACCGAATTATCAAGCATTGTTTCATCACCCCTCTTCATCTTCAATCGAACCCCTGCAACTTGTATAGA
TTAGTGAGATACTTATCCCAAATGGATCTCATTAGTATCGGAGAATGTTTGAATGAAGCAACTGTTTCATTAACA
GGCACATCCAAGTTACTACTTAGAAACCAAGAAAAGAGTTTAATTGATTGGGTATTGGCAATTTCTTGCGAAATG
ATGGTTGTTGTTTGGCACGAACCTAAGTAGCTCTGTTTCAACTCCTGCGGATGAGCCTCCAATCTTCCAGAAGGTT
CATGGTAAAAATGCTTTAGAATTAGCAGGGGAATTTCCAGAATGGAATGATCTGATCAACAATGCTATGACTAGT
GATACTAGTGTAACCTAAGCCAGCGCTAATACAAGGATGTGGCAAAATCCTGAACGGAGTTACATCGTTAATTGAT
GTCGGTGGTGGTCACGGTGCCACTATGGCCTACATAGTTGAAGCTTTTCTCACATAAAAAGGTGCGGTAATCGAT
TTACCACATGTTGTTGAAGCCGCTCCGGAGCGTCCAGGTGTTGAGTTTCATCAGCGGTGATATATTCAAGTCCATT
TCTAACGCTGATGCTGTGTTGTTGAAGTATGTCTGCACAATTGGGAAGATACGGAATGTGTGAATTTACTGAAG
AGATGTAAGGAAGCAGTTCCGGCAGACAAAGGAAAAGTGATCATAATGGATTTAGTAATAGACGACGATGATAAC
AGTATTTTAACGCAGGCAAAGTTGAGCCTTGATCTCACTGTGATGAACCATGGAGGAGGTAGAGAAAGGACTAAA
GAAGATTGGAGAAATCTAATTGAGATGTCTGGATTTAGTAGGCATGAAATAATTCCAATATCTGCCATGCCATCA
ATTATTGTTGCTTATCCTTAG

Figure 1c

ATGGAAGTAGTAAGTAAGATTGATCAAGAAAACCAAGCAAAAATTTGGAAACAAATTTTTGGTTTTGCAGAATCA
TTAGTTCTAAAATGTGCAGTTCAGTTAGAGATTGCTGAAACACTTCATAATAATGTAAACCCATGTCTTTATCC
GAGTTAGCATCTAAACTTCCGGCTCAACCCGTTAATGAAGACCGTTTGTACCGAATTCTGCATTTCTTAGTTTAC
ATGAAACTCTTCAACAAAGATGCTACCACACAGAAATATTCAATTAGCTCCACCAGCAAAGTATTTGCTAAAAGGC
TGGGAAAAATCAATGGTTCTTCAATATTAAGCGTGACTGATAAAGATTTTACAGCTCCATGGAATCATCTTGGG
GACGGTTTTAACGGTAACGTGTAACGCTTTTGTAGAAAAGCGTTAGGAAAAGGCATTCCGGGTTTATATGAGAGAAAAT
CCTGAAAAAGATCAATTGTTTAAATGAAGGAATGGCTTGTGATACTAGATTATTTGCTTCAGCATTGGTTAACGAG
TGCAAAAGTATTTTCAGTGACGGGATCAATACACTTGCCGGTGTGGCCGTGGTACTGGTACTGCAGTGAAAGCC
ATATCCAAAGCTTTTCCGGATATTAAGTGCACAATCCATGATCTTCTGAAGTTACCAGTAAAAATAGTAAATTT
CCAAGAGATGTTTTTAAGTCCGTTCTAGTGCAGACGCCATCTTTATGAAGAGCATTCTTCACGAATGGAACGAT
GAGGAATGTATTTCAAATCTTGAAACGATGCAAGAAGCAATACCAAAAGGGGGCAAAGTTATCATTGCGGATGTC
GTAATAGACATGGACTCGACTCATCCGTATTCAAAATCTAGACTCGCAATGGATTTGGCTATGATGCTCCACACG
GGTGGAAAAGAGAGAACTGAAGAAGATTGGAAAAAATTTATTGATGCTGCAGGTTTTGCTAGCTGTAAAATTACT
AAACTATCTGCTCTCCAGTCTGTTATTGAGGCTTACCCCTATTGA

Figure 2

2/14

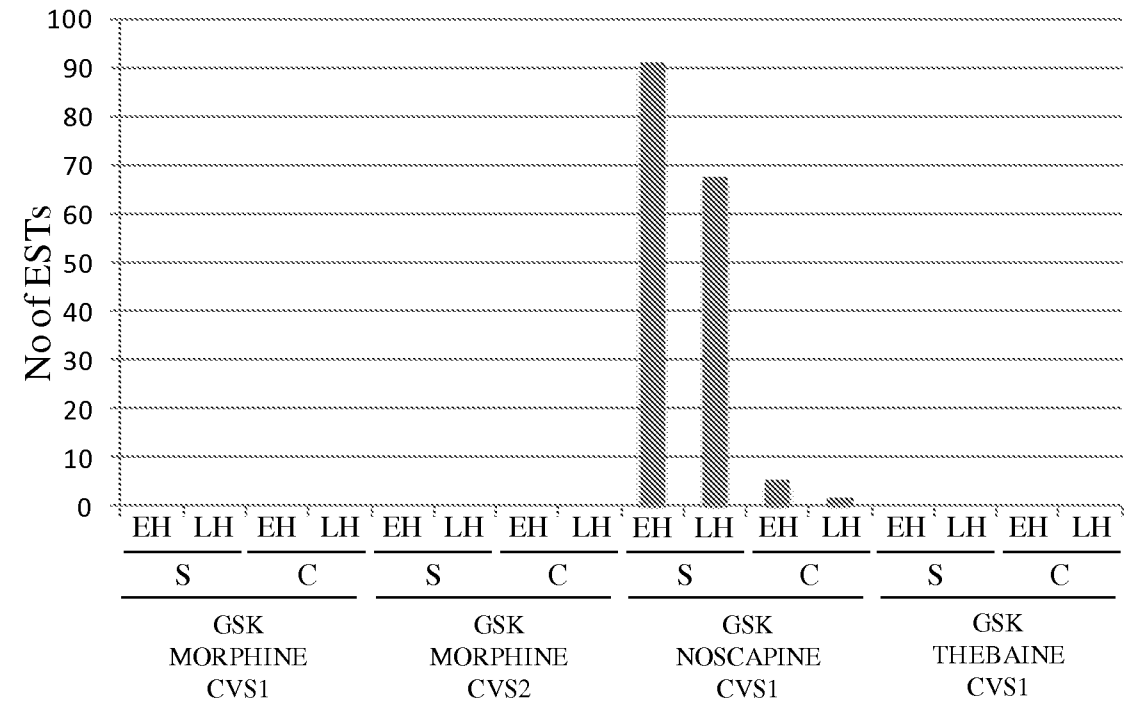


Figure 3a

3/14

Start	End	Feature
1	66	5' untranslated region
67	877	exon 1
878	993	intron 1
994	1058	exon 2
1059	1841	intron 2
1842	2138	exon 3
2139	2306	3' untranslated region

