



US010385363B2

(12) **United States Patent**  
**Zhang et al.**

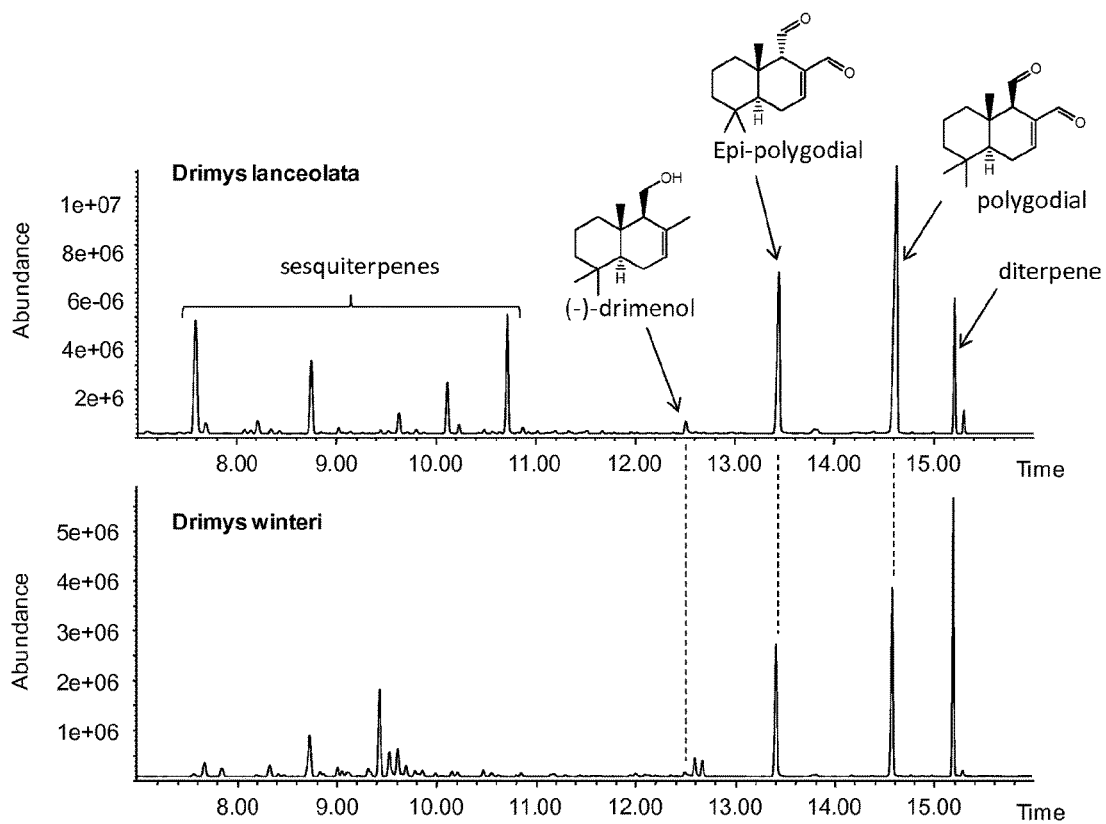
(10) **Patent No.:** **US 10,385,363 B2**  
(45) **Date of Patent:** **Aug. 20, 2019**

- (54) **DRIMENOL SYNTHASES II**
- (71) Applicant: **Firmenich SA**, Geneva (CH)
- (72) Inventors: **Yu-Hua Zhang**, Shanghai (CN); **Michel Schalk**, Geneva (CH); **Olivier Haefliger**, Shanghai (CN); **Xiu-Feng He**, Shanghai (CN); **Fabienne Deguerry**, Geneva (CH)
- (73) Assignee: **Firmenich SA**, Geneva (CH)
- (\* ) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 156 days.
- (21) Appl. No.: **15/330,814**
- (22) PCT Filed: **May 6, 2015**
- (86) PCT No.: **PCT/EP2015/059988**  
§ 371 (c)(1),  
(2) Date: **Nov. 7, 2016**
- (87) PCT Pub. No.: **WO2015/169871**  
PCT Pub. Date: **Nov. 12, 2015**
- (65) **Prior Publication Data**  
US 2018/0251797 A1 Sep. 6, 2018
- (30) **Foreign Application Priority Data**  
May 6, 2014 (WO) ..... PCT/CN2014/076890
- (51) **Int. Cl.**  
**C12N 9/02** (2006.01)  
**C12N 9/16** (2006.01)  
**C12P 7/02** (2006.01)  
**C12P 7/04** (2006.01)
- (52) **U.S. Cl.**  
CPC ..... **C12P 7/02** (2013.01); **C12N 9/0004** (2013.01); **C12N 9/16** (2013.01); **C12P 7/04** (2013.01); **C12Y 301/07007** (2013.01)
- (58) **Field of Classification Search**  
None  
See application file for complete search history.

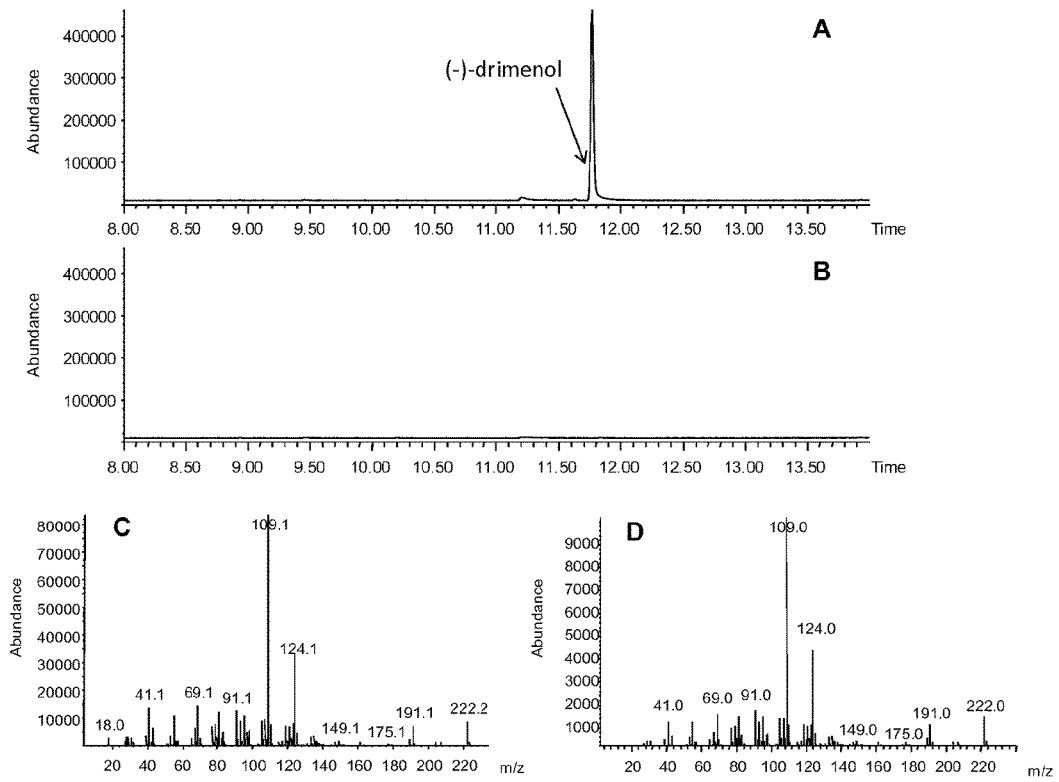
- (56) **References Cited**
- FOREIGN PATENT DOCUMENTS
- WO WO2012058636 5/2012
- WO WO2013058655 4/2013
- OTHER PUBLICATIONS
- Whisstock et al. Quaterly Reviews of Biophysics, 2003, "Prediction of protein function from protein sequence and structure", 36(3):307-340.\*
- Witkowski et al. Conversion of a beta-ketoacyl synthase to a malonyl decarboxylase by replacement of the active-site cysteine with glutamine, Biochemistry, Sep. 7, 1999;38(36):11643-50.\*
- International Search Report and Written Opinion, application PCT/EP2015/059988 dated Nov. 26, 2015.
- Altschul, Stephen F., et al; J. Mol. Biol. (1990) 215, 403-410.
- Schalk, Michel, et al; Journal of the American Chemical Society (2012), 134, 18900-18903.
- Tatiana A. Tatusova, et al; FEMS Microbiology Letters 174 (1999) 247-250.
- Munoz-Concha, et al, Biochemical Systematics and Ecology, vol. 35, No. 7, 2007 p. 434-438.
- XP55211211, Calgary, Alberta; Retrieved from the Internet, URL: [http://theses.ucalgary.ca/bitstream/11023/129/2/ucalgary\\_2012\\_pyle\\_bryan.pdf](http://theses.ucalgary.ca/bitstream/11023/129/2/ucalgary_2012_pyle_bryan.pdf).
- \* cited by examiner
- Primary Examiner* — Iqbal H Chowdhury
- (74) *Attorney, Agent, or Firm* — Armstrong Teasdale LLP

- (57) **ABSTRACT**
- The present invention relates to a method of producing drimenol and/or drimenol derivatives by contacting at least one polypeptide with farnesyl diphosphate. The method may be performed in vitro or in vivo. The present invention also provides amino acid sequences of polypeptides useful in the method of the invention and nucleic acid encoding the polypeptides of the invention. The method further provides host cells or organisms genetically modified to express the polypeptides of the invention and useful to produce drimenol and/or drimenol derivatives.
- 8 Claims, 7 Drawing Sheets**
- Specification includes a Sequence Listing.**