CACACCAAACCTTGATCATTGTCATAAAAAACAGTCCTAATTGTCATCAATCAAAAAACAGTCCTAACATGGCTACC
AATGGCGAAATTTTCAATACCTATGGTCATAATCATCAATCAGCCACAGTCCTAAAACTACTGCTTCTAATGAA
AGCAGCAATGGTGTCTGTTATCTTTTCAGAAACGGCTAACTTGGGGAAGTTAATATGCATTCCAATGGCACTAAGA
GCTGCGATGGAGCTAAATGTGTTCCAACCTTATCTCAAAGTTCGGAACCTGACGCAAAAGTTTCGGCTTCTGAAATT
GCCTCTAAAAATGCCAAACGCGAAGAATAAATCTGAAGCAGCTATGTATTTGGATAGAATTCTTCGACTGCTCGGG
GCAAGTTCTATTCTTTCTGTTTCTACTACAAAAAATCAATCAACAGAGGAGGAGATGATGTAGTAGTACATGAG
AAGCTTTATGGGTAAACAAATTCGTTCGTGTTGTTTGGTCCCTCGACAAGAAGACGGGGTGTCTATTAGTCGAAGAA
TTGCTATTACATCTGACAAGGTTGTTGTTGGATAGTTTTTTTCAAACCTGAAATGTGTGGTGGGAAGAAAAAGACAGT
GTGCCATTTGAGGTTGCTCATGGTGCTAAAAATCTTTGAGTATGCTGCTACAGAACCAAGAATGAATCAAGTATTT
AACGATGGAATGGCAGTTTTCTCTATTGTTGTTTTTGGAGGCTGTTTTTAGAGTTTACGATGGATTTCTTGATATG
AAAGAATTGTTAGATGTTGGTGGTGGTATTGGTACTTCGGTTAGTAAGATTGTTGCTAAATACCCTTTGATTTCGC
GGTGTCAACTTCGACTTGCCTCATGTTATTTCTGTTGCCCTCAATACCCAGGTATACCTTCTTCTTCTTTTTTTC
TGAAAAGAACGGGTTTCAATTTTTTACAGAATTTTTTTTCTCATTCGATACTCAAGCAACTCTATTAAAGTATACT
GTGTAATAATGCATGCAGGTGTAGAGCATGTTGCAGGAGATATGTTTCGAGGAAGTCCCAAAGGGTCAAAACATGT
TGCTAAAAGTAAGCTAACCATACTCAATTTTCTTAATAATTAGGAAAAATTGCAAAAACCGTCACAATATTATAAA
GGCATCTGAAGTGCCATCACTCAGATACCGATGCTATGTACTCTATACATTGACAAAAATCCATGGTATCAAGTC
TCAACCTGCCGGTTATAATAATTTTTTTTCAGGCTTCTTTTAAAGAAATATTTTGAATGGTAAAAATCATCATT
ATATTGGAGAAAAGTGCAGATCTTGCTACATTAATAATTTATAATATAATAAAACATTTGTTTATGGTTGTTTGAA
AAAAAAATCTCATTGTTAATGCATCTTTCTAAGTTAATGGTGATTAATGGTGAATAATATGATATCTTATTACC
GTCTTGACACTTTTTTTTTTGTCTGAGACAAAATATTTCCAACCTTTCTATATTAATAAAATCAGAAATATTTCA
TTTATATGAATATTAATAAATAAGAAGGTGCATGAGTAATATTCCAAATTTCTTAAAGCGTTTTTTATAGCAGTACG
GCGTTTTCTCAAATCTTATTAACCCATAATTAAGGGTTTCCGTAAATTAATTTGAGGGATATCAAAACAAAAAC
AAAAAATAGGGTTATTTTGCAGTAAATCAATAACCCCTTATCATATGAAAAGGATAACTTAGTCTACCCCAATT
TGGAGAGATATGGGCAATTATTGTATTACTAGTTCGTTTGGAGCATTGATAATATTTTTTCATTAGATTTATACTCA
ATAAAATATATGAACTATATTGATAAAGATTAATAATGCAGTGGGTACTGCACGATTGGGGTGATGAACGATGTG
TGAAGCTGTAAAGAATTGTTGGAACCTATTACCTGTGGGTGGAAAAAGTTTGGATAATCGAGTTTGTTCTCCCGA
ATGAACTTGGTAACAATGCTGAATCATTCAATGCGTTGATTCCCGATTACTCCTGATGGCTCTGAATCCAGGCG
GTAAAGAGCGAACGATTTCCGAATACGATGATTTAGGCAAAGCAGCTGGATTCATAAAAACTATACCTATCCCTA
TCTCCAATGGTCTTCATGTCATTGAGTTTCACAAATGAATGGTTATTGAGTGCTTTGGTAATTAACTACCAAGA
TAACTACATCCATTTTCATGCATTTGCTTTTTTTTTTTCTTTTTTTTTCTTTTTTTTTCTTTTTGTTTTGTATTCCA
GGTGTGAAGTGTAGTGTGTTGAGTGGACAAAAGTAAGTAATCGTATTTTGTGTT

Figure 3b

4/14

Start	End	Feature
1	74	5' untranslated region
75	851	exon 1
852	938	intron
939	1232	exon 2
1233	1440	3' untranslated region