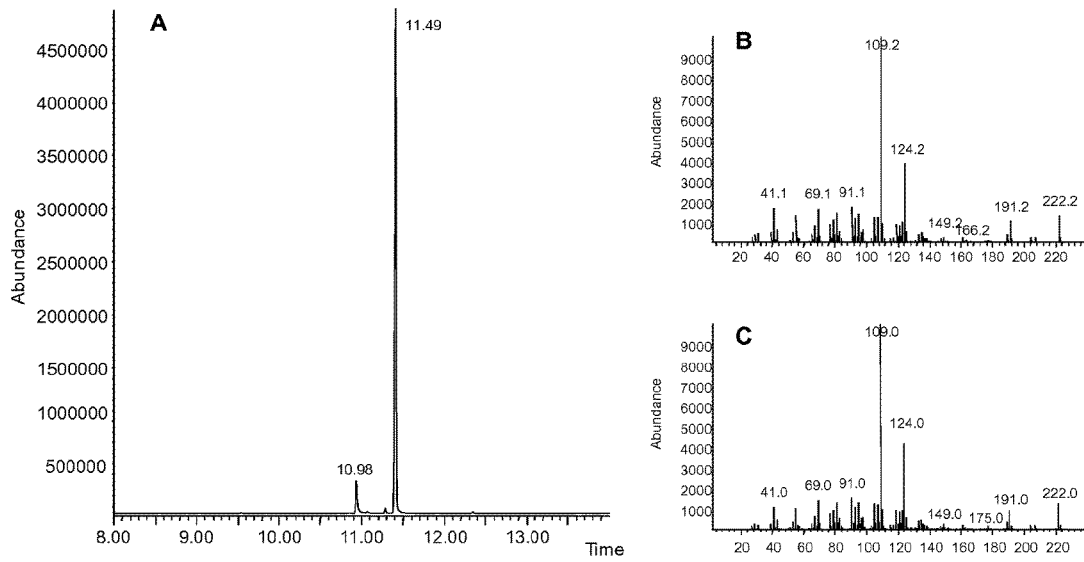
- Figure 1 -



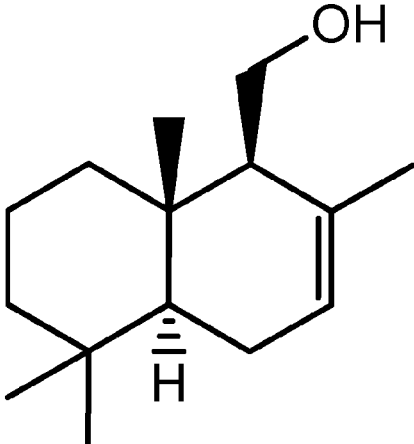
- Figure 2 -



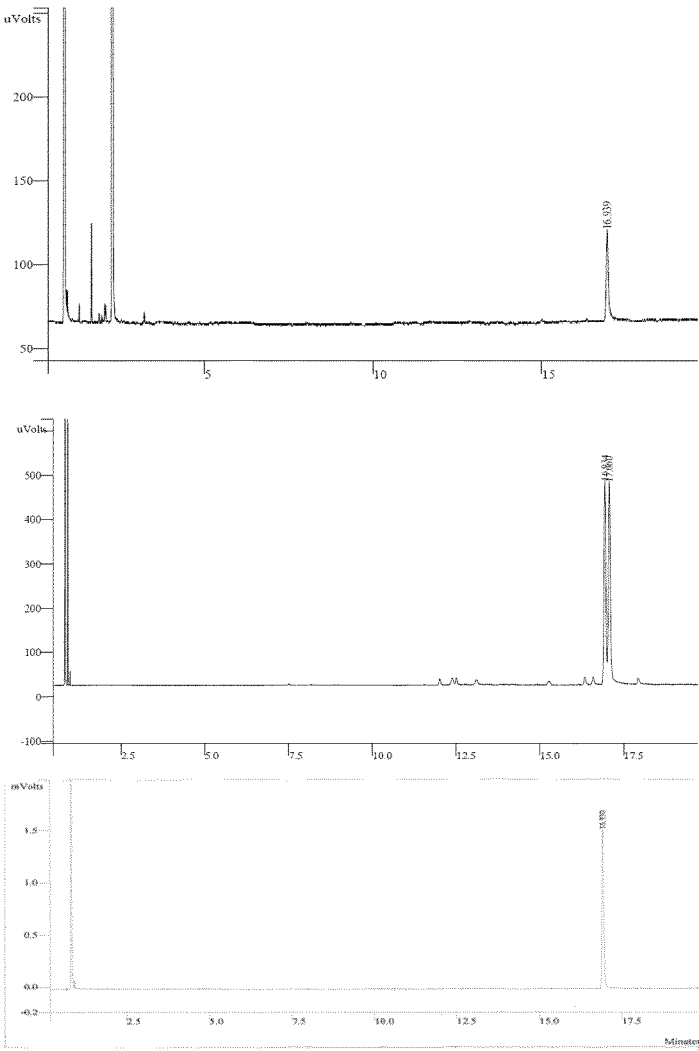
- Figure 3 -



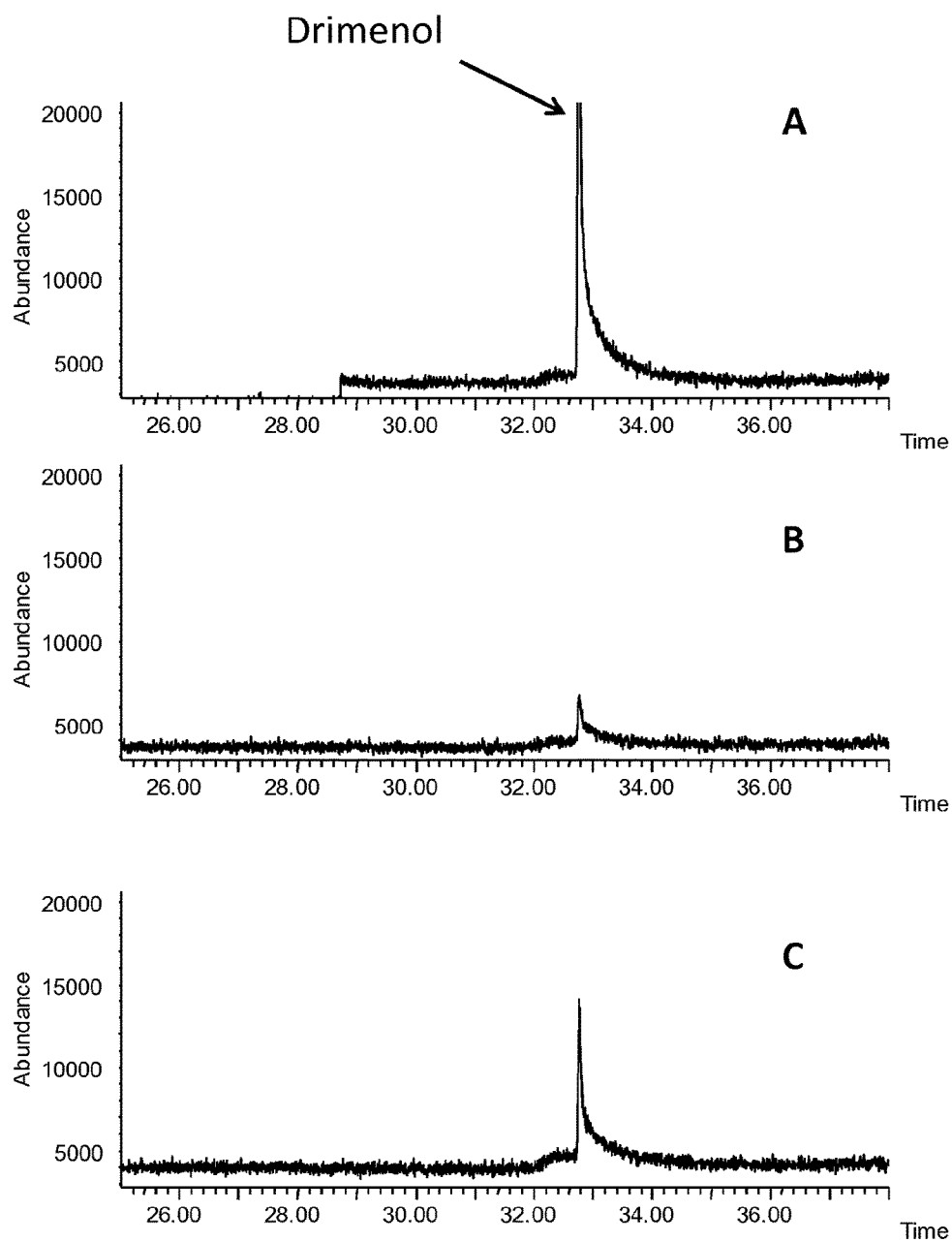
- Figure 4 -



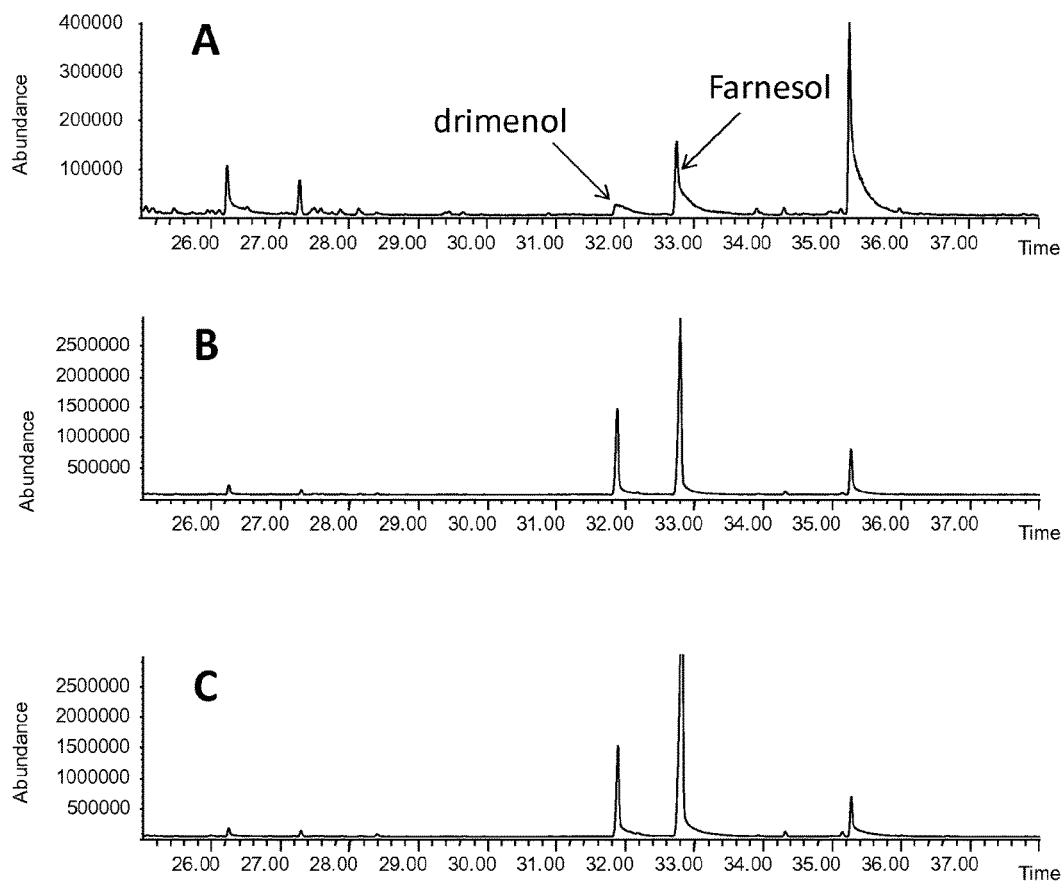
- Figure 5 -



- Figure 6 -



- Figure 7 -



## DRIMENOL SYNTHASES II

## TECHNICAL FIELD

The field relates to methods of producing Drimenol, said method comprising contacting at least one polypeptide with farnesyl pyrophosphate (FPP). In particular, said method may be carried out in vitro or in vivo to produce Drimenol, a very useful compound in the fields of perfumery. Also provided herein is an amino acid sequence of a polypeptide useful in the methods provided herein. A nucleic acid encoding the polypeptide of an embodiment herein and an expression vector containing said nucleic acid are also provided herein. A non-human host organism or a cell transformed to be used in the method of producing Drimenol is further provided herein.

## BACKGROUND

Terpenes are found in most organisms (microorganisms, animals and plants). These compounds are made up of five carbon units called isoprene units and are classified by the number of these units present in their structure. Thus monoterpenes, sesquiterpenes and diterpenes are terpenes containing 10, 15 and 20 carbon atoms respectively. Sesquiterpenes, for example, are widely found in the plant kingdom. Many sesquiterpene molecules are known for their flavor and fragrance properties and their cosmetic, medicinal and antimicrobial effects. Numerous sesquiterpene hydrocarbons and sesquiterpenoids have been identified.

Biosynthetic production of terpenes involves enzymes called terpene synthases. There is virtually an infinity of sesquiterpene synthases present in the plant kingdom, all using the same substrate (farnesyl pyrophosphate, FPP) but having different product profiles. Genes and cDNAs encoding sesquiterpene synthases have been cloned and the corresponding recombinant enzymes characterized.

Currently the main source for Drimenol are plants naturally containing Drimenol and the contents of Drimenol in these natural sources are low. Chemical synthesis approaches have been developed but are still complex and not cost-effective.

## SUMMARY

Provided herein is a method of producing Drimenol comprising:

- i) contacting a acyclic terpene pyrophosphate with a polypeptide having Drimenol synthase activity and having at least, or at least about 70% sequence identity to a sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 11 and SEQ ID NO: 14 to produce the Drimenol; and
- ii) optionally isolating the Drimenol.

Further provided herein is an isolated polypeptide having Drimenol activity comprising an amino acid sequence having at least or at least about 70%, or more identity to amino acid sequence of a sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 11 and SEQ ID NO: 14.

Also provided herein is an isolated nucleic acid molecule encoding a polypeptide having at least, or at least about 70% sequence identity to a sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 11 and SEQ ID NO: 14

## DESCRIPTION OF THE DRAWINGS

FIG. 1. GCMS analysis of the leaves of *Drimys lanceolata* and *Drimys winteri*.

FIG. 2. GCMS analysis of the sesquiterpenes produced by the recombinant DITps589 in in-vitro assays. A. Total ion chromatogram of the sesquiterpene profile of an incubation of the recombinant DITps589 protein with FPP. B. Negative control performed in the same conditions with *E. coli* cells transformed with an empty plasmid. C. Mass spectrum of the peak at 11.76 min. D. Mass spectrum of an authentic standard of (-)-drimenol.

FIG. 3. GCMS analysis of the sesquiterpenes produced in vivo by the recombinant DITps589 in engineered bacteria cells. A. Total ion chromatogram. B. Mass spectrum of the peak at 11.49 min. C. Mass spectrum of an authentic standard of (-)-drimenol. The compound eluting at 10.98 min is farnesol produced by the DITps589 enzyme or resulting from the hydrolysis of excess FPP produced by the *E. coli* cells.

FIG. 4. Structure of (-)-drimenol produced by the recombinant DITps589 synthase.

FIG. 5. Chiral GC/FID chromatograms of (-)-drimenol produced by the recombinant enzyme (upper), racemic drimenol obtained chemically (middle) and authentic (-)-drimenol (lower).

FIG. 6. Total ion chromatogram of GCMS analysis of the sesquiterpenes produced in in-vitro assays by the recombinant proteins SCH51-3228-9 (A), SCH51-998-28 (B) or SCH52-13163-6 (C).

FIG. 7. Total ion chromatogram of GCMS analysis of the sesquiterpenes produced in vivo by engineered bacteria cells expressing the different recombinant proteins SCH51-3228-9 (A), SCH51-998-28 (B) or SCH52-13163-6 (C). The farnesol detected results from the hydrolysis of excess FPP produced by the *E. coli* cells or could be in part produced by the recombinant proteins.

## DETAILED DESCRIPTION

For the descriptions herein and the appended claims, the use of "or" means "and/or" unless stated otherwise. Similarly, "comprise," "comprises," "comprising" "include," "includes," and "including" are interchangeable and not intended to be limiting.

It is to be further understood that where descriptions of various embodiments use the term "comprising," those skilled in the art would understand that in some specific instances, an embodiment can be alternatively described using language "consisting essentially of" or "consisting of." In one aspect, provided here is a method of producing Drimenol comprising:

- i) contacting a acyclic terpene pyrophosphate, particularly farnesyl diphosphate (FPP) with a polypeptide having Drimenol synthase activity and having at least, or at least about 70%, particularly 75%, particularly 80%, particularly 85%, particularly 90%, particularly 95%, particularly 96%, particularly 97%, particularly 98% or particularly 99% or more sequence identity to a sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 11 and SEQ ID NO: 14 to produce the Drimenol; and

ii) optionally isolating the Drimenol.

In one aspect, the Drimenol is isolated.

Further provided here is an isolated polypeptide having Drimenol activity comprising an amino acid sequence having at least or at least about 70%, particularly 75%, particu-

larly 80%, particularly 85%, particularly 90%, particularly 95%, particularly 96%, particularly 97%, particularly 98% or more particularly 99% or more identity to amino acid sequence of a sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 11 and SEQ ID NO: 14.

Further provided herein is an isolated nucleic acid molecule encoding a polypeptide comprising an amino acid sequence having at least or at least about 70%, particularly 75%, particularly 80%, particularly 85%, particularly 90%, particularly 95%, particularly 96%, particularly 97%, particularly 98% or more particularly 99% or more identity to amino acid sequence of a sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 11 and SEQ ID NO: 14.

Further provided herein a nucleic acid molecule comprising the sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13 and SEQ ID NO: 15.

Further provided here is a method as recited in claim 1 comprising the steps of transforming a host cell or non-human organism with a nucleic acid encoding a polypeptide having at least, or at least about, 70%, particularly 75%, particularly 80%, particularly 85%, particularly 90%, particularly 95%, particularly 96%, particularly 97%, particularly 98% or particularly 99% or more sequence identity of the sequence of a sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 11 and SEQ ID NO: 14 and culturing the host cell or organism under conditions that allow for the production of the polypeptide.

Further provided is at least one vector comprising the nucleic acid molecules described.

Further provided herein is a vector selected from the group of a prokaryotic vector, viral vector and a eukaryotic vector.

Further provided here is a vector that is an expression vector.

As a "Drimenol synthase" or as a "polypeptide having a Drimenol synthase activity", we mean here a polypeptide capable of catalyzing the synthesis of Drimenol, in the form of any of its stereoisomers or a mixture thereof, starting from an acyclic terpene pyrophosphate, particularly FPP. Drimenol may be the only product or may be part of a mixture of sesquiterpenes.

The ability of a polypeptide to catalyze the synthesis of a particular sesquiterpene (for example Drimenol) can be simply confirmed by performing the enzyme assay as detailed in Example 2 to 5.

Polypeptides are also meant to include truncated polypeptides provided that they keep their Drimenol synthase activity.

As intended herein below, "a nucleotide sequence obtained by modifying SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13 or SEQ ID NO: 15 or the complement thereof" encompasses any sequence that has been obtained by changing the sequence of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 10 or SEQ ID NO: 12 or of the complement thereof using any method known in the art, for example by introducing any type of mutations such as deletion, insertion or substitution mutations. Examples of such methods are cited in the part of the description relative to the variant polypeptides and the methods to prepare them.

The percentage of identity between two peptidic or nucleotidic sequences is a function of the number of amino acids or nucleotide residues that are identical in the two sequences when an alignment of these two sequences has been generated. Identical residues are defined as residues that are the same in the two sequences in a given position of the alignment. The percentage of sequence identity, as used herein, is calculated from the optimal alignment by taking the number of residues identical between two sequences dividing it by the total number of residues in the shortest sequence and multiplying by 100. The optimal alignment is the alignment in which the percentage of identity is the highest possible. Gaps may be introduced into one or both sequences in one or more positions of the alignment to obtain the optimal alignment. These gaps are then taken into account as non-identical residues for the calculation of the percentage of sequence identity. Alignment for the purpose of determining the percentage of amino acid or nucleic acid sequence identity can be achieved in various ways using computer programs and for instance publicly available computer programs available on the world wide web. Preferably, the BLAST program (Tatiana et al, *FEMS Microbiol Lett.*, 1999, 174:247-250, 1999) set to the default parameters, available from the National Center for Biotechnology Information (NCBI), can be used to obtain an optimal alignment of peptidic or nucleotidic sequences and to calculate the percentage of sequence identity.

#### ABBREVIATIONS USED

bp base pair  
 kb kilo base  
 BSA bovine serum albumin  
 DNA deoxyribonucleic acid  
 cDNA complementary DNA  
 DTT dithiothreitol  
 FID Flame ionization detector  
 FPP farnesyl pyrophosphate  
 GC gaseous chromatograph  
 IPTG isopropyl-D-thiogalacto-pyranoside  
 LB lysogeny broth  
 MS mass spectrometer  
 MVA mevalonic acid  
 PCR polymerase chain reaction  
 RMCE recombinase-mediated cassette exchange  
 3'-/5'-RACE 3' and 5' rapid amplification of cDNA ends  
 RNA ribonucleic acid  
 mRNA messenger ribonucleic acid  
 miRNA micro RNA  
 siRNA small interfering RNA  
 rRNA ribosomal RNA  
 tRNA transfer RNA

#### Definitions

The term "polypeptide" means an amino acid sequence of consecutively polymerized amino acid residues, for instance, at least 15 residues, at least 30 residues, at least 50 residues. In some embodiments provided herein, a polypeptide comprises an amino acid sequence that is an enzyme, or a fragment, or a variant thereof.

The term "isolated" polypeptide refers to an amino acid sequence that is removed from its natural environment by any method or combination of methods known in the art and includes recombinant, biochemical and synthetic methods.

The term "protein" refers to an amino acid sequence of any length wherein amino acids are linked by covalent

peptide bonds, and includes oligopeptide, peptide, polypeptide and full length protein whether naturally occurring or synthetic.

The terms “Drimenol synthase” or “Drimenol synthase protein” refer to an enzyme that is capable of converting farnesyl diphosphate (FPP) to Drimenol.

The terms “biological function,” “function,” “biological activity” or “activity” refer to the ability of the Drimenol synthase to catalyze the formation of Drimenol from FPP.

The terms “nucleic acid sequence,” “nucleic acid,” and “polynucleotide” are used interchangeably meaning a sequence of nucleotides. A nucleic acid sequence may be a single-stranded or double-stranded deoxyribonucleotide, or ribonucleotide of any length, and include coding and non-coding sequences of a gene, exons, introns, sense and anti-sense complimentary sequences, genomic DNA, cDNA, miRNA, siRNA, mRNA, rRNA, tRNA, recombinant nucleic acid sequences, isolated and purified naturally occurring DNA and/or RNA sequences, synthetic DNA and RNA sequences, fragments, primers and nucleic acid probes. The skilled artisan is aware that the nucleic acid sequences of RNA are identical to the DNA sequences with the difference of thymine (T) being replaced by uracil (U).

An “isolated nucleic acid” or “isolated nucleic acid sequence” is defined as a nucleic acid or nucleic acid sequence that is in an environment different from that in which the nucleic acid or nucleic acid sequence naturally occurs. The term “naturally-occurring” as used herein as applied to a nucleic acid refers to a nucleic acid that is found in a cell in nature. For example, a nucleic acid sequence that is present in an organism, for instance in the cells of an organism, that can be isolated from a source in nature and which has not been intentionally modified by a human in the laboratory is naturally occurring.

“Recombinant nucleic acid sequence” are nucleic acid sequences that result from the use of laboratory methods (molecular cloning) to bring together genetic material from more than one source, creating a nucleic acid sequence that does not occur naturally and would not be otherwise found in biological organisms.

“Recombinant DNA technology” refers to molecular biology procedures to prepare a recombinant nucleic acid sequence as described, for instance, in Laboratory Manuals edited by Weigel and Glazebrook, 2002 Cold Spring Harbor Lab Press; and Sambrook et al., 1989 Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press.

The term “gene” means a DNA sequence comprising a region, which is transcribed into a RNA molecule, e.g., an mRNA in a cell, operably linked to suitable regulatory regions, e.g., a promoter. A gene may thus comprise several operably linked sequences, such as a promoter, a 5' leader sequence comprising, e.g., sequences involved in translation initiation, a coding region of cDNA or genomic DNA, introns, exons, and/or a 3' non-translated sequence comprising, e.g., transcription termination sites.

A “chimeric gene” refers to any gene, which is not normally found in nature in a species, in particular, a gene in which one or more parts of the nucleic acid sequence are present that are not associated with each other in nature. For example the promoter is not associated in nature with part or all of the transcribed region or with another regulatory region. The term “chimeric gene” is understood to include expression constructs in which a promoter or transcription regulatory sequence is operably linked to one or more coding sequences or to an antisense, i.e., reverse complement of the sense strand, or inverted repeat sequence (sense

and antisense, whereby the RNA transcript forms double stranded RNA upon transcription).

A “3' UTR” or “3' non-translated sequence” (also referred to as “3' untranslated region,” or “3'end”) refers to the nucleic acid sequence found downstream of the coding sequence of a gene, which comprises for example a transcription termination site and (in most, but not all eukaryotic mRNAs) a polyadenylation signal such as AAUAAA or variants thereof. After termination of transcription, the mRNA transcript may be cleaved downstream of the polyadenylation signal and a poly(A) tail may be added, which is involved in the transport of the mRNA to the site of translation, e.g., cytoplasm.

“Expression of a gene” involves transcription of the gene and translation of the mRNA into a protein. Overexpression refers to the production of the gene product as measured by levels of mRNA, polypeptide and/or enzyme activity in transgenic cells or organisms that exceeds levels of production in non-transformed cells or organisms of a similar genetic background.

“Expression vector” as used herein means a nucleic acid molecule engineered using molecular biology methods and recombinant DNA technology for delivery of foreign or exogenous DNA into a host cell. The expression vector typically includes sequences required for proper transcription of the nucleotide sequence. The coding region usually codes for a protein of interest but may also code for an RNA, e.g., an antisense RNA, siRNA and the like.

An “expression vector” as used herein includes any linear or circular recombinant vector including but not limited to viral vectors, bacteriophages and plasmids. The skilled person is capable of selecting a suitable vector according to the expression system. In one embodiment, the expression vector includes the nucleic acid of an embodiment herein operably linked to at least one regulatory sequence, which controls transcription, translation, initiation and termination, such as a transcriptional promoter, operator or enhancer, or an mRNA ribosomal binding site and, optionally, including at least one selection marker. Nucleotide sequences are “operably linked” when the regulatory sequence functionally relates to the nucleic acid of an embodiment herein. “Regulatory sequence” refers to a nucleic acid sequence that determines expression level of the nucleic acid sequences of an embodiment herein and is capable of regulating the rate of transcription of the nucleic acid sequence operably linked to the regulatory sequence. Regulatory sequences comprise promoters, enhancers, transcription factors, promoter elements and the like.

“Promoter” refers to a nucleic acid sequence that controls the expression of a coding sequence by providing a binding site for RNA polymerase and other factors required for proper transcription including without limitation transcription factor binding sites, repressor and activator protein binding sites. The meaning of the term promoter also include the term “promoter regulatory sequence”. Promoter regulatory sequences may include upstream and downstream elements that may influence transcription, RNA processing or stability of the associated coding nucleic acid sequence. Promoters include naturally-derived and synthetic sequences. The coding nucleic acid sequences is usually located downstream of the promoter with respect to the direction of the transcription starting at the transcription initiation site.

The term “constitutive promoter” refers to an unregulated promoter that allows for continual transcription of the nucleic acid sequence it is operably linked to.

As used herein, the term “operably linked” refers to a linkage of polynucleotide elements in a functional relationship. A nucleic acid is “operably linked” when it is placed into a functional relationship with another nucleic acid sequence. For instance, a promoter, or rather a transcription regulatory sequence, is operably linked to a coding sequence if it affects the transcription of the coding sequence. Operably linked means that the DNA sequences being linked are typically contiguous. The nucleotide sequence associated with the promoter sequence may be of homologous or heterologous origin with respect to the plant to be transformed. The sequence also may be entirely or partially synthetic. Regardless of the origin, the nucleic acid sequence associated with the promoter sequence will be expressed or silenced in accordance with promoter properties to which it is linked after binding to the polypeptide of an embodiment herein. The associated nucleic acid may code for a protein that is desired to be expressed or suppressed throughout the organism at all times or, alternatively, at a specific time or in specific tissues, cells, or cell compartment. Such nucleotide sequences particularly encode proteins conferring desirable phenotypic traits to the host cells or organism altered or transformed therewith. More particularly, the associated nucleotide sequence leads to the production of Drimenol in the organism. Particularly, the nucleotide sequence encodes Drimenol synthase.

“Target peptide” refers to an amino acid sequence which targets a protein, or polypeptide to intracellular organelles, i.e., mitochondria, or plastids, or to the extracellular space (secretion signal peptide). A nucleic acid sequence encoding a target peptide may be fused to the nucleic acid sequence encoding the amino terminal end, e.g., N-terminal end, of the protein or polypeptide, or may be used to replace a native targeting polypeptide.

The term “primer” refers to a short nucleic acid sequence that is hybridized to a template nucleic acid sequence and is used for polymerization of a nucleic acid sequence complementary to the template.

As used herein, the term “host cell” or “transformed cell” refers to a cell (or organism) altered to harbor at least one nucleic acid molecule, for instance, a recombinant gene encoding a desired protein or nucleic acid sequence which upon transcription yields a Drimenol synthase protein useful to produce Drimenol. The host cell is particularly a bacterial cell, a fungal cell or a plant cell. The host cell may contain a recombinant gene which has been integrated into the nuclear or organelle genomes of the host cell. Alternatively, the host may contain the recombinant gene extra-chromosomally. Homologous sequences include orthologous or paralogous sequences. Methods of identifying orthologs or paralogs including phylogenetic methods, sequence similarity and hybridization methods are known in the art and are described herein.

Paralogs result from gene duplication that gives rise to two or more genes with similar sequences and similar functions. Paralogs typically cluster together and are formed by duplications of genes within related plant species. Paralogs are found in groups of similar genes using pair-wise Blast analysis or during phylogenetic analysis of gene families using programs such as CLUSTAL. In paralogs, consensus sequences can be identified characteristic to sequences within related genes and having similar functions of the genes.

Orthologs, or orthologous sequences, are sequences similar to each other because they are found in species that descended from a common ancestor. For instance, plant species that have common ancestors are known to contain

many enzymes that have similar sequences and functions. The skilled artisan can identify orthologous sequences and predict the functions of the orthologs, for example, by constructing a polygenic tree for a gene family of one species using CLUSTAL or BLAST programs. A method for identifying or confirming similar functions among homologous sequences is by comparing of the transcript profiles in plants overexpressing or lacking (in knockouts/knock-downs) related polypeptides. The skilled person will understand that genes having similar transcript profiles, with greater than 50% regulated transcripts in common, or with greater than 70% regulated transcripts in common, or greater than 90% regulated transcripts in common will have similar functions. Homologs, paralogs, orthologs and any other variants of the sequences herein are expected to function in a similar manner by making plants producing Drimenol synthase proteins.

An embodiment of the provided herein provides amino acid sequences of Drimenol synthase proteins including orthologs and paralogs as well as methods for identifying and isolating orthologs and paralogs of the Drimenol synthases in other organisms. Particularly, so identified orthologs and paralogs of the Drimenol synthase retain Drimenol synthase activity and are capable of producing Drimenol starting from FPP precursors.

The term “selectable marker” refers to any gene which upon expression may be used to select a cell or cells that include the selectable marker. Examples of selectable markers are described below. The skilled artisan will know that different antibiotic, fungicide, auxotrophic or herbicide selectable markers are applicable to different target species.

“Drimenol” for purposes of this application refers to (–)-drimenol (CAS: 468-68-8).

The term “organism” refers to any non-human multicellular or unicellular organisms such as a plant, or a micro-organism. Particularly, a micro-organism is a bacterium, a yeast, an algae or a fungus. The term “plant” is used interchangeably to include plant cells including plant protoplasts, plant tissues, plant cell tissue cultures giving rise to regenerated plants, or parts of plants, or plant organs such as roots, stems, leaves, flowers, pollen, ovules, embryos, fruits and the like. Any plant can be used to carry out the methods of an embodiment herein.

The polypeptide to be contacted with an acyclic pyrophosphate, e.g. FPP, in vitro can be obtained by extraction from any organism expressing it, using standard protein or enzyme extraction technologies. If the host organism is an unicellular organism or cell releasing the polypeptide of an embodiment herein into the culture medium, the polypeptide may simply be collected from the culture medium, for example by centrifugation, optionally followed by washing steps and re-suspension in suitable buffer solutions. If the organism or cell accumulates the polypeptide within its cells, the polypeptide may be obtained by disruption or lysis of the cells and further extraction of the polypeptide from the cell lysate.

The polypeptide having a Drimenol synthase activity, either in an isolated form or together with other proteins, for example in a crude protein extract obtained from cultured cells or microorganisms, may then be suspended in a buffer solution at optimal pH. If adequate, salts, DTT, inorganic cations and other kinds of enzymatic co-factors, may be added in order to optimize enzyme activity. The precursor FPP is added to the polypeptide suspension, which is then incubated at optimal temperature, for example between 15 and 40° C., particularly between 25 and 35° C., more particularly at 30° C. After incubation, the Drimenol pro-

duced may be isolated from the incubated solution by standard isolation procedures, such as solvent extraction and distillation, optionally after removal of polypeptides from the solution.

According to another particularly embodiment, the method of any of the above-described embodiments is carried out in vivo. In this case, step a) comprises cultivating a non-human host organism or cell capable of producing FPP and transformed to express at least one polypeptide comprising an amino acid sequence at least 70% identical to a sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 11 and SEQ ID NO: 14 and having a Drimenol synthase activity, under conditions conducive to the production of Drimenol.

According to a more particular embodiment, the method further comprises, prior to step a), transforming a non human organism or cell capable of producing FPP with at least one nucleic acid encoding a polypeptide comprising an amino acid sequence at least 70% identical to a sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 11 and SEQ ID NO: 14 and having a Drimenol synthase activity, so that said organism expresses said polypeptide.

These embodiments provided herein are particularly advantageous since it is possible to carry out the method in vivo without previously isolating the polypeptide. The reaction occurs directly within the organism or cell transformed to express said polypeptide.

According to a more particular embodiment at least one nucleic acid used in any of the above embodiments comprises a nucleotide sequence that has been obtained by modifying SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 10 SEQ ID NO: 12, SEQ ID NO: 13 or SEQ ID NO: 15 or the complement thereof. According to another embodiment, the at least one nucleic acid is isolated from a plant of the Winteraceae family or the Canellaceae family, particularly from *Drimys Winteri* or *Drimys lanceolata*.