GAATCAGAAACTTTCTTCTAAAATCTTTCAATACCAGTACTGTTAGTTTCCGATAAGAGCCACACTAA
TCCATTATGGAAATTCATTTAGAAAGCCAAGAACAAGAAATGAAATATCAATCTCAAATCTGGAACCA
AATATGTGGCACTGTTGATACCTCTGTTCTGAGATGTGCAATTCAATTAGGTATATTTGATGCCATTC
ATAACTCTGGCAAACCAATGATTACCTTAACCGAATTATCAAGCATTGTTTCATCACCCCTCTTCATCT
TCAATCGAACCCTGCAACTTGTATAGATTAGTGAGATACTTATCCCAAATGGATCTCATTAGTATCGG
AGAATGTTTGAATGAAGCAACTGTTTCATTAACAGGCACATCCAAGTTACTACTTAGAAACCAAGAAA
AGAGTTTAATTGATTGGGTATTGGCAATTTCTTGCGAAATGATGGTTGTTGTTTGGCACGAACCTAAGT
AGCTCTGTTTCAACTCCTGCGGATGAGCCTCCAATCTTCCAGAAGGTTTCATGGTAAAAATGCTTTAGA
ATTAGCAGGGGAATTTCCAGAATGGAATGATCTGATCAACAATGCTATGACTAGTGATACTAGTGTA
CTAAGCCAGCGCTAATACAAGGATGTGGCAAAATCCTGAACGGAGTTACATCGTTAATTGATGTCGGT
GGTGGTCACGGTGCCACTATGGCCTACATAGTTGAAGCTTTTCCTCACATAAAAGGTGCGGTAATCGA
TTTACCACATGTTGTTGAAGCCGCTCCGGAGCGTCCAGGTGTTGAGTTCATCAGCGGTGATATATTCA
AGTCCATTTCTAACGCTGATGCTGTGTTGTTGAAGGTATGTAAAGAGTAGCTAACCTTAGTGCGTCTA
ATTTATTCCACAAATTTTTCTGATGCATTTTATTCTTATTTTTGGTTTTTGCAGTATGTCCTGCACAA
TTGGGAAGATACGGAATGTGTGAATTTACTGAAGAGATGTAAGGAAGCAGTTCGGGCAGACAAAGGAA
AAGTGATCATAATGGATTTAGTAATAGACGACGATGATAACAGTATTTTAACGCAGGCAAAGTTGAGC
CTTGATCTCACTGTGATGAACCATGGAGGAGGTAGAGAAAGGACTAAAGAAGATTGGAGAAATCTAAT
TGAGATGTCTGGATTTAGTAGGCATGAAATAATTCCAATATCTGCCATGCCATCAATTATTGTTGCTT
ATCCTTAGTTAAGTCACCCGCATGTTTACTTGAACGGGAATAAGTTGGGGGCGTGTTGAATCTGTAA
CATCGCAATTGTGCCTTTACTTTATGCATTCTCATTCCGGTAGAAACTGTTTGGGGCATTCGGATTCT
GCTGAGCCCTTTTATGTATGTTTGTGTTGGTTGGTTGGTTTTCAAGTAACTGAAGTTTCTTCTCTG
TTTTCAAGGCAT

Figure 3c**5/14**

Start	End	Feature
1	120	5' untranslated region
121	846	exon 1
847	994	intron
995	1288	exon 2
1289	1436	3' untranslated region

AAGTTGCAGGTAGGGTTATGAGCAAGCTCAATTATCTCTCCTATAAAAAGCTAACATTAGAAAACTAA
TAAGCACACAAACCGTAAAAGTTCTGAAGATAGACAAAACAAGAGAAAAAAGATGGAAGTAGTAAGTA
AGATTGATCAAGAAAACCAAGCAAAAATTTGGAAACAAATTTTTGGTTTTGCAGAATCATTAGTTCTA
AAATGTGCAGTTCAGTTAGAGATTGCTGAAACACTTCATAATAATGTAAAACCCATGTCTTTATCCGA
GTTAGCATCTAAACTTCCGGCTCAACCCGTTAATGAAGACCGTTTGTACCGAATTCTGCATTTCTTAG
TTCACATGAAACTCTTCAACAAAGATGCTACCACACAGAAATATTCATTAGCTCCACCAGCAAAGTAT
TTGCTAAAAGGCTGGGAAAAATCAATGGTTCCCTTCAATATTAAGCGTGACTGATAAAGATTTTACAGC
TCCATGGAATCATCTTGGGGACGGTTTAAACCGGTAACGTGAACGCTTTTGAGAAAGCGTTAGGAAAGG
GCATTCGGGTTTATATGAGAGAAAATCCTGAAAAAGATCAATTGTTTAAATGAAGGAATGGCTTGTGAT
ACTAGATTATTTGCTTCAGCATTGGTTAACGAGTGCAAAAGTATTTTCAGTGACGGGATCAATACACT
TGCCGGTGTTGGCCGTGGTACTGGTACTGCAGTGAAAGCCATATCCAAAGCTTTTCCGGATATTAAGT
GCACAATCCATGATCTTCCCTGAAGTTACCAGTAAAAATAGTAAATTTCCAAGAGATGTTTTTAAGTCC
GTTCCTAGTGCAGACGCCATCTTTATGAAGGTAACCTCTAAGAAATTTTGTTTTAGAATATTCGTTGC
AACTCTAATTGACAACATTCATAAAAAATATGTTAATGGTCTTAATTTATTAATTCTAGTAGAGTTAC
TTAAATGATATACAAAATTCAAAATCATATAACATTTGCAGAGCATTCTTCACGAATGGAACGATGA
GGAATGTATTCAAATCTTGAAACGATGCAAAGAAGCAATACCAAAGGGGGCAAAGTTATCATTGCGG
ATGTCGTAATAGACATGGACTCGACTCATCCGTATTCAAAATCTAGACTCGCAATGGATTTGGCTATG
ATGCTCCACACGGGTGGAAAAGAGAGAACTGAAGAAGATTGGAAAAAAGTTATTGATGCTGCAGGTTT
TGCTAGCTGTAAAATTACTAACTATCTGCTCTCCAGTCTGTTATTGAGGCTTACCCTCATTGAGGAT
AATTTTTATCCTTCTGTTTTCCCTTTGGTTAATTGTTGCCTTCTCTTTGGATCATGGTTGCGTTTATA
ATAAATGCAGCGTTTCTTTCCCTGGCGGTAAGTGCAAGAAAGAAAAAGCTTCCAGAACTTCCTTGAGT
ATGCCTGG

Figure 4a

6/14

MATNGEIFNTYGHNNQSATVTKITASNESSNGVCYLSETANLGKLCIPMALRAAMELNVFQLISKFGTDAKVSA
SEIASKMPNAKNNPEAAMYLDRI LRLLGASSILSVSTTKKSINRGDDVVVHEKLYGLTNSSCCLVPRQEDGVSL
VEELLFTSDKVVDVSFFKLKCVVEEKDSVPFEVAHGAKIFEYAATEPRMNQVFNDGMAVFSIVVFEAVFRVYDGF
LDMKELLDVGGGIGTSSVSKIVAKYPLIRGVNFDLPHVISVAPQYPGVEHVAGDMFEEV PKGQNMLLKWVLHDWGD
ERCVKLLKNCWNSLPVGGKVLIIEFVLPNELGNNAESFNALIPDLLLMALNPGGKERTISEYDDLGAAGFIKTI
PIPI SNGLHVIEFHK.