The organism or cell is meant to “express” a polypeptide, provided that the organism or cell is transformed to harbor a nucleic acid encoding said polypeptide, this nucleic acid is transcribed to mRNA and the polypeptide is found in the host organism or cell. The term “express” encompasses “heterologously express” and “over-express”, the latter referring to levels of mRNA, polypeptide and/or enzyme activity over and above what is measured in a non-transformed organism or cell. A more detailed description of suitable methods to transform a non-human host organism or cell will be described later on in the part of the specification that is dedicated to such transformed non-human host organisms or cells.

A particular organism or cell is meant to be “capable of producing FPP” when it produces FPP naturally or when it does not produce FPP naturally but is transformed to produce FPP, either prior to the transformation with a nucleic acid as described herein or together with said nucleic acid. Organisms or cells transformed to produce a higher amount of FPP than the naturally occurring organism or cell are also encompassed by the “organisms or cells capable of producing FPP”. Methods to transform organisms, for example microorganisms, so that they produce FPP are already known in the art.

To carry out an embodiment herein in vivo, the host organism or cell is cultivated under conditions conducive to the production of Drimenol. Accordingly, if the host is a transgenic plant, optimal growth conditions are provided, such as optimal light, water and nutrient conditions, for

example. If the host is a unicellular organism, conditions conducive to the production of Drimenol may comprise addition of suitable cofactors to the culture medium of the host. In addition, a culture medium may be selected, so as to maximize Drimenol synthesis. Optimal culture conditions are described in a more detailed manner in the following Examples.

Non-human host organisms suitable to carry out the method of an embodiment herein in vivo may be any non-human multicellular or unicellular organisms. In a particular embodiment, the non-human host organism used to carry out an embodiment herein in vivo is a plant, a prokaryote or a fungus. Any plant, prokaryote or fungus can be used. Particularly useful plants are those that naturally produce high amounts of terpenes. In a more particular embodiment the non-human host organism used to carry out the method of an embodiment herein in vivo is a microorganism. Any microorganism can be used but according to an even more particular embodiment said microorganism is a bacteria or yeast. Most particularly, said bacteria is *E. coli* and said yeast is *Saccharomyces cerevisiae*.

Some of these organisms do not produce FPP naturally. To be suitable to carry out the method of an embodiment herein, these organisms have to be transformed to produce said precursor. They can be so transformed either before the modification with the nucleic acid described according to any of the above embodiments or simultaneously, as explained above.

Isolated higher eukaryotic cells can also be used, instead of complete organisms, as hosts to carry out the method of an embodiment herein in vivo. Suitable eukaryotic cells may be any non-human cell, but are particularly plant or fungal cells.

In another particular embodiment, the polypeptide consists of an amino acid sequence at least at least 70%, particularly at least 75%, particularly at least 80%, particularly at least 85%, particularly at least 90%, particularly at least 95%, particularly at least 96%, particularly at least 97%, particularly at least 98% and even more particularly at least 99% sequence identity to a sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 11 and SEQ ID NO: 14. In an even more particular embodiment, said polypeptide consists of SEQ ID.

According to another particular embodiment, the at least one polypeptide having a Drimenol synthase activity used in any of the above-described embodiments or encoded by the nucleic acid used in any of the above-described embodiments comprises an amino acid sequence that is a variant of SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 11 or SEQ ID NO: 14 obtained by genetic engineering, provided that said variant keeps its Drimenol synthase activity, as defined above and has the required percentage of identity to SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 11 or SEQ ID NO: 14. In other terms, said polypeptide particularly comprises an amino acid sequence encoded by a nucleotide sequence that has been obtained by modifying SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 10 SEQ ID NO: 12, SEQ ID NO: 13 or SEQ ID NO: 15 or the complement thereof. According to a more particular embodiment, the at least one polypeptide having a Drimenol synthase activity used in any of the above-described embodiments or encoded by the nucleic acid used in any of the above-described embodiments consists of an amino acid sequence that is a variant of SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 11 or SEQ ID NO: 14 obtained by genetic engineering, i.e. an amino acid sequence encoded

by a nucleotide sequence that has been obtained by modifying SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 10 SEQ ID NO: 12, SEQ ID NO: 13 or SEQ ID NO: 15 or the complement thereof.

According to another particular embodiment, the at least one polypeptide having a Drimenol synthase activity used in any of the above-described embodiments or encoded by the nucleic acid used in any of the above-described embodiments is a variant of SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 11 or SEQ ID NO: 14 that can be found naturally in other organisms, such as other plant species, provided that it keeps its Drimenol synthase activity as defined above and has the required percentage of identity to of SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 11 or SEQ ID NO: 14.

As used herein, the polypeptide is intended as a polypeptide or peptide fragment that encompasses the amino acid sequences identified herein, as well as truncated or variant polypeptides, provided that they keep their Drimenol synthase activity as defined above and that they share at least the defined percentage of identity with the corresponding fragment of SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 11 or SEQ ID NO: 14.

Examples of variant polypeptides are naturally occurring proteins that result from alternate mRNA splicing events or from proteolytic cleavage of the polypeptides described herein. Variations attributable to proteolysis include, for example, differences in the N- or C-termini upon expression in different types of host cells, due to proteolytic removal of one or more terminal amino acids from the polypeptides of an embodiment herein. Polypeptides encoded by a nucleic acid obtained by natural or artificial mutation of a nucleic acid of an embodiment herein, as described thereafter, are also encompassed by an embodiment herein.

Polypeptide variants resulting from a fusion of additional peptide sequences at the amino and carboxyl terminal ends can also be used in the methods of an embodiment herein. In particular such a fusion can enhance expression of the polypeptides, be useful in the purification of the protein or improve the enzymatic activity of the polypeptide in a desired environment or expression system. Such additional peptide sequences may be signal peptides, for example. Accordingly, encompassed herein are methods using variant polypeptides, such as those obtained by fusion with other oligo- or polypeptides and/or those which are linked to signal peptides. Polypeptides resulting from a fusion with another functional protein, such as another protein from the terpene biosynthesis pathway, can also be advantageously be used in the methods of an embodiment herein.

According to another embodiment, the at least one polypeptide having a Drimenol synthase activity used in any of the above-described embodiments or encoded by the nucleic acid used in any of the above-described embodiments is isolated from a plant of the Winteraceae family or the Canellaceae family, particularly from *Drimys Winteri* or *Drimys lanceolata*.

An important tool to carry out the method of an embodiment herein is the polypeptide itself. A polypeptide having a Drimenol synthase activity and comprising an amino acid sequence at least 70% identical to SEQ ID NO:2 is therefore provided herein.

According to a particular embodiment, the polypeptide is capable of producing a mixture of sesquiterpenes wherein Drimenol represents at least 20%, particularly at least 30%, particularly at least 35%, particularly at least 90%, particularly at least 95%, more particularly at least 98% of the

sesquiterpenes produced. In another aspect provided here, the Drimenol is produced with greater than or equal to 95%, more particularly 98% selectivity.

According to a particular embodiment, the polypeptide comprises an amino acid sequence at least 70%, particularly at least 75%, particularly at least 80%, particularly at least 85%, particularly at least 90%, particularly at least 95%, particularly at least 96%, particularly at least 97%, particularly at least 98% and even more particularly at least 99% identical to a sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 11 and SEQ ID NO: 14. According to a more particular embodiment, the polypeptide comprises amino acid sequence selected from the group consisting of of SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 11 and SEQ ID NO: 14

According to another particular embodiment, the polypeptide consists of an amino acid sequence at least 70%, particularly at least 75%, particularly at least 80%, particularly at least 85%, particularly at least 90%, particularly at least 95%, particularly at least 96%, particularly at least 97%, particularly at least 98% and even more particularly at least 99% identical to a sequence selected from the group consisting of of SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 11 or SEQ ID NO: 14. According to a more particular embodiment, the polypeptide consists of an amino acid selected from the group consisting of of SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 11 or SEQ ID NO: 14.

The at least one polypeptide comprises an amino acid sequence that is a variant of SEQ ID NO:2, either obtained by genetic engineering or found naturally in *Drimys* plants or in other plant species. In other terms, when the variant polypeptide is obtained by genetic engineering, said polypeptide comprises an amino acid sequence encoded by a nucleotide sequence that has been obtained by modifying SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 10 SEQ ID NO: 12, SEQ ID NO: 13 or SEQ ID NO: 15 or the complement thereof. According to a more particular embodiment, the at least one polypeptide having a Drimenol synthase activity consists of an amino acid sequence that is a variant of SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 11 or SEQ ID NO: 14 obtained by genetic engineering, i.e. an amino acid sequence encoded by a nucleotide sequence that has been obtained by modifying SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 10 SEQ ID NO: 12, SEQ ID NO: 13 or SEQ ID NO: 15 or the complement thereof.

According to another embodiment, the polypeptide is isolated from a plant of the Winteraceae family or the Canellaceae family, particularly from *Drimys Winteri* or *Drimys lanceolata*. As used herein, the polypeptide is intended as a polypeptide or peptide fragment that encompasses the amino acid sequence identified herein, as well as truncated or variant polypeptides, provided that they keep their activity as defined above and that they share at least the defined percentage of identity with the corresponding fragment of SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 11 or SEQ ID NO: 14.

As mentioned above, the nucleic acid encoding the polypeptide of an embodiment herein is a useful tool to modify non-human host organisms or cells intended to be used when the method is carried out in vivo.

A nucleic acid encoding a polypeptide according to any of the above-described embodiments is therefore also provided herein.

According to a particular embodiment, the nucleic acid comprises a nucleotide sequence at least 50%, particularly at least 55%, particularly at least 60%, particularly at least 65%, particularly at least 70%, particularly at least 75%, particularly at least 80%, particularly at least 85%, particularly at least 90%, more particularly at least 95% particularly at least 96%, particularly at least 97%, particularly at least 98%, and even more particularly at least 99% identical to a sequence selected from the group consisting of a sequence selected from the group consisting of NO: 1, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 10 SEQ ID NO: 12, SEQ ID NO: 13 or SEQ ID NO: 15 or the complement thereof. According to a more particular embodiment, the nucleic acid comprises the nucleotide sequence selected from the group consisting of NO: 1, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 10 SEQ ID NO: 12, SEQ ID NO: 13 or SEQ ID NO: 15 thereof.

According to another particular embodiment, the nucleic acid consists of a nucleotide sequence at least 70%, particularly at least 75%, particularly at least 80%, particularly at least 85%, particularly at least 90%, particularly at least 95%, particularly at least 96%, particularly at least 97%, particularly at least 98% and even more particularly at least 99% or more identity to a sequence selected from the group consisting NO: 1, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 10 SEQ ID NO: 12, SEQ ID NO: 13 or SEQ ID NO: 15 or the complement thereof. According to an even more particular embodiment, the nucleic acid consists of a sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 10 SEQ ID NO: 12, SEQ ID NO: 13 or SEQ ID NO: 15 or the complement thereof.

The nucleic acid of an embodiment herein can be defined as including deoxyribonucleotide or ribonucleotide polymers in either single- or double-stranded form (DNA and/or RNA). The terms "nucleotide sequence" should also be understood as comprising a polynucleotide molecule or an oligonucleotide molecule in the form of a separate fragment or as a component of a larger nucleic acid. Nucleic acids of an embodiment herein also encompass certain isolated nucleotide sequences including those that are substantially free from contaminating endogenous material. The nucleic acid of an embodiment herein may be truncated, provided that it encodes a polypeptide encompassed herein, as described above.

In one embodiment, the nucleic acid of an embodiment herein can be either present naturally in plants of the *Drimys* species or other species, or be obtained by modifying SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 10 SEQ ID NO: 12, SEQ ID NO: 13 or SEQ ID NO: 15 or the complement thereof. Particularly said nucleic acid consists of a nucleotide sequence that has been obtained by modifying SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 10 SEQ ID NO: 12, SEQ ID NO: 13 or SEQ ID NO: 15 or the complement thereof.

The nucleic acids comprising a sequence obtained by mutation of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 10 SEQ ID NO: 12, SEQ ID NO: 13 or SEQ ID NO: 15 or the complement thereof are encompassed by an embodiment

herein, provided that the sequences they comprise share at least the defined percentage of identity with the corresponding fragments of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 10 SEQ ID NO: 12, SEQ ID NO: 13 or SEQ ID NO: 15 or the complement thereof and provided that they encode a polypeptide having a Drimenol synthase activity, as defined in any of the above embodiments. Mutations may be any kind of mutations of these nucleic acids, such as point mutations, deletion mutations, insertion mutations and/or frame shift mutations. A variant nucleic acid may be prepared in order to adapt its nucleotide sequence to a specific expression system. For example, bacterial expression systems are known to more efficiently express polypeptides if amino acids are encoded by particular codons.

Due to the degeneracy of the genetic code, more than one codon may encode the same amino acid sequence, multiple nucleic acid sequences can code for the same protein or polypeptide, all these DNA sequences being encompassed by an embodiment herein. Where appropriate, the nucleic acid sequences encoding the Drimenol synthase may be optimized for increased expression in the host cell. For example, nucleotides of an embodiment herein may be synthesized using codons particular by a host for improved expression.

Another important tool for transforming host organisms or cells suitable to carry out the method of an embodiment herein *in vivo* is an expression vector comprising a nucleic acid according to any embodiment of an embodiment herein. Such a vector is therefore also provided herein.

The expression vectors provided herein may be used in the methods for preparing a genetically transformed host organism and/or cell, in host organisms and/or cells harboring the nucleic acids of an embodiment herein and in the methods for making polypeptides having a Drimenol synthase activity, as disclosed further below.

Recombinant non-human host organisms and cells transformed to harbor at least one nucleic acid of an embodiment herein so that it heterologously expresses or over-expresses at least one polypeptide of an embodiment herein are also very useful tools to carry out the method of an embodiment herein. Such non-human host organisms and cells are therefore also provided herein.

A nucleic acid according to any of the above-described embodiments can be used to transform the non-human host organisms and cells and the expressed polypeptide can be any of the above-described polypeptides.

Non-human host organisms of an embodiment herein may be any non-human multicellular or unicellular organisms. In a particular embodiment, the non-human host organism is a plant, a prokaryote or a fungus. Any plant, prokaryote or fungus is suitable to be transformed according to the methods provided herein. Particularly useful plants are those that naturally produce high amounts of terpenes.

In a more particular embodiment the non-human host organism is a microorganism. Any microorganism is suitable to be used herein, but according to an even more particular embodiment said microorganism is a bacteria or yeast. Most particularly, said bacteria is *E. coli* and said yeast is *Saccharomyces cerevisiae*.

Isolated higher eukaryotic cells can also be transformed, instead of complete organisms. As higher eukaryotic cells, we mean here any non-human eukaryotic cell except yeast cells. Particular higher eukaryotic cells are plant cells or fungal cells.