Figure 4b

MEIHLESQEQEMKYQSQIWNQICGTVDTSVLRCAIQLGIFDAIHNSGKPMITLTELSSIVSSPSSSSSIEPCNLYR
LVRYLSQMDLISIGECLNEATVSLTGTSKLLLRNQEKS LIDWVLAISCEMMVVVWHELSSSVSTPADEPPIFQKV
HGKNALELAGEFPEWNDLINNAMTSDTSVTKPALIQGCGKI LINGVTS LIDVGGGHGATMAYIVEAFPHIKGAVID
LPHVVEAAPERPGVEFISGDIFKSISNADAVLLKYVLHNWEDTECVNLLKRCKEAVPADKGV IIMDLVIDDDDN
SILTQAKLSLDLTVMNHGGGRERTKEDWRNLIEMSGFSRHEIIPISAMPSIIVAYP

Figure 4c

MEVVS KIDQENQAKIWKQIFGFAESLVLKCAVQLEIAETLHNNVKPMSLSELASKLP AQPVNEDRLYRILHFLVH
MKLFNKDATTQKYS LAPPAKYLLKGWEKSMVPSILSVTDKDF TAPWNHLGDGLTGNCNAFEKALGKGIRVYMREN
PEKDQLFNEGMACDTRLFASALVNECKSIFSDGINTLAGVGRGTGTAVKAISKAFPDIKCTI HDLPEVTSKNSKI
PRDVFKSVPSADAI FMKSILHEWNDEECIQILKRCKEAI PKGGKV I IADVVIDMDSTHPYKSR LAMDLAMMLHT
GGKERT EEDWKKLIDAAGFASCKITKLSALQSVIEAYPH

Figure 5

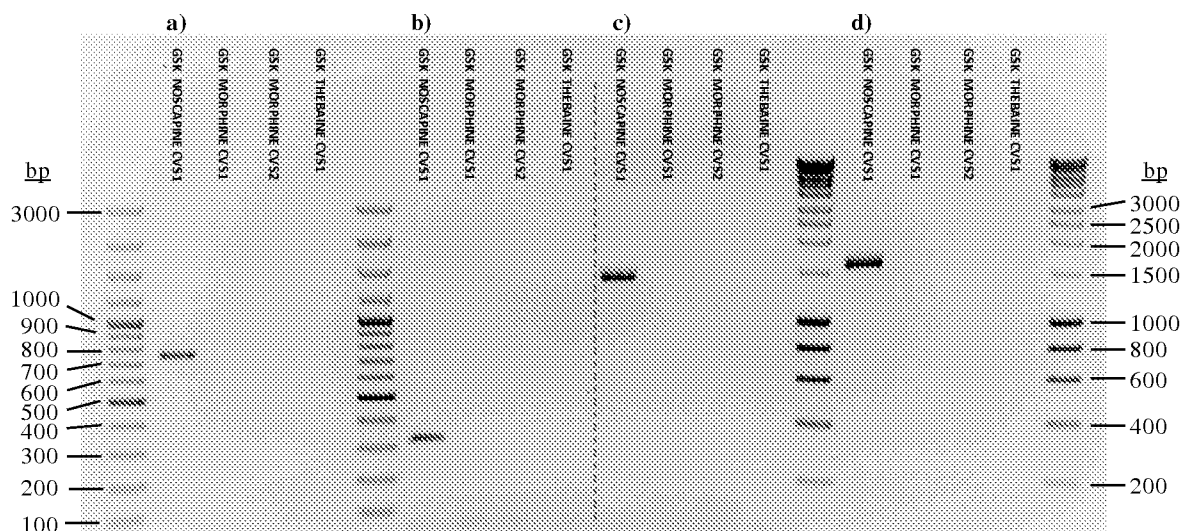


Figure 6a

7/14

F2: GSK THEBAINE CVS1 x GSK NOSCAPINE CVS				
SUM	Noscapine +		Noscapine -	
	61		215	
276	<i>PSMT1</i> +	<i>PSMT1</i> -	<i>PSMT1</i> +	<i>PSMT1</i> -
	61	0	131	84
SUM	<i>PSMT1</i> +		<i>PSMT1</i> -	
	192		84	
276	Noscapine +	Noscapine -	Noscapine +	Noscapine -
	61	131	0	84

Figure 6b

F2: GSK THEBAINE CVS1 x GSK NOSCAPINE CVS				
SUM	Noscapine +		Noscapine -	
	61		215	
276	<i>PSMT2</i> +	<i>PSMT2</i> -	<i>PSMT2</i> +	<i>PSMT2</i> -
	59*	0	127**	84
SUM	<i>PSMT2</i> +		<i>PSMT2</i> -	
	186***		84	
276	Noscapine +	Noscapine -	Noscapine +	Noscapine -
	61	131	0	84

* two genotyping assays failed; **four genotyping assays failed; ***in total the *PSMT2* genotyping assay failed on DNA samples from six F2 plants that were positive for *PSMT1* and *PSMT3* (Figure 6a and 6c)

Figure 6c

F2: GSK THEBAINE CVS1 x GSK NOSCAPINE CVS				
SUM	Noscapine +		Noscapine -	
	61		215	
276	<i>PSMT3</i> +	<i>PSMT3</i> -	<i>PSMT3</i> +	<i>PSMT3</i> -
	61	0	131	84
SUM	<i>PSMT3</i> +		<i>PSMT3</i> -	
	192		84	
276	Noscapine +	Noscapine -	Noscapine +	Noscapine -
	61	131	0	84

Figure 7

8/14

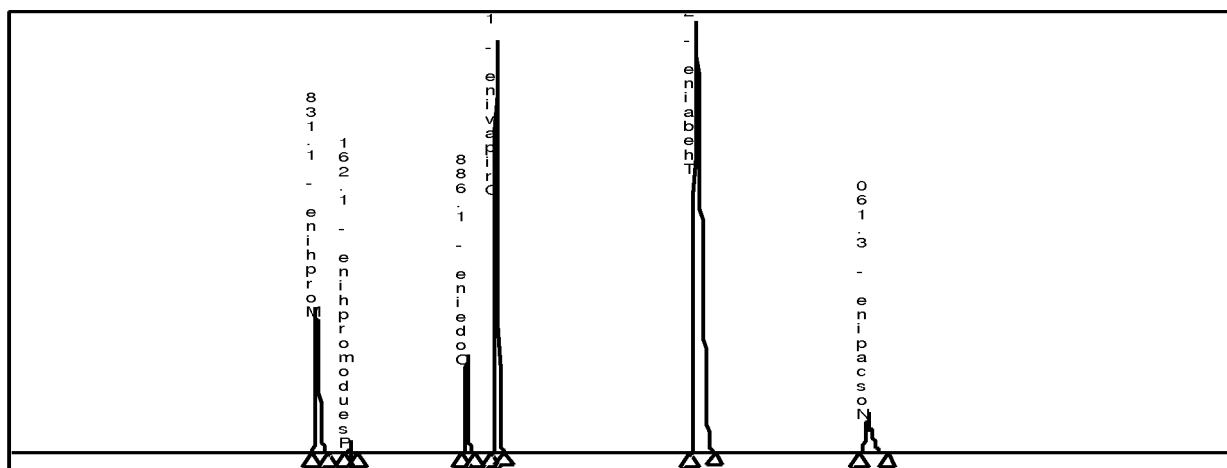


Figure 8

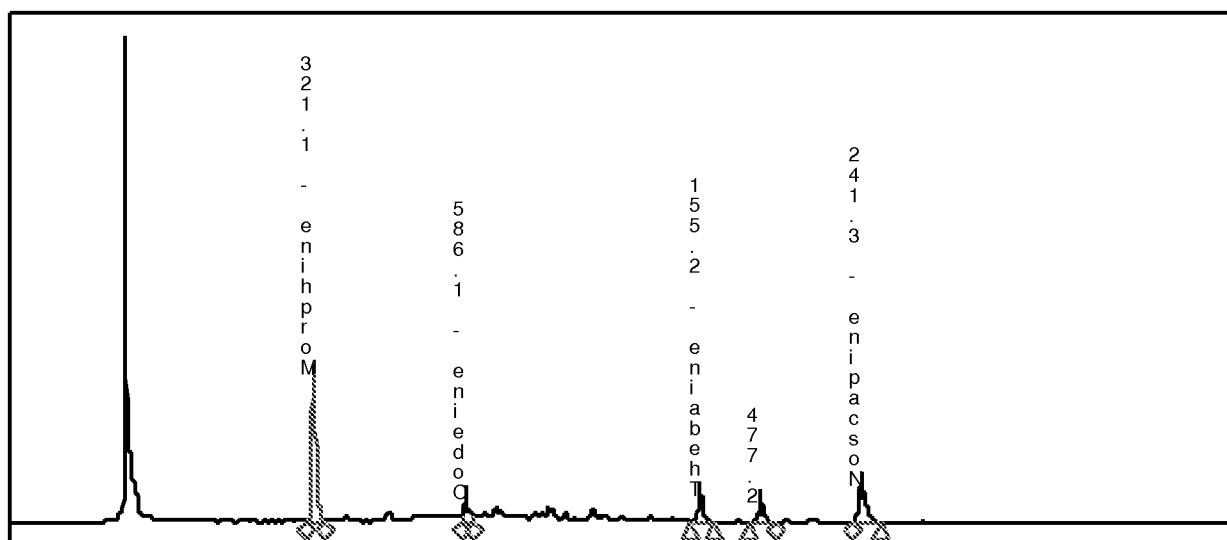


Figure 9

9/14

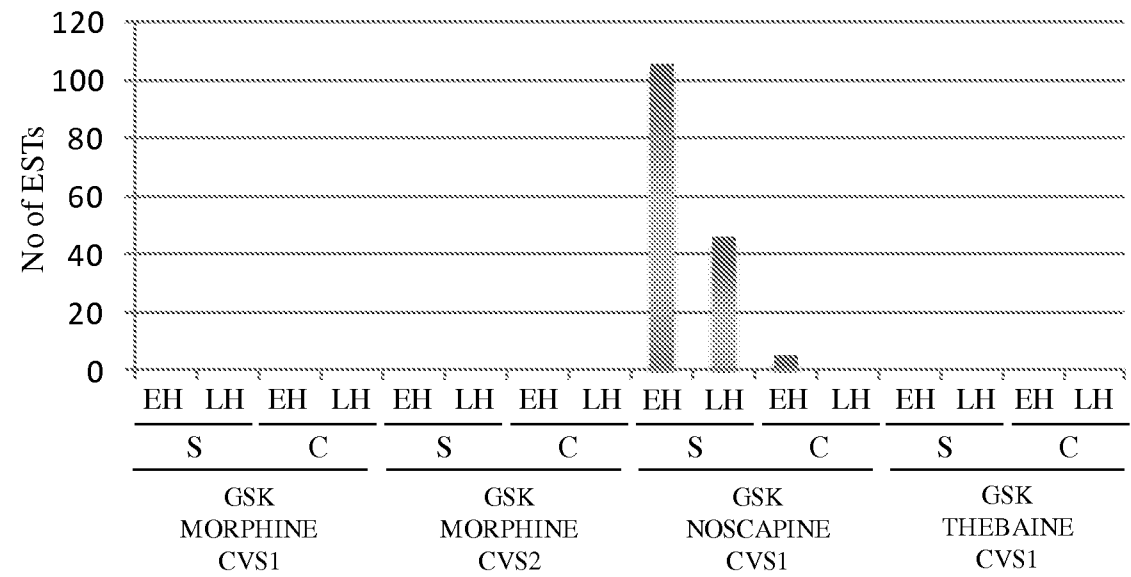


Figure 10

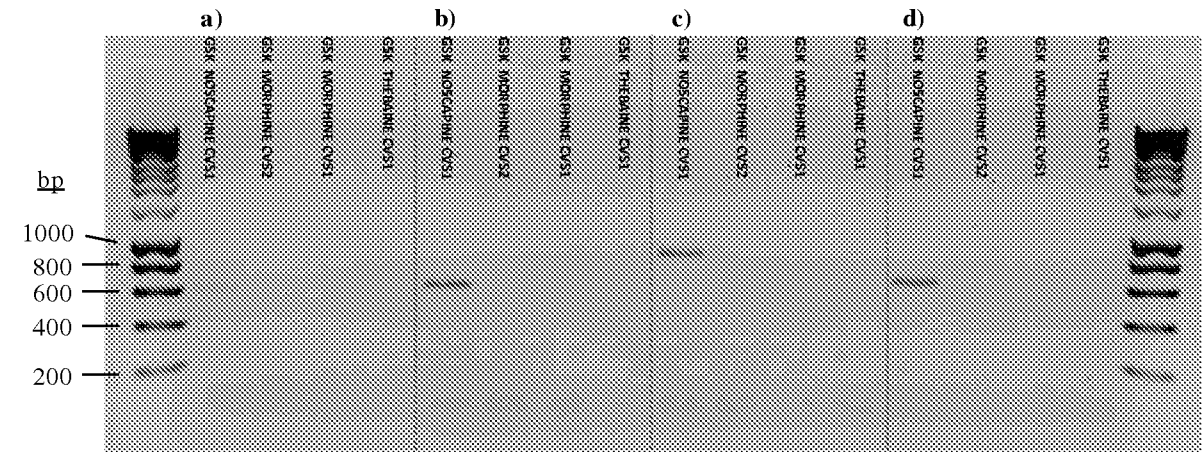


Figure 11

10/14

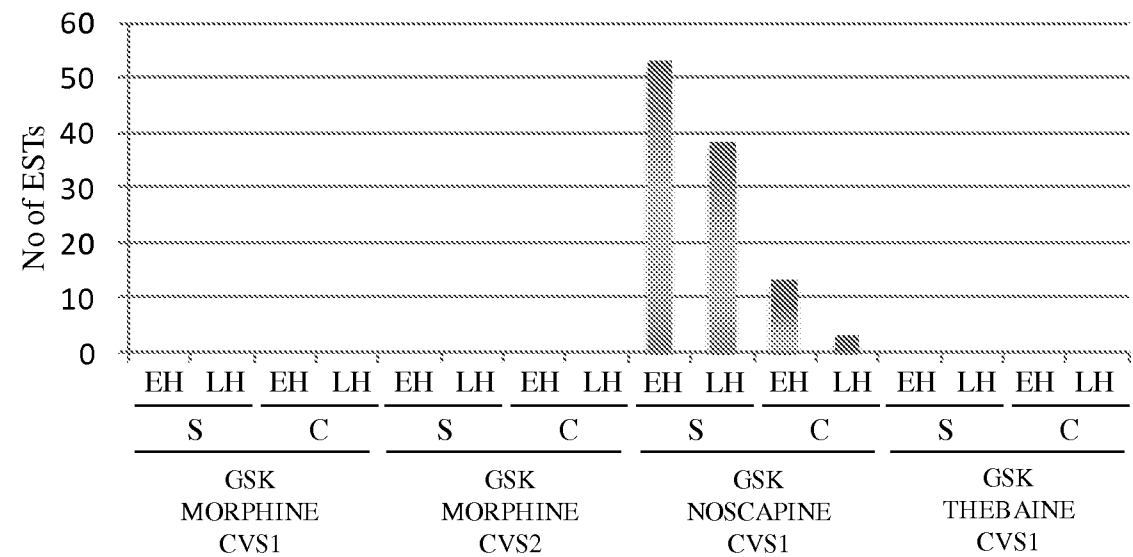


Figure 12

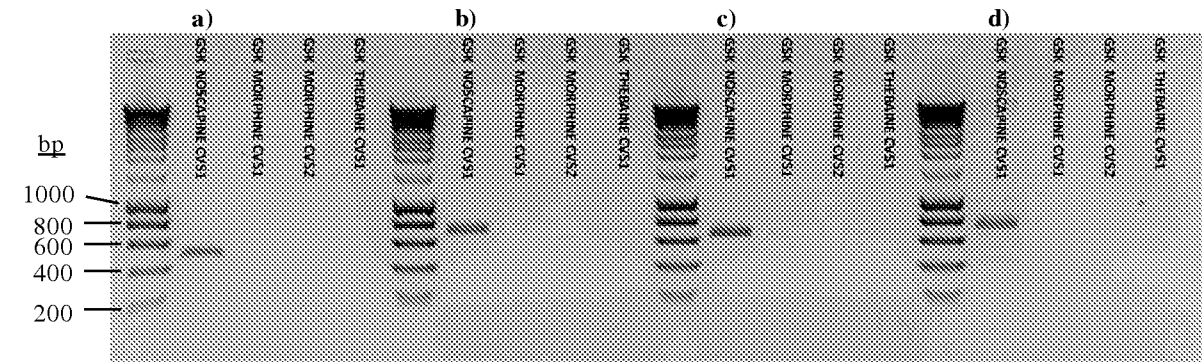


Figure 13

GAGGTGTTTCATTGCCATGTCAAAGGCATTAAACTTCATAAACCCAGATGAGCTTTCGATGCAGTGCATTTTGATA
GCTTTGAACCGTTTTCCTTCAGGAAAAGCATGGTTCCAAGATGGCCTTTTGTAGATGGTAATCCTCCCGAGAGACTT
TGCAAGCCGGTCGTGGATCATATAGAGTCACTTGGCGGTGAAGTCCGTCTCAATTCCAGGATTAAAAAGATTGAG
CTTAAAAAAGATGGTACTGTGAAACGTCTAATGCTCACCAACGGTGATGCAATAGAAGGAGATGCTTATGTCATT