A variant may also differ from the polypeptide of an embodiment herein by attachment of modifying groups

which are covalently or non-covalently linked to the polypeptide backbone. The variant also includes a polypeptide which differs from the polypeptide described herein by introduced N-linked or O-linked glycosylation sites, and/or an addition of cysteine residues. The skilled artisan will recognise how to modify an amino acid sequence and preserve biological activity.

The functionality or activity of any Drimenol synthase protein, variant or fragment, may be determined using various methods. For example, transient or stable overexpression in plant, bacterial or yeast cells can be used to test whether the protein has activity, i.e., produces Drimenol from FPP precursors. Drimenol synthase activity may be assessed in a microbial expression system, such as the assay described in Example 2 or 3 herein on the production of Drimenol, indicating functionality. A variant or derivative of a Drimenol synthase polypeptide of an embodiment herein retains an ability to produce Drimenol from FPP precursors. Amino acid sequence variants of the Drimenol synthases provided herein may have additional desirable biological functions including, e.g., altered substrate utilization, reaction kinetics, product distribution or other alterations.

An embodiment herein provides polypeptides of an embodiment herein to be used in a method to produce Drimenol by contacting an FPP precursor with the polypeptides of an embodiment herein either in vitro or in vivo.

Provided herein is also an isolated, recombinant or synthetic polynucleotide encoding a polypeptide or variant polypeptide provided herein.

An embodiment of an embodiment herein provides an isolated, recombinant or synthetic nucleic acid sequence of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 10 SEQ ID NO: 12, SEQ ID NO: 13 or SEQ ID NO: 15 or a variant thereof encoding for a Drimenol synthase having the amino acid sequence which is at least 70%, 75%, 80%, 85%, 90%, 92%, 95%, 96%, 97%, 98% or 99% identical to a amino acid sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 11 or SEQ ID NO: 14 or fragments thereof that catalyze production of Drimenol in a cell from a FPP precursor. Embodiments provided herein include, but are not limited to cDNA, genomic DNA and RNA sequences. Any nucleic acid sequence encoding the Drimenol synthase or variants thereof is referred herein as a Drimenol synthase encoding sequence.

According to a particular embodiment, the nucleic acid of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 10 SEQ ID NO: 12, SEQ ID NO: 13 or SEQ ID NO: 15 SEQ is the coding sequence of a Drimenol synthase gene encoding the Drimenol synthase obtained as described in the Examples.

A fragment of a polynucleotide of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 10 SEQ ID NO: 12, SEQ ID NO: 13 or SEQ ID NO: 15 refers to contiguous nucleotides that is particularly at least 15 bp, at least 30 bp, at least 40 bp, at least 50 bp and/or at least 60 bp in length of the polynucleotide of an embodiment herein herein. Particularly the fragment of a polynucleotide comprises at least 25, more particularly at least 50, more particularly at least 75, more particularly at least 100, more particularly at least 150, more particularly at least 200, more particularly at least 300, more particularly at least 400, more particularly at least 500, more particularly at least 600, more particularly at least 700, more particularly at least 800, more particularly at least 900, more particularly at least 1000 contiguous nucleotides of the

polynucleotide of the an embodiment herein. Without being limited, the fragment of the polynucleotides herein may be used as a PCR primer, and/or as a probe, or for anti-sense gene silencing or RNAi.

It is clear to the person skilled in the art that genes, including the polynucleotides of an embodiment herein, can be cloned on basis of the available nucleotide sequence information, such as found in the attached sequence listing, by methods known in the art. These include e.g. the design of DNA primers representing the flanking sequences of such gene of which one is generated in sense orientations and which initiates synthesis of the sense strand and the other is created in reverse complementary fashion and generates the antisense strand. Thermo stable DNA polymerases such as those used in polymerase chain reaction are commonly used to carry out such experiments. Alternatively, DNA sequences representing genes can be chemically synthesized and subsequently introduced in DNA vector molecules that can be multiplied by e.g. compatible bacteria such as e.g. *E. coli*.

In a related embodiment of an embodiment herein, PCR primers and/or probes for detecting nucleic acid sequences encoding a Drimenol synthase are provided. The skilled artisan will be aware of methods to synthesize degenerate or specific PCR primer pairs to amplify a nucleic acid sequence encoding the Drimenol synthase or fragments thereof, based on SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 10 SEQ ID NO: 12, SEQ ID NO: 13 or SEQ ID NO: 15. A detection kit for nucleic acid sequences encoding the Drimenol synthase may include primers and/or probes specific for nucleic acid sequences encoding the Drimenol synthase, and an associated protocol to use the primers and/or probes to detect nucleic acid sequences encoding the Drimenol synthase in a sample. Such detection kits may be used to determine whether a plant has been modified, i.e., transformed with a sequence encoding the Drimenol synthase.

Provided herein are nucleic acid sequences obtained by mutations of NO: 1, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 10 SEQ ID NO: 12, SEQ ID NO: 13 or SEQ ID NO: 15 such mutations can be routinely made. It is clear to the skilled artisan that mutations, deletions, insertions, and/or substitutions of one or more nucleotides can be introduced into the DNA sequence of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 10 SEQ ID NO: 12, SEQ ID NO: 13 or SEQ ID NO: 15. Generally, a mutation is a change in the DNA sequence of a gene that can alter the amino acid sequence of the polypeptide produced.

To test a function of variant DNA sequences according to an embodiment herein, the sequence of interest is operably linked to a selectable or screenable marker gene and expression of the reporter gene is tested in transient expression assays with protoplasts or in stably transformed plants. The skilled artisan will recognize that DNA sequences capable of driving expression are built as modules. Accordingly, expression levels from shorter DNA fragments may be different than the one from the longest fragment and may be different from each other. Further provided herein are also functional equivalents of the nucleic acid sequence coding the Drimenol synthase proteins, i.e., nucleotide sequences that hybridize under stringent conditions to the nucleic acid sequence of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 10 SEQ ID NO: 12, SEQ ID NO: 13 or SEQ ID NO: 15. The skilled artisan will be aware of methods to identify homolo-

gous sequences in other organisms and methods (identified in the Definition section herein) to determine the percentage of sequence identity between homologous sequences. Such newly identified to DNA molecules then can be sequenced and the sequence can be compared with the nucleic acid sequence of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 10 SEQ ID NO: 12, SEQ ID NO: 13 or SEQ ID NO: 15 and tested for functional equivalence. Provided herein are DNA molecules having at least 70% particularly 75%, particularly 80%, particularly 85%, particularly 90%, particularly 95%, particularly 96% particularly 97% particularly 98%, or more particularly 99% or more sequence identity to the nucleotide sequence of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 10 SEQ ID NO: 12, SEQ ID NO: 13 or SEQ ID NO: 15

A related embodiment provides a nucleic acid sequence which is complementary to the nucleic acid sequence according to SEQ ID NO:1 or SEQ ID NO:3, such as inhibitory RNAs, or nucleic acid sequence which hybridizes under stringent conditions to at least part of the nucleotide sequence according to NO: 1, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 10 SEQ ID NO: 12, SEQ ID NO: 13 or SEQ ID NO: 15

An alternative embodiment of provided herein provides a method to alter gene expression in a host cell. For instance, the polynucleotide of an embodiment herein may be enhanced or overexpressed or induced in certain contexts (e.g. following insect bites or stings or upon exposure to a certain temperature) in a host cell or host organism.

Alteration of expression of a polynucleotide provided hereinafter results in "ectopic expression" which is a different expression pattern in an altered and in a control or wild-type organism. Alteration of expression occurs from interactions of polypeptide of an embodiment herein with exogenous or endogenous modulators, or as a result of chemical modification of the polypeptide. The term also refers to an altered expression pattern of the polynucleotide of an embodiment herein which is altered below the detection level or completely suppressed activity.

In one embodiment, several Drimenol synthase encoding nucleic acid sequences are co-expressed in a single host, particularly under control of different promoters. Alternatively, several Drimenol synthase protein encoding nucleic acid sequences can be present on a single transformation vector or be co-transformed at the same time using separate vectors and selecting transformants comprising both chimeric genes. Similarly, one or more Drimenol synthase encoding genes may be expressed in a single plant together with other chimeric genes, for example encoding other proteins which enhance insect pest resistance, or others.

The nucleic acid sequences of an embodiment herein encoding Drimenol synthase proteins can be inserted in expression vectors and/or be contained in chimeric genes inserted in expression vectors, to produce Drimenol synthase proteins in a host cell or host organism. The vectors for inserting transgenes into the genome of host cells are well known in the art and include plasmids, viruses, cosmids and artificial chromosomes. Binary or co-integration vectors into which a chimeric gene is inserted are also used for transforming host cells.

An embodiment of the provided herein provides recombinant expression vectors comprising a nucleic acid sequence of a Drimenol synthase gene, or a chimeric gene comprising a nucleic acid sequence of a Drimenol synthase

gene, operably linked to associated nucleic acid sequences such as, for instance, promoter sequences. For example, a chimeric gene comprising a nucleic acid sequence of SEQ ID NO:1 or SEQ ID NO:3 may be operably linked to a promoter sequence suitable for expression in plant cells, bacterial cells or fungal cells, optionally linked to a 3' non-translated nucleic acid sequence.

Alternatively, the promoter sequence may already be present in a vector so that the nucleic acid sequence which is to be transcribed is inserted into the vector downstream of the promoter sequence. Vectors are typically engineered to have an origin of replication, a multiple cloning site, and a selectable marker.

The following examples are illustrative only and are not intended to limit the scope of the claims or embodiments provided herein.

## EXAMPLES

### Example 1

*Drimys lanceolata* and *Drimys winteri* Plant Material and Leaf Transcriptome Sequencing.

*Drimys winteri* and *Drimys lanceolata* plants were obtained from Bluebell Nursery (Leicestershire, UK). For analysis of the composition in terpene molecules, the leaves were collected and solvent extracted using MTBE (methyl tert-butyl ether). The extract was analyzed by GCMS using an Agilent 6890 Series GC system connected to an Agilent 5975 mass detector. The GC was equipped with 0.25 mm inner diameter by 30 m DB-1 ms capillary column (Agilent). The carrier gas was He at a constant flow of 1 mL/min. The initial oven temperature was 50° C. (1 min hold) followed by a gradient of 10° C./min to 300° C. The injection was made in a split/splitless injector set at 260° C. and used in splitless mode. The identification of the products was based on the comparison of the mass spectra and retention indices with authentic standards and internal mass spectra databases. The leaves of the two plants contained significant quantities of drimane sesquiterpene compounds including (-)-drimenol, polygodial and epipolygodial (FIG. 1).

Small leaves of *D. winteri* and *D. lanceolata* were thus taken for transcriptome analysis. Total RNA was extracted using the Concert™ Plant RNA Reagent (Invitrogen). This total RNA was processed using the Illumina Total RNA-Seq technique and the Illumina HiSeq 2000 sequencer. A total of 101 and 105 millions of paired-reads of 2x100 bp were generated for *D. winteri* and *D. lanceolata*, respectively. The reads were assembled using the Velvet de novo genomic assembler and the Oases software. For *D. winteri* 40,586 contigs with an average size of 1,080 bp were assembled and for *D. lanceolata* 28,255 contigs with an average size of 1,179 bp were obtained. The contigs were search using the tBlastn algorithm (Altschul et al. J. Mol. Biol. 215, 403-410, 1990) and using as query the amino acid sequences of known sesquiterpene synthases. This approach provided the sequences for 37 new putative sesquiterpene synthases. The enzymatic activity of these synthases were evaluated as described in the following examples for the synthases showing Drimenol synthase activity.

### Example 2. Functional Expression and Characterization of DITps589 from *D. Lanceolata*

The DNA sequences of one of the selected sesquiterpene synthases DITps589 was codon-optimized, synthesized in-

vitro and cloned in the pJ444-SR expression plasmid (DNA2.0, Menlo Park, Calif., USA).

Heterologous expression of the DITps589 synthases was performed in KRX *E. coli* cells (Promega). Single colonies of cells transformed with the pJ444SR-DITps589 expression plasmid were used to inoculate 5 ml LB medium. After 5 to 6 hours incubation at 37° C., the cultures were transferred to a 20° C. incubator and left 1 hour for equilibration. Expression of the protein was then induced by the addition of 1 mM IPTG and 0.2% L-rhamnose and the culture was incubated over-night at 20° C. The next day, the cells were collected by centrifugation, resuspended in 0.1 volume of 50 mM MOPSO pH 7, 10% glycerol and lysed by sonication. The extracts were cleared by centrifugation (30 min at 20,000 g) and the supernatants containing the soluble proteins were used for further experiments.

The crude *E. coli* protein extracts containing the recombinant protein were used for the characterization of the enzymatic activities. The assays were performed in 2 mL of 50 mM MOPSO pH 7, 10% glycerol, 1 mM DTT, 15 mM MgCl<sub>2</sub> in the presence of 80 μM of farnesyl-diphosphate (FPP, Sigma) and 0.1 to 0.5 mg of crude protein. The tubes were incubated 12 to 24 hours at 30° C. and extracted twice with one volume of pentane. After concentration under a nitrogen flux, the extracts were analysed by GC and GC-MS and compared to extracts from assays with control proteins. The analysis of the products formed by the enzymes was made by GCMS as described in example 1. A negative control was performed in the same conditions using *E. coli* cells transformed with an empty pJ444 plasmid. In these conditions, the DITps589 recombinant enzyme produced (-)-drimenol as major product with a selectivity over 98% (FIG. 2). The identity of (-)-drimenol was confirmed by matching of the mass spectrum and retention time of an authentic Drimenol standard isolated from Sandalwood Oil West (*Amyris balsamifera*).

#### Example 3. In Vivo Production of (-)-Drimenol in *E. coli* Cells Using DITps589

To evaluate the in-vivo production of (-)-drimenol in heterologous cells, *E. coli* cells were transformed with the pJ444SR-DITps589 expression plasmid and the production of sesquiterpenes from the endogenous FPP pool was evaluated. To increase the productivity of the cells, a heterologous FPP synthase and an the enzymes from a complete heterologous mevalonate (MVA) pathway were also expressed in the same cells. The construction of the expression plasmid containing an FPP synthase gene and the gene for a complete MVA pathway was described in patent WO2013064411 or in Schalk et al (2013) *J. Am. Chem. Soc.* 134, 18900-18903. Briefly, an expression plasmid was prepared containing two operons composed of the genes encoding the enzymes for a complete mevalonate pathway. A first synthetic operon consisting of an *E. coli* acetoacetyl-CoA thiolase (atoB), a *Staphylococcus aureus* HMG-CoA synthase (mvaS), a *Staphylococcus aureus* HMG-CoA reductase (mvaA) and a *Saccharomyces cerevisiae* FPP synthase (ERG20) genes was synthesized in-vitro (DNA2.0, Menlo Park, Calif., USA) and ligated into the NcoI-BamHI digested pACYCDuet-1 vector (Invitrogen) yielding pACYC-29258. A second operon containing a mevalonate kinase (MvaK1), a phosphomevalonate kinase (MvaK2), a mevalonate diphosphate decarboxylase (MvaD), and an isopentenyl diphosphate isomerase (idi) was amplified from genomic DNA of *Streptococcus pneumoniae* (ATCC BAA-334) and ligated into the second multicloning site of

pACYC-29258 providing the plasmid pACYC-29258-4506. This plasmid thus contains the genes encoding all enzymes of the biosynthetic pathway leading from acetyl-coenzyme A to FPP.

KRX *E. coli* cells (Promega) were co-transformed with the plasmid pACYC-29258-4506 and the plasmid pJ444SR-DITps589. Transformed cells were selected on carbenicillin (50 μg/ml) and chloramphenicol (34 μg/ml) LB-agarose plates. Single colonies were used to inoculate 5 mL liquid LB medium supplemented with the same antibiotics. The culture was incubated overnight at 37° C. The next day 2 mL of TB medium supplemented with the same antibiotics were inoculated with 0.2 mL of the overnight culture. After 6 hours incubation at 37° C., the culture was cooled down to 28° C. and 0.1 mM IPTG and 0.2% rhamnose were added to each tube. The cultures were incubated for 48 hours at 28° C. The cultures were then extracted twice with 2 volumes of MTBE, the organic phase were concentrated to 500 μL and analyzed by GC-MS as described above in Example 1.

In this in-vivo conditions the DITps589 recombinant enzyme produced (-)-drimenol as major product with the same apparent selectivity as in the in vitro assay described in example 2 (FIG. 3). Using these engineered *E. Coli* cells larger (1 L) culture were used to purified the sesquiterpene produced by the enzyme in sufficient quantity to confirm the structure by NMR analysis and specific rotation measurement as being the structure of (-)-drimenol shown in FIG. 4. The enantiopurity was confirmed by chiral GC analysis on a Varian CP-3800 GC system equipped with a ChiraSil column (Agilent) and using oven gradient temperature of 3.0° C./min from 125 to 180° C. (FIG. 5).

#### Example 4. Functional Expression and Characterization of SCH51-3228-9 and SCH51-3228-11 from *D. Winteri*

SCH51-3228-9 and SCH51-3228-11 are two other DNA sequences putatively encoding for sesquiterpene synthases and isolated from the *Drymis winteri* transcriptome sequences. The deduced amino acid sequences share 92.6 and 95.1% identity, respectively with the DITps589 amino acid sequence. The two sequences were codon optimized, synthesized in-vitro (Invitrogen) and cloned between the NdeI and KpnI restriction enzyme recognition sites of the pETDuet-1 (Novagen) expression plasmid (Invitrogen).

Heterologous expression of the SCH51-3228-9 and SCH51-3228-11 synthases was performed in BL21 (DE3) *E. coli* cells (Invitrogen). Single colonies of cells transformed with the pETDuet-SCH51-3228-9 or the pETDuet-SCH51-3228-11 expression plasmids were used to produce the recombinant enzymes as described in example 2. The crude *E. coli* protein extracts containing the recombinant proteins were used for the characterization of the enzymatic activities as described in example 2 except for the the GCMS analysis conditions which were performed as follows. The GCMS analysis was made on an Agilent 6890 Series GC system connected to an Agilent 5975 mass detector. The GC was equipped with 0.25 mm inner diameter by 30 m DB-1 ms capillary column (Agilent). The carrier gas was He at a constant flow of 1 mL/min. The initial oven temperature was 50° C. (5 min hold) followed by a gradient of 5° C./min to 300° C. The injection was made in split mode at 250° C. with a split ratio of 5:1.

The The two recombinant enzymes produced (-)-drimenol as major product with high selectivity. The identity of (-)-drimenol was confirmed by matching of the mass spec-

21

trum and retention time of an authentic Drimenol standard isolated from Sandalwood Oil West (*Amyris balsamifera*) (FIG. 6).

Using the whole *E. Coli* cell system and method described in example 3 (except for the GCMS analysis conditions which were as described above) Drimenol could also be produced in vivo in bacteria cultures using the SCH51-3228-9 and SCH51-3228-11 recombinant proteins (FIG. 7).

Example 5. Functional Expression and Characterization of SCH51-998-28 from *D. Winteri* and SCH52-13163-6 from *D. Lanceolata*

Similarly to example 4, the two cDNAs SCH51-998-28 and SCH52-13163-6 were optimized and cloned in the pETDuet expression plasmid.

The recombinant proteins were produced in in BL21 (DE3) *E. coli* cells (Invitrogen) and the in vitro assays using FPP as substrate were performed as described in example 2 and 4. These assays showed Drimenol synthase activity for SCH51-998-28 and SCH52-13163-6 (FIG. 6). Using *E. Coli* cells overproducing FPP from a recombinant mevalonate

22

pathway (example 2 and 4), Drimenol could also be produced in vivo using the SCH51-998-28 and SCH52-13163-6 proteins (FIG. 7).

Example 6. Sequence Comparison of the Drimenol Synthases from *Drimys* Species

The amino acid sequences of the Drimenol synthases from *Drimys winteri* and *Drimys lanceolata* were aligned using the ClustalW Multiple alignment program (Thompson et al, 1994, Nucleic Acid Res. 22(22), 4673-4680 and the sequence identities were calculated based on this alignment.

Percent identity (%) between the different Drimenol synthases from *Drimys* species:

	DITps589	SCH51-3228-11	SCH51-3228-9	SCH51-998-28	SCH52-13163-6
DITps589	ID	95.1	92.6	70.5	88
SCH51_3228_11	95.1	ID	97.1	70.6	87.6
SCH51_3228_9	92.6	97	ID	71	90.1
SCH51_998_28	70.5	70.6	0.71	ID	72.5
SCH52_13163_6	88	87.6	90.1	72.5	ID

SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 15

<210> SEQ ID NO 1

<211> LENGTH: 1680

<212> TYPE: DNA

<213> ORGANISM: *Drimys lanceolata*

<400> SEQUENCE: 1

```

atggatctta ttaatccctc cccagcggct tccaccctcc ctctcccagt tgatggagat      60
tcagaagttg ttaggcgatc tgccgggttt catccgacta tctggggcga tcacttcctc      120
tctacaagc ccgatccaaa gaaaatagat gcatggaata aaagggttga agagctgaag      180
gaagaagtga agaagatatt aagcaatgca aaagggacgg tggaagagct gaatttgatt      240
gatgatctcg tacaccttgg gattagtatt cattttgaga aggagattga tgatgcteta      300
caacacatct ttgatacca tcttgatgat tttcctaagg atgatctata tgcgcgcgct      360
ctccgatttg gcgtcttaag gaaacagggg caccgtgttt ctccagatgt attcaaaaaa      420
ttcaaagatg agcaggggaa tttcaaggca gaggttgagca ccgatcgcaa aggtttgcta      480
tgtttaaatg atgtggctta tctcagcaca agaggggaag atatcttgga tgaagccatt      540
cctttcactg aggagcacct taggtcttgt attagccatg tagatttcca tatggcagca      600
aaaattgaac attctctcga gcttcccctt catcatcgca taccagggt agagaacagg      660
cactacatct cagtctatga aggagacaag gaaaggaacg aagttgtcct tgagcttgcc      720
aatttagatt tcaatctgat tcaaatcttg caccaaagag agctgagaga catcacaatg      780
tggtggaagg agattgacct tgcagcaaag ctgcctttta ttaggatag gttggtggag      840
tgctactact ggatcatggg ggtctathtt gaaccaatat actcgagggc tagggttttt      900
tccacaaaaa tgacaatggt ggtctcagtt gtggacgaca tatatgatgt gtatgctacc      960
gaggatgagc ttcaactatt cactgatgcc atctataggt gggatgctga tgacattgat      1020
cagctgecte agtacttgaa agatgctttt atggctactc acaacactgt gaagacteta      1080
gaagaagaac ttgaaccaga aggaaactct tatcgtggat tctatgtaaa agatgcaatg      1140
aaggttttgg caagggatta ctttgtggag cacaaatggt ataacagaaa aattgtgcca      1200
    
```

-continued

---

```

tccgtagagg aatacttgaa aatttcttgc atcagtggtg ccgttcatat ggctacagtt 1260
cactgtattg ctgggatgta tgaattgca accaaagagg cattcgaatg gttgatgact 1320
gagcccaaac ttgttattga tgcattctctg attggctctc tccttgatga catgcagtcc 1380
acctcgtttg agcaacagag aggccactg tcatcagcag tacagtgta catggctgaa 1440
tatggtgtaa cagcgaaga agcatgtgaa aagctccgag atatggctgc aattgcttgg 1500
aaagatgtga acgaggcatg ccttaggccc acggttttcc ctatgcctat ccttttgccct 1560
tctatcaact tggcacgtgt ggcagaagtc atctacctac gtggagatgg atacacgcac 1620
gctgggggtg agaccaagaa acacatcacg gccatgcttg ttaagccaat tgaagtctga 1680

```

&lt;210&gt; SEQ ID NO 2

&lt;211&gt; LENGTH: 559

&lt;212&gt; TYPE: PRT

<213> ORGANISM: *Drimys lanceolata*

&lt;400&gt; SEQUENCE: 2

```

Met Asp Leu Ile Asn Pro Ser Pro Ala Ala Ser Thr Leu Pro Leu Pro
 1           5           10          15
Val Asp Gly Asp Ser Glu Val Val Arg Arg Ser Ala Gly Phe His Pro
 20          25          30
Thr Ile Trp Gly Asp His Phe Leu Ser Tyr Lys Pro Asp Pro Lys Lys
 35          40          45
Ile Asp Ala Trp Asn Lys Arg Val Glu Glu Leu Lys Glu Glu Val Lys
 50          55          60
Lys Ile Leu Ser Asn Ala Lys Gly Thr Val Glu Glu Leu Asn Leu Ile
 65          70          75          80
Asp Asp Leu Val His Leu Gly Ile Ser Tyr His Phe Glu Lys Glu Ile
 85          90          95
Asp Asp Ala Leu Gln His Ile Phe Asp Thr His Leu Asp Asp Phe Pro
100          105          110
Lys Asp Asp Leu Tyr Val Ala Ala Leu Arg Phe Gly Val Leu Arg Lys
115          120          125
Gln Gly His Arg Val Ser Pro Asp Val Phe Lys Lys Phe Lys Asp Glu
130          135          140
Gln Gly Asn Phe Lys Ala Glu Leu Ser Thr Asp Ala Lys Gly Leu Leu
145          150          155          160
Cys Leu Asn Asp Val Ala Tyr Leu Ser Thr Arg Gly Glu Asp Ile Leu
165          170          175
Asp Glu Ala Ile Pro Phe Thr Glu Glu His Leu Arg Ser Cys Ile Ser
180          185          190
His Val Asp Ser His Met Ala Ala Lys Ile Glu His Ser Leu Glu Leu
195          200          205
Pro Leu His His Arg Ile Pro Arg Leu Glu Asn Arg His Tyr Ile Ser
210          215          220
Val Tyr Glu Gly Asp Lys Glu Arg Asn Glu Val Val Leu Glu Leu Ala
225          230          235          240
Asn Leu Asp Phe Asn Leu Ile Gln Ile Leu His Gln Arg Glu Leu Arg
245          250          255
Asp Ile Thr Met Trp Trp Lys Glu Ile Asp Leu Ala Ala Lys Leu Pro
260          265          270
Phe Ile Arg Asp Arg Leu Val Glu Cys Tyr Tyr Trp Ile Met Gly Val
275          280          285

```

-continued

---

Tyr Phe Glu Pro Ile Tyr Ser Arg Ala Arg Val Phe Ser Thr Lys Met  
 290 295 300

Thr Met Leu Val Ser Val Val Asp Asp Ile Tyr Asp Val Tyr Ala Thr  
 305 310 315 320

Glu Asp Glu Leu Gln Leu Phe Thr Asp Ala Ile Tyr Arg Trp Asp Ala  
 325 330 335

Asp Asp Ile Asp Gln Leu Pro Gln Tyr Leu Lys Asp Ala Phe Met Val  
 340 345 350

Leu Tyr Asn Thr Val Lys Thr Leu Glu Glu Glu Leu Glu Pro Glu Gly  
 355 360 365

Asn Ser Tyr Arg Gly Phe Tyr Val Lys Asp Ala Met Lys Val Leu Ala  
 370 375 380

Arg Asp Tyr Phe Val Glu His Lys Trp Tyr Asn Arg Lys Ile Val Pro  
 385 390 395 400

Ser Val Glu Glu Tyr Leu Lys Ile Ser Cys Ile Ser Val Ala Val His  
 405 410 415

Met Ala Thr Val His Cys Ile Ala Gly Met Tyr Glu Ile Ala Thr Lys  
 420 425 430

Glu Ala Phe Glu Trp Leu Met Thr Glu Pro Lys Leu Val Ile Asp Ala  
 435 440 445

Ser Leu Ile Gly Arg Leu Leu Asp Asp Met Gln Ser Thr Ser Phe Glu  
 450 455 460

Gln Gln Arg Gly His Val Ser Ser Ala Val Gln Cys Tyr Met Ala Glu  
 465 470 475 480

Tyr Gly Val Thr Ala Glu Glu Ala Cys Glu Lys Leu Arg Asp Met Ala  
 485 490 495

Ala Ile Ala Trp Lys Asp Val Asn Glu Ala Cys Leu Arg Pro Thr Val  
 500 505 510

Phe Pro Met Pro Ile Leu Leu Pro Ser Ile Asn Leu Ala Arg Val Ala  
 515 520 525

Glu Val Ile Tyr Leu Arg Gly Asp Gly Tyr Thr His Ala Gly Gly Glu  
 530 535 540

Thr Lys Lys His Ile Thr Ala Met Leu Val Lys Pro Ile Glu Val  
 545 550 555

<210> SEQ ID NO 3  
 <211> LENGTH: 1680  
 <212> TYPE: DNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Codon optimized DNA sequence of D1Tps589 from  
 D. lanceolata

<400> SEQUENCE: 3

atggacctga ttaaccggag cctgtctgca tccaccctgc cactgccagt cgatggtgat 60  
 agcgaagttg tgcgccgtag cgcgggtttc catccgacca tctgggggtga ccactttctg 120  
 tcttataagc cggaccggaa aaagattgat gcgtggaaca agcgtgttga ggaactgaaa 180  
 gaagaggtca aaaagatttt gagcaatgag aaaggcacgg ttgaggaact gaatttgatt 240  
 gacgacctgg tacacctggg tattagctat cactttgaga aagaaatcga cgacgcgctg 300  
 cagcatatct tcgatacgca cctggatgat ttcccgaag atgacctcta cgtggctgag 360  
 ctgctgtttg gcgtcctgag taagcaagc catcgtgtca gcccgagcgt ctttaagaaa 420  
 ttcaaagaag agcaaggcaa cttcaaagcg gagctgtcaa ccgatgcaaa gggcctggtg 480  
 tgcttgaagc atgtggcgta cctgagcacc cgtggtgagg atatcctgga cgaagcgatc 540