GCAACCCAGTGGACATCCTAAAGCTGCTTATACCCGAGGAGTGAAAGAAGTTGGGTACTTTAAAAAGATTGGAT
AAATTAGTTGGAGTTCTGTGATTAACGTCCATATATGGTTTGACAGGAAATTGAAAAATACATATGATCATCTT
CTCTTCAGCAGAAGTCCCCTCTTAAGCGTATACGCTGACATGTCAGTGACATGCAAGGAATATTATGACCCAAAC
AAATCCATGCTTGAGTTGGTATTTGCACCCGCTGAGGAATGGATCTCGCGCAGTGACTCTGAAATTATTGAAGCT
ACTATGCAGGAGCTTGCGAAAC

Figure 14

TGGTCATAATCATCAATCAGCCACAGTCACTAAAATCACTGCTTCTAATGAAAGCAGCAATGGTGTCTGTTATCT
TTCAGAAACGGCTAACTTGGGGAAGTTAATATGCATTCCAATGGCACTAAGAGCTGCGATGGAGCTAAATGTGTT
CCAATTATCTCAAAGTTCGGAAGTACGCAAAAAGTTTCGGCTTCTGAAATTGCCTCTAAAAATGCCAAACGCGAA
GAATAATCCTGAAGCAGCTATGTATTTGGATAGAATTCTTCGACTGCTCGGGGCAAGTTCTATTCTTTCTGTTTC
TACTACAAAAAATCAATCAACAGAGGAGGAGATGATGTAGTAGTACATG

Figure 15

11/14

GTGTAAC TAAGCCAGCGCTAATACAAGGATGTGGCAAAATCCTGAACGGAGTTACATCGTTAATTGATGTCGGTG
GTGGTCACGGTGCCACTATGGCCTACATAGTTGAAGCTTTTCCTCACATAAAAGGTGCGGTAATCGATTACCAC
ATGTTGTTGAAGCCGCTCCGGAGCGTCCAGGTGTTGAGTTCATCAGCGGTGATATATTCAAGT

Figure 16

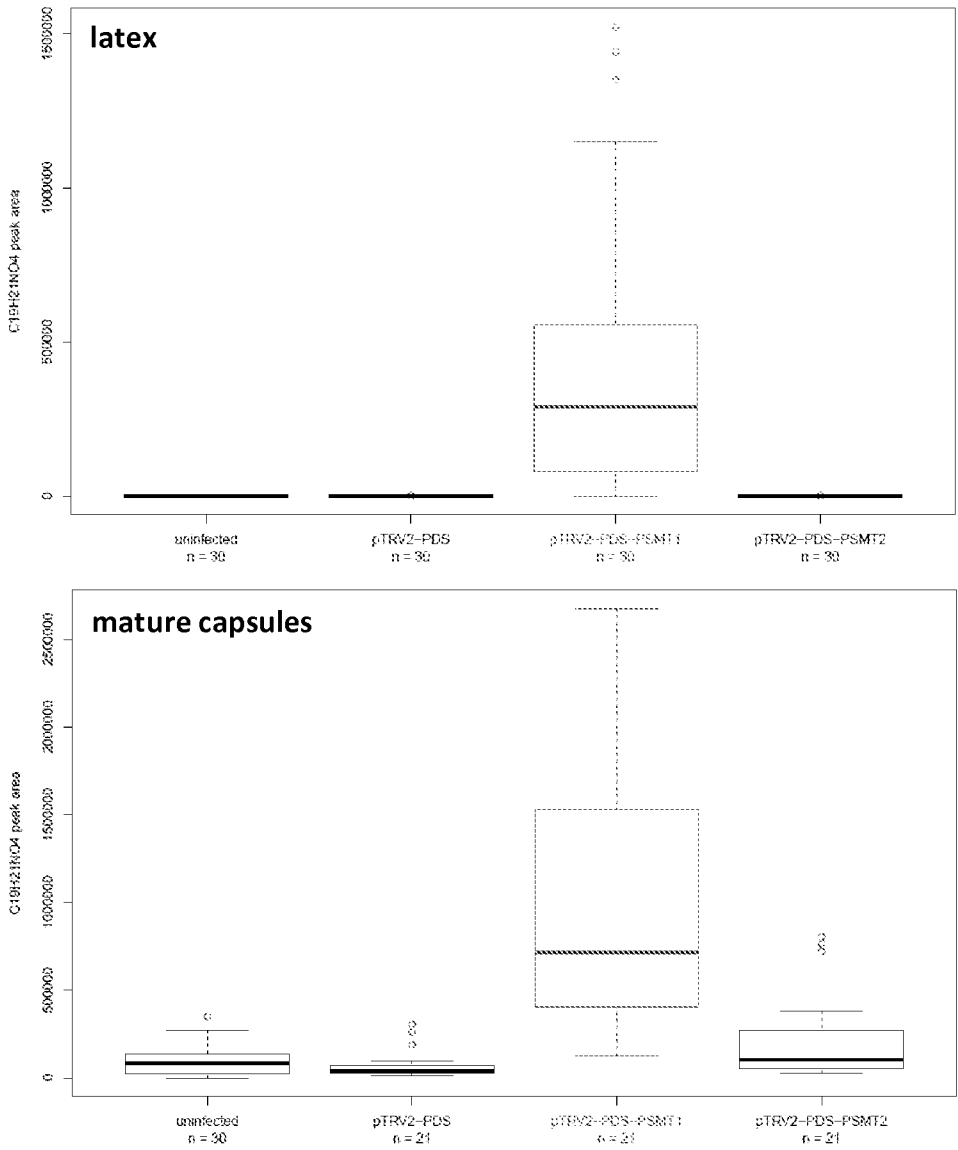
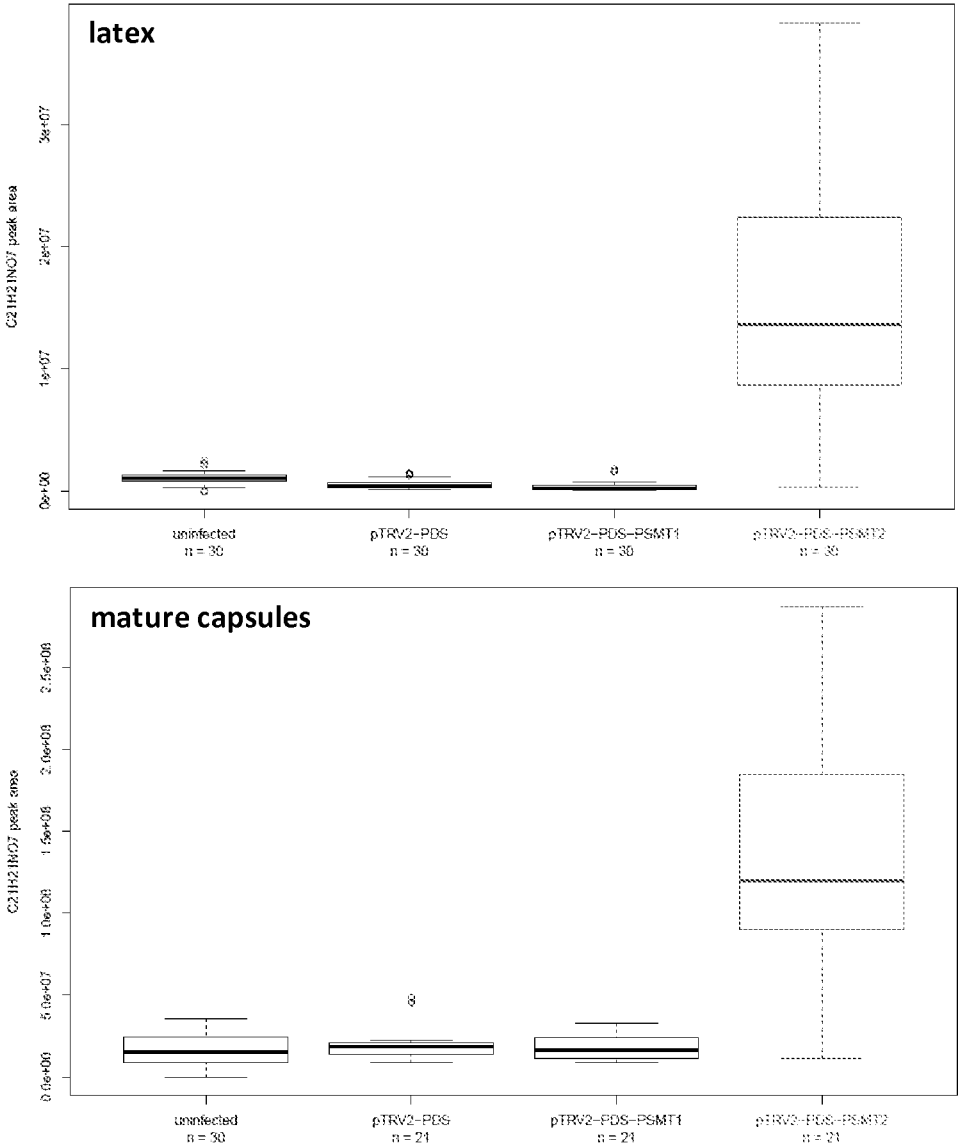