-continued

---

cogttcacgg aagaacatth gogctogtgc attagccaog ttgatagcca catggcagcg	600
aagattgagc actctctgga gctgocgctg caccatcgca ttccgogtht agagaatcg	660
cattacatct ccgtgtacga gggtgacaaa gagcgtaatg aagtoctct ggagttggct	720
aacttggaact ttaatcttat ccagatcctg caccagcgg agctgogcga catcacgatg	780
tggtggaaaag aaattgatct ggccgcaaag ctgocgthta ttogtgaccg tctggtggag	840
tgttactatt ggattatggg cgtgtacttc gagocgatct acagocgogc gogcogtht	900
agcaccaaga tgaccatgct ggttagcgtg gtggatgaca tctatgatgt ctacgctacg	960
gaagatgagt tgcagctgth taccgaoccc atttacagat gggacogcga tgacattgat	1020
caactgoccg aatatctgaa agacgcctth atggtthctg acaacacogt caaaacctg	1080
gaagaagaac tggagocgga aggtaactct tatcogtht tctacogthaa agatgoggatg	1140
aaagthctgg cgcgtgacta thctgthgag cataagthgt acaatogthaa gatcogthccg	1200
thcgtthgaa agthactthgaa gattagctgt atcagcogthc cagthccacat gggcagcogthg	1260
cactgtatcg ccggcatgta tgagatogcc acgaaagaag cattogagthg gctgatgacc	1320
gagccgaaaac tgggtgattga cgcaagocthg attgthcogcc tgothggacga ththcagagc	1380
acgagctthg agcagcagcg cggthcatgth agctcogcag thcaatgcta catggctgag	1440
thcggthgta ctgocgaaaga agcatgogag aagthcogthg atatggcggc cattgogthgg	1500
aaagatgthga atgaagcatg cctgococcg accgththcc ogatgocgat ththactgocct	1560
agcatcaacc tggcacgthg ggcggaagth atctatctgc thggcogcgg thatacgcac	1620
gogggthggthg agactaagaa gcacatcacc gogatgctgg thcaagcogat cgaagthgaa	1680

&lt;210&gt; SEQ ID NO 4

&lt;211&gt; LENGTH: 1656

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Drimys winteri

&lt;400&gt; SEQUENCE: 4

atggctthca ccctccctct cccagcttht ggagattcag aagthgthtag gogatctgoc	60
gggththcatc cgaogatctg gggcogatcac thctctctct acaagcctga thcaacgaaa	120
atagatgaaat ggaataaaa ggtthgaagag ctgaaggaag aagthgaagaa gatathaaagc	180
aatgcaaaa ggcagctgga agagctgaaat thgctthgatg atctogthaca cctthgggatt	240
agthtatcatt thgagaagga gattgatgat gctthacaac aaatctthga thcccatctt	300
gatgthththc ctaaggatga tctatathgoc accgctctcc gattthggcgt cthaaaggaaa	360
caggggcacc gthgththctc agatgthattc aaaaaattca aagatgagca ggggaatttc	420
aaggcagagth tgagcaccga thcgaagggth thgctatgth thatagatgth ggetthattc	480
agcacaagag gggaaagatath cthggatgaa gccattcctt thcactaagga gcacctagg	540
thctthgthta gccatgthoga thctcatathg cagcaaaaa thgagcattc thtagagctt	600
cccctthcatc atcgcatacc aaggctagag aacaggcact acatctcagth ctatgaagga	660
gacaagggaaa ggaatgaagth thgctctthgag cthgoccaaat thagaththcaa thctgattcaa	720
atctthgcacc aaagagagct gagggacatc acaacgthggth ggaagggagat thgaccthga	780
gcaaaogthac cthththattag gthagagthgth thgthgagthgct actathggat catgggagthc	840
ththththgaa caatathactc aagggctgaa gththththcga ccaaaatgac aatctthggthc	900
thcagthgthgg acgacathata thgatthata gthcagaggg atgagcthca actththcact	960

-continued

---

```

gatgcaatct ataggtggga tgctgaggac attgagcagc ttccacagta cttgaaagat 1020
gcttttcttg tactctataa cactgtgaag gacctagaag aggaattgga accagaagga 1080
aactcttatac gtggatacta tgtaaaagat gcgatgaagg ttttggcaag ggattacttt 1140
gtggagcaca aatggtataa cagaaaaatt gtgccatcag tagaggacta cctgccaatt 1200
tcttgcatca gtgttgccgt tcatatggcc acagttcatt gtattgctgg gatgatgaa 1260
attgcaacca aagaggcatt cgaatggttg aagacggaac ctaaacttgt tatagatgca 1320
tcactgattg ggcgtctoct cgatgacatg cagtccacct cgtttgagca acagagaggt 1380
catgtgtcat cagcggatca gtgttcatcg atccaatagc gggatcaca cgaagaagcg 1440
tgtgagaagt tgcgagaaat ggctgcaatt gcgtggaaag atgtaaacca agcatgcctt 1500
aggcccactg ttttccctat gcctattctt ctgccctcca tcaaccttgc acgtgtggca 1560
gaagtgattt acctacgocg agatggatat acacatgocg gtggtgagac caaaaaacat 1620
atcacggcca tgcttgttga tccaatcaaa gtctga 1656
    
```

```

<210> SEQ ID NO 5
<211> LENGTH: 551
<212> TYPE: PRT
<213> ORGANISM: Drimys winteri
    
```

<400> SEQUENCE: 5

```

Met Ala Ser Thr Leu Pro Leu Pro Ala Tyr Gly Asp Ser Glu Val Val
1          5          10          15
Arg Arg Ser Ala Gly Phe His Pro Thr Ile Trp Gly Asp His Phe Leu
20         25         30
Ser Tyr Lys Pro Asp Pro Thr Lys Ile Asp Glu Trp Asn Lys Arg Val
35         40         45
Glu Glu Leu Lys Glu Glu Val Lys Lys Ile Leu Ser Asn Ala Lys Gly
50         55         60
Thr Val Glu Glu Leu Asn Leu Leu Asp Asp Leu Val His Leu Gly Ile
65         70         75         80
Ser Tyr His Phe Glu Lys Glu Ile Asp Asp Ala Leu Gln Gln Ile Phe
85         90         95
Asp Thr His Leu Asp Val Phe Pro Lys Asp Asp Leu Tyr Ala Thr Ala
100        105        110
Leu Arg Phe Gly Val Leu Arg Lys Gln Gly His Arg Val Ser Pro Asp
115        120        125
Val Phe Lys Lys Phe Lys Asp Glu Gln Gly Asn Phe Lys Ala Glu Leu
130        135        140
Ser Thr Asp Ala Lys Gly Leu Leu Cys Leu Tyr Asp Val Ala Tyr Leu
145        150        155        160
Ser Thr Arg Gly Glu Asp Ile Leu Asp Glu Ala Ile Pro Phe Thr Lys
165        170        175
Glu His Leu Arg Ser Cys Ile Ser His Val Asp Ser His Met Ala Ala
180        185        190
Lys Ile Glu His Ser Leu Glu Leu Pro Leu His His Arg Ile Pro Arg
195        200        205
Leu Glu Asn Arg His Tyr Ile Ser Val Tyr Glu Gly Asp Lys Glu Arg
210        215        220
Asn Glu Val Val Leu Glu Leu Ala Lys Leu Asp Phe Asn Leu Ile Gln
225        230        235        240
Ile Leu His Gln Arg Glu Leu Arg Asp Ile Thr Thr Trp Trp Lys Glu
245        250        255
    
```

-continued

Ile Asp Leu Ala Ala Lys Leu Pro Phe Ile Arg Asp Arg Leu Val Glu  
 260 265 270

Cys Tyr Tyr Trp Ile Met Gly Val Tyr Phe Glu Pro Ile Tyr Ser Arg  
 275 280 285

Ala Arg Val Phe Ser Thr Lys Met Thr Ile Leu Val Ser Val Val Asp  
 290 295 300

Asp Ile Tyr Asp Val Tyr Ala Thr Glu Asp Glu Leu Gln Leu Phe Thr  
 305 310 315 320

Asp Ala Ile Tyr Arg Trp Asp Ala Glu Asp Ile Glu Gln Leu Pro Gln  
 325 330 335

Tyr Leu Lys Asp Ala Phe Leu Val Leu Tyr Asn Thr Val Lys Asp Leu  
 340 345 350

Glu Glu Glu Leu Glu Pro Glu Gly Asn Ser Tyr Arg Gly Tyr Tyr Val  
 355 360 365

Lys Asp Ala Met Lys Val Leu Ala Arg Asp Tyr Phe Val Glu His Lys  
 370 375 380

Trp Tyr Asn Arg Lys Ile Val Pro Ser Val Glu Asp Tyr Leu Arg Ile  
 385 390 395 400

Ser Cys Ile Ser Val Ala Val His Met Ala Thr Val His Cys Cys Ala  
 405 410 415

Gly Met Asp Glu Ile Ala Thr Lys Glu Ala Phe Glu Trp Leu Lys Thr  
 420 425 430

Glu Pro Lys Leu Val Ile Asp Ala Ser Leu Ile Gly Arg Leu Leu Asp  
 435 440 445

Asp Met Gln Ser Thr Ser Phe Glu Gln Gln Arg Gly His Val Ser Ser  
 450 455 460

Ala Val Gln Cys Tyr Met Ile Gln Tyr Gly Val Ser His Glu Glu Ala  
 465 470 475 480

Cys Glu Lys Leu Arg Glu Met Ala Ala Ile Ala Trp Lys Asp Val Asn  
 485 490 495

Gln Ala Cys Leu Arg Pro Thr Val Phe Pro Met Pro Ile Leu Leu Pro  
 500 505 510

Ser Ile Asn Leu Ala Arg Val Ala Glu Val Ile Tyr Leu Arg Gly Asp  
 515 520 525

Gly Tyr Thr His Ala Gly Gly Glu Thr Lys Lys His Ile Thr Ala Met  
 530 535 540

Leu Val Asp Pro Ile Lys Val  
 545 550

&lt;210&gt; SEQ ID NO 6

&lt;211&gt; LENGTH: 1656

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Codon optimized DNA sequence of SCH51-3228-9

&lt;400&gt; SEQUENCE: 6

```

atggcaagca ccctgcccct gctgcctat ggtgatagcg aagttgttcg tcgtagegca    60
ggttttcacg cgaccatttg gggtgatcat tttctgagct ataaaccgga tccgaccaa    120
attgatgaat ggaataaacg tgtcgaagaa ctgaaagaag aagtgaaaaa aatcctgagc    180
aatgccaaag gcaccgttga ggaactgaat ctgctggatg atctggttca tctgggtatc    240
agctatcact ttgagaaaga aatcgatgat gcactgcagc agatttttga taccatctg    300
gatgttttcc cgaagatga tctgtatgca accgcactgc gttttggtgt tctgcgtaaa    360

```

-continued

---

```

cagggtcatc gtgtagtcc ggatgtgttc aaaaaattca aagatgaaca gggcaacttc 420
aaagcagaac tgagcaccga tgcaaaaggc ctgctgtgtc tgtatgatgt tgcatactcg 480
agcaccocgtg gtgaagatat tctggatgaa gcaattccgt ttaccaaaga acatctgcgt 540
agctgtatta gccatgttga tagccacatg gcagcgaaaa ttgaacatag cctggaactg 600
cctctgcatc accgtattcc gcgtctggaa aatcgctact atattagcgt ttatgagggc 660
gataaagaac gcaatgaagt tgtgctggaa ctggcaaaac tggattttaa cctgattcag 720
attctgcatc agcgtgaact gcgtgatatt accacctggg ggaaagaaat tgatctggca 780
gcaaaactgc cgtttattcg tgatcgtctg gttgaatgct attattggat tatggcgctg 840
tatttcgaac cgatttatag ccgtgcacgt gtttttagca ccaaaatgac cattctgggt 900
agcgtgggtg atgatatcta tgatgtttat gccaccgaag atgaactgca gctgtttacc 960
gatgccattt atcgttggga tgcagaagat attgaacagc tgccgcagta tctgaaagat 1020
gcatttctgg ttctgtacaa caccgtgaaa gatctggaag aagaactgga accggaaggc 1080
aatagctatc gtggttatta tgttaaagat gccatgaaag ttctggcagc cgattatctt 1140
gttgagcaca aatggtataa ccgcaaaatt gttccgagcg tggaagatta tctgcgtatt 1200
agctgcatta gcgttgacgt tcacatggca accgttcatt gttgtgcagg tatggatgaa 1260
attgcaacca aagaagcatt tgagtggctg aaaaccgaac cgaaactggc tattgatgca 1320
agcctgattg gtcgtctgct ggacgatatg cagagcacca gctttgaaca gcagcgtggc 1380
catgttagca gcgcagttca gtgttatatg attcagatg gtgttagcca tgaagaagca 1440
tgcgaaaaac tgcgcgaaat ggcagcaatt gcatggaag atgttaatca ggcagtctcg 1500
cgtccgacgc tttttccgat gccgattctg ctgccgagca ttaatctggc acgtgttgcc 1560
gaagttatct atctgcgtgg tgatggttat acccatgccg gtggtgaaac caaaaaacat 1620
attaccgcaa tgctggtoga tccgattaaa gtttaa 1656

```

&lt;210&gt; SEQ ID NO 7

&lt;211&gt; LENGTH: 1656

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Drimys winteri

&lt;400&gt; SEQUENCE: 7

```

atggcttcca ccctccctct cccagcttat ggagattcag aagttgttag gcgatctgcc 60
gggtttcatc cgacgatctg gggcgatcac ttcctctcct acaagcctga tccaacgaaa 120
atagatgaat ggaataaaag ggttgaagag ctgaaggaag aagtgaagaa gatattaagc 180
aatgcaaaag ggacagtgga agagctgaat ttgcttgatg atctcgtaca ccttgggatt 240
agttatcatt ttgagaagga gattgatgat gctttacaac aaatcttga taccatctt 300
gatgtttttc ctaaggatga tctatatgcc accgctctcc gatttggcgt cttaaggaaa 360
caggggcacc gtgtttctcc agatgtatc aaaaaattca aagatgagca ggggaatttc 420
aaggcagagt tgagcaccga tgcgaagggt ttgctatggt tatatgatgt ggcttatctc 480
agcacaagag ggaagatat cttggatgaa gccattcctt tcaactaagga gcaccttagg 540
tcttgattta gccatgtoga ttctcatatg gcagcaaaaa ttgagcattc tctagagctt 600
ccccctcatc atcgcatacc aaggctagag aacaggcact acatctcagt ctatgaagga 660
gacaagggaa ggaatgaagt tgccttgag cttgccaaat tagatttcaa tctgattcaa 720
atcttgacc aaagagagct gagggacatc acaatgtggt ggaaggagat tgaccttgca 780