Figure 17



13/14

Table 1: Sequences of forward and reverse primers used to amplify the methyltransferase genes from genomic DNA

Methyltransferase gene	Primer name	Oligonucleotide sequences (5'– to 3'-)
<i>PSMT1</i>	PSMT1_F2	CTCTAAATGCCAAACGCG
	PSMT1_F3	CGTGTGTTGTTGGTCCCTCG
	PSMT1_F4	GATTCCCGATTACTCCTGATGG
	PSMT1_F5	TGCCTCATGTTATTTCTGTTGCC
	PSMT1_F6	CACACCAAACCTTGATCATTGTC
	PSMT1_R1	GCACACTGTCTTTTCTTCCACC
	PSMT1_R2	GACCCTTTGGGACTTCCTCG
	PSMT1_R3	AAATCGTTCGCTCTTACCGC
	PSMT1_R4	AACACAAAATACGATTACTTACTTTTGTCC
<i>PSMT2</i>	PSMT1_R5	GCATGAAATGGATGTAGTTATCTTGG
	PSMT2_F1	ATTGTTGATATTGAATCAGAACTTTC
	PSMT2_F2	TCAATACCAGTACTGTAGTTTCCG
	PSMT2_F4	GCAACTGTTTCATTAAACAGGCACATCC
	PSMT2_F5	ATTGATGTCGGTGGTGGTCACG
	PSMT2_R3	CCACATCCTTGTATTAGCGCTGGC
	PSMT2_R4	CAGTAAATTCACACATTCGGTATCTTCCC
	PSMT2_R5	ATTCCCGTTCAAGTAAACATGCGG
	PSMT2_R6	ACCGGAATGAGAATGCATAAAGTAAAGG
<i>PSMT3</i>	PSMT3_F1	ATTGTATAGCCAAAGTTGCAGGTAGGG
	PSMT3_F2	AGACCGTTTGTACCGAATTCTGC
	PSMT3_F3	GCTTCAGCATTGGTTAACGAGTGC
	PSMT3_F4	GCAGTGAAAGCCATATCCAAAGC
	PSMT3_R1	AACCGTCCCCAAGATGATTCC
	PSMT3_R2	TCGTTCCATTCTGTGAAGAATGC
	PSMT3_R3	GAGGGTAAGCCTCAATAACAGACTGG
	PSMT3_R5	TTAAGATCACCAGGCATACTCAAGG

Table 2: Primer combinations used to amplify and Sanger-sequence methyltransferases from genomic DNA

Methyltransferase	Primer combination	Annealing temperature [°C]	Extension time [s]	Sequencing primers used for Sanger sequencing of purified PCR product
<i>PSMT1</i>	PSMT1_F2/R2	62	60	PSMT1_F2, PSMT1_R2
	PSMT1_F3/R3	68	60	PSMT1_F3, PSMT1_F5, PSMT1_R3
	PSMT1_F4/R4	66	60	PSMT1_F4, PSMT1_R4
	PSMT1_F5/R5	63.5	60	PSMT1_F5, PSMT1_R3, PSMT1_R5
	PSMT1_F6/R1	62.8	60	PSMT1_F6, PSMT1_R1
<i>PSMT2</i>	PSMT2_F1/R6	60	60	PSMT2_F1, PSMT2_F2, PSMT2_F4, PSMT2_F5, PSMT2_R1, PSMT2_R4, PSMT2_R6
<i>PSMT3</i>	PSMT3_F1/R5	68.5	60	PSMT3_F2, PSMT3_F4, PSMT3_F5, PSMT3_R1, PSMT3_R2, PSMT3_R3

Table 3: Primer combinations used to amplify PS Methyltransferases from genomic DNA

Methyltransferase	Primer combination	Annealing temperature [°C]	Extension time [s]	Expected fragment size [bp]	Figure
<i>PSMT1</i>	PSMT1_F1/R2	62	60	739	Figure 5a
	PSMT1_F3/R3	66	60	1568	Figure 5b
	PSMT1_F4/R4	68	60	318	Figure 5c
	PSMT1_F5/R5	63.5	60	1354	Figure 5d
	PSMT1_F6/R1	62.8	60	604	Figure 5e
<i>PSMT2</i>	PSMT2_F2/R3	60.3	60	613	Figure 10a
	PSMT2_F4/R4	60.9	60	627	Figure 10b
	PSMT2_F4/R5	61.7	60	909	Figure 10c
	PSMT2_F5/R6	62.8	60	664	Figure 10d
<i>PSMT3</i>	PSMT3_F1/R1	68.5	50	497	Figure 12a
	PSMT3_F2/R2	68.5	50	564	Figure 12b
	PSMT3_F3/R3	68.5	50	511	Figure 12c
	PSMT3_F4/R5	68.5	50	589	Figure 12d

Table 4: Primers used to amplify sequences selected for virus induced gene silencing

Target gene to be silenced	Primer name	Oligonucleotide sequences (5'- to 3'-) (in capitals: gene-specific sequence; in lower case: added sequence; underlined: restriction sites)
<i>PS PHYTOENE DESATURASE</i>	ps_pds_F	GAGGTGTTTCATTGCCATGTCAA
	ps_pds_R4	GTTTCGCAAGCTCCTGCATAGT
<i>PSMT1</i>	VIGS_PSMT1_F	aaactcgaagctTGGTCATAATCATCAATCAG
	VIGS_PSMT1_R	aaaggtagcCATGTACTACTACATCATCTCC
<i>PSMT2</i>	VIGS_PSMT2_F	aaactcgaagcttGTGTAAGTAAGCCAGCGC
	VIGS_PSMT2_R	aaaggtagcACTTGAATATATCACCGC