```

-continued

---

```

gcaaagctac cttttattag agataggttg gtggagtgct actactggat catgggggtc 840
tattttgaac caatatactc cagggctagg gttttttcca ctaaaatgac aatcttggtc 900
tcagttgtgg acgacatata tgatgtctat gctacggagg atgagcttca actattcact 960
gatgcaatct ataggtggga tgctgatgac attgatcagc tgcctcagta cttgaaagat 1020
gcttttatgg tactctataa cactgtgaag actctagaag aagaactga accagaagga 1080
aactcttatac gtggatacta cgtaaaagat gcaatgaagg ttttggcaag agattacttt 1140
gtggaacaca aatggtataa cagacaaatt gtgccatccg tagaggaata cttgaaaatt 1200
tcttgcatca gtgtggctgt tcatatggct acagttcatt gtattgtctg gatgtatgaa 1260
attgtacca aagaggcatt cgaatggttg aagactgaac ccaaactgtt tatcgatgca 1320
tctctgatcg gtcgtcttct tgatgacatg cagtctacct cgtttgagca acaaagaggg 1380
cacgtgtcat cagcagtaca gtgttacatg gcccaatatg gagtaacagc agaagaagca 1440
tgtgaaaagc tacgagaaat ggctgcaatt gcttggaag atgtgaaatga agcatgcctt 1500
aggcccacgg tattccctat gcctatctctc ttgcctteta tcaacttggc acgtgtggca 1560
gaagtgatct acctacgtgg agatggatac acgcacgctg ggggtgagac caaaaaaac 1620
atcacggcca tgcttgtaa gccaatgaa gtctga 1656
    
```

```

<210> SEQ ID NO 8
<211> LENGTH: 551
<212> TYPE: PRT
<213> ORGANISM: Drimys winteri
    
```

<400> SEQUENCE: 8

```

Met Ala Ser Thr Leu Pro Leu Pro Ala Tyr Gly Asp Ser Glu Val Val
1          5          10          15
Arg Arg Ser Ala Gly Phe His Pro Thr Ile Trp Gly Asp His Phe Leu
20        25        30
Ser Tyr Lys Pro Asp Pro Thr Lys Ile Asp Glu Trp Asn Lys Arg Val
35        40        45
Glu Glu Leu Lys Glu Glu Val Lys Lys Ile Leu Ser Asn Ala Lys Gly
50        55        60
Thr Val Glu Glu Leu Asn Leu Leu Asp Asp Leu Val His Leu Gly Ile
65        70        75        80
Ser Tyr His Phe Glu Lys Glu Ile Asp Asp Ala Leu Gln Gln Ile Phe
85        90        95
Asp Thr His Leu Asp Val Phe Pro Lys Asp Asp Leu Tyr Ala Thr Ala
100       105       110
Leu Arg Phe Gly Val Leu Arg Lys Gln Gly His Arg Val Ser Pro Asp
115       120       125
Val Phe Lys Lys Phe Lys Asp Glu Gln Gly Asn Phe Lys Ala Glu Leu
130       135       140
Ser Thr Asp Ala Lys Gly Leu Leu Cys Leu Tyr Asp Val Ala Tyr Leu
145       150       155       160
Ser Thr Arg Gly Glu Asp Ile Leu Asp Glu Ala Ile Pro Phe Thr Lys
165       170       175
Glu His Leu Arg Ser Cys Ile Ser His Val Asp Ser His Met Ala Ala
180       185       190
Lys Ile Glu His Ser Leu Glu Leu Pro Leu His His Arg Ile Pro Arg
195       200       205
Leu Glu Asn Arg His Tyr Ile Ser Val Tyr Glu Gly Asp Lys Glu Arg
210       215       220
    
```

-continued

---

Asn Glu Val Val Leu Glu Leu Ala Lys Leu Asp Phe Asn Leu Ile Gln  
 225 230 235 240

Ile Leu His Gln Arg Glu Leu Arg Asp Ile Thr Met Trp Trp Lys Glu  
 245 250 255

Ile Asp Leu Ala Ala Lys Leu Pro Phe Ile Arg Asp Arg Leu Val Glu  
 260 265 270

Cys Tyr Tyr Trp Ile Met Gly Val Tyr Phe Glu Pro Ile Tyr Ser Arg  
 275 280 285

Ala Arg Val Phe Ser Thr Lys Met Thr Ile Leu Val Ser Val Val Asp  
 290 295 300

Asp Ile Tyr Asp Val Tyr Ala Thr Glu Asp Glu Leu Gln Leu Phe Thr  
 305 310 315 320

Asp Ala Ile Tyr Arg Trp Asp Ala Asp Asp Ile Asp Gln Leu Pro Gln  
 325 330 335

Tyr Leu Lys Asp Ala Phe Met Val Leu Tyr Asn Thr Val Lys Thr Leu  
 340 345 350

Glu Glu Glu Leu Glu Pro Glu Gly Asn Ser Tyr Arg Gly Tyr Tyr Val  
 355 360 365

Lys Asp Ala Met Lys Val Leu Ala Arg Asp Tyr Phe Val Glu His Lys  
 370 375 380

Trp Tyr Asn Arg Gln Ile Val Pro Ser Val Glu Glu Tyr Leu Lys Ile  
 385 390 395 400

Ser Cys Ile Ser Val Ala Val His Met Ala Thr Val His Cys Ile Ala  
 405 410 415

Gly Met Tyr Glu Ile Ala Thr Lys Glu Ala Phe Glu Trp Leu Lys Thr  
 420 425 430

Glu Pro Lys Leu Val Ile Asp Ala Ser Leu Ile Gly Arg Leu Leu Asp  
 435 440 445

Asp Met Gln Ser Thr Ser Phe Glu Gln Gln Arg Gly His Val Ser Ser  
 450 455 460

Ala Val Gln Cys Tyr Met Ala Gln Tyr Gly Val Thr Ala Glu Glu Ala  
 465 470 475 480

Cys Glu Lys Leu Arg Glu Met Ala Ala Ile Ala Trp Lys Asp Val Asn  
 485 490 495

Glu Ala Cys Leu Arg Pro Thr Val Phe Pro Met Pro Ile Leu Leu Pro  
 500 505 510

Ser Ile Asn Leu Ala Arg Val Ala Glu Val Ile Tyr Leu Arg Gly Asp  
 515 520 525

Gly Tyr Thr His Ala Gly Gly Glu Thr Lys Lys His Ile Thr Ala Met  
 530 535 540

Leu Val Lys Pro Ile Glu Val  
 545 550

<210> SEQ ID NO 9  
 <211> LENGTH: 1656  
 <212> TYPE: DNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Codon optimized DNA sequence of SCH51-3228-11  
 <400> SEQUENCE: 9

atggcatceta ctcttccact gccggettat ggtgattctg aggttggttcg tcgttccgcg 60  
 ggttttcacc ctaccatctg gggcgatcac tttctgtcct ataagccaga cccgaccaag 120  
 attgacgagt ggaataagcg tgtcgaggaa ctgaaagaag aagtgaaaaa gatcctgtcc 180

-continued

---

aacgcaaaag gtactgtoga ggagctgaat ctgctggatg acctggtgca tctgggcatc	240
agctatcaact tcgaaaagga aattgacgac gctttgcagc aaatttttga tacgcacctg	300
gacgtctttc cgaagatga cctgtatgcg accgcgctgc gctttggtgt gctgcgtaaa	360
cagggtcatc gcgtgtctcc tgatgtgttc aagaaattta aagatgaaca gggcaatttc	420
aaggccgagt tgagcacgga cgccaaaggt ttgctctgcc tgtacgacgt tgcatactg	480
agcaccctgtg gtgaagatat cctggacgaa gcgattccgt tcaccaagga acatctgcgc	540
tcgtgcattt cccatgtaga tagccacatg gcggccaaga tcgagcacag cctggagctg	600
cctttgcaacc atcgtattcc gcgcctggag aatcgccatt acattagcgt ctatgagggg	660
gacaaagagc gcaacgaagt cgtgttagag ctggcgaagc tggacttcaa cctgattcaa	720
attctgcac c aacgcgagct gcgcgacatt accatgtggt ggaaagagat tgatctggca	780
gcgaagctgc cgttcatccg cgatcgtctg gttgagtgct actactggat catgggcgtc	840
tacttcgagc cgatctacag ccgcgctcgt gtgttttcga cgaagatgac catcctgggt	900
agcgttgttg atgacattta tgacgtttac gcgaccgaag atgaactgca gctgtttacg	960
gacgcaatct accgttggga gcgcgatgat atcgaccagc tgccgcaata cttgaaagat	1020
gcgttcatgg ttttgtacaa caccgtcaaa acgctggaag aagaactgga gccggaaggc	1080
aacagctacc gtggttacta tgttaaagat gcgatgaaag ttctggcgcg cgactacttc	1140
gtcgagcaca agtggataaa ccgtcagatt gtgccgagcg tcgaggaata cctgaagatt	1200
agctgtatca gcgttgccgt tcacatggca acggtgcact gcacgcgcgg tatgtacgag	1260
attgcgacga aagaagcctt cgaatgggtg aaaaccgagc cgaagctggt tatcgacgcc	1320
agcctgatcg gtcgtttgct ggacgacatg caaagcacga gcttcgagca gcagcgcggc	1380
catgtgagca gcgctgttca gtgttatatg gcgcaaatg gcgtgaccgc agaagaagcg	1440
tcgagagaagc tcgctgagat ggcagcaatt gcgtggaag atgtgaatga agcctgtctg	1500
cgcccgactg tgtttccgat gccgatcctg ctgccgagca ttaacctggc gcgtgtggca	1560
gaggteatct atctgcgtgg tgacggttac acccagcggc gtggcgaaac caagaaacat	1620
atcaccgcaa tgctggttaa gccgattgaa gtgtaa	1656

&lt;210&gt; SEQ ID NO 10

&lt;211&gt; LENGTH: 1677

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Drimys winteri

&lt;400&gt; SEQUENCE: 10

atggatctta gtacttcacc tgttctttct tctctcccc ttcgggtgga agacggaaaa	60
aatccggccg ttcgcgcttc agctggattt caccocagta tttggggtga tcatttctc	120
tctacactg aagatcacia gaagctggat gcatggagcg aaaggactca agtggtgaag	180
gaagaggtga ggagaatttt aatcaatgcc aaggggtcac tagaagagtt ggatttggtg	240
gatgcaatcc aacgccttgg ggtgaaatat cactttgaga aagagattga agaggcatta	300
caccatattt atgttgacga aactcatggt tctactgatg acttatattc cgtttctctc	360
cggtttcgac ttcttagaca acaagggtag aatgtatctg ctgatgtatt taaaaagttc	420
aaagatgaga ggggcaactt caaggcaagc ttaagtactg atgccagggg gttgctaagc	480
ttgtatgaag ctgcatttct cagcatacga ggagatgata tcttagatga agccataact	540
ttcacaagag agcagcttaa gtcttctatg acccatgttg atgccctct tgccaaacaa	600

-continued

---

```

atagcccatg ccttagaggt accagcgcac aagcgcatac aaagactaga gaacattcgc 660
tacctcacia tctaccaaga agaaaagga aggaatgatg tgggtgctga gcttgccaag 720
ttggatttca atatcttaca acaattgcat aagaaagaac tgagagacct tacaaagtgg 780
tggaaggaca cagacgttgc aggaaagta cctttcatca gagataggtt ggtggaatgc 840
tattattgga tcttgggtgt gtattatgag ccagaatact ccagagctag aatTTTTTct 900
acccaaatga caatcatggt ctcagttggt gatgacatat atgacgtata tgctactgaa 960
gatgagctcc aactattcac tgatgcaatc tataggtggg atctggaggg cctagatcaa 1020
ctcccacagt tcttgaaga ctgttttctt gtactctatg acaccgtcaa ggaattagaa 1080
gacgaactag aaccggaagg aaaatcctat cgtggatact atgtaaagga tgcgatgaa 1140
gttttggtca gagattactt cgttgagcac aaatggtata acagaaacat agtgccaagt 1200
gtagaagaat atctccgtgt ttcttgcac agtggtgcag tccatattggc taacgtccat 1260
tgctgtgctg ggtggggaga tgtaatgagc aaagaggcat tcgaatggtt gaagagtgaa 1320
ccaaaggttg taatggatgc atcactaatt ggccgactgc tcgatgacat gcagtccacc 1380
gagtttgagc aaaagagagg ccatgttgca tcggctgtcc aatgttcat gaatgagtat 1440
ggagtgactt acaagaagc gtgtgaaaag ctgcatgaaa tggctgcoct tgcattgaaa 1500
gacgtaaac aggcttgcoct taaaccaact gttttccctc tccctgtatt tatgctgca 1560
atcaaccttg cgcgagtggc tgaagtcatc taccttcgtg gagatgggta tactcattca 1620
ggaggagaga ctaaagaaaa tatcacgttg atgcttgcac atccaatctc tgtgtgaa 1677

```

```

<210> SEQ ID NO 11
<211> LENGTH: 558
<212> TYPE: PRT
<213> ORGANISM: Drimys winteri

```

```

<400> SEQUENCE: 11

```

```

Met Asp Leu Ser Thr Ser Pro Val Leu Ser Ser Ser Pro Leu Pro Val
1          5          10          15
Glu Asp Gly Lys Asn Pro Ala Val Arg Arg Ser Ala Gly Phe His Pro
20         25         30
Ser Ile Trp Gly Asp His Phe Leu Ser Tyr Thr Glu Asp His Lys Lys
35         40         45
Leu Asp Ala Trp Ser Glu Arg Thr Gln Val Leu Lys Glu Glu Val Arg
50         55         60
Arg Ile Leu Ile Asn Ala Lys Gly Ser Leu Glu Glu Leu Asp Leu Leu
65         70         75         80
Asp Ala Ile Gln Arg Leu Gly Val Lys Tyr His Phe Glu Lys Glu Ile
85         90         95
Glu Glu Ala Leu His His Ile Tyr Val Ala Glu Thr His Val Ser Thr
100        105        110
Asp Asp Leu Tyr Ser Val Ser Leu Arg Phe Arg Leu Leu Arg Gln Gln
115        120        125
Gly Tyr Asn Val Ser Ala Asp Val Phe Lys Lys Phe Lys Asp Glu Arg
130        135        140
Gly Asn Phe Lys Ala Ser Leu Ser Thr Asp Ala Arg Gly Leu Leu Ser
145        150        155        160
Leu Tyr Glu Ala Ala Phe Leu Ser Ile Arg Gly Asp Asp Ile Leu Asp
165        170        175
Glu Ala Ile Thr Phe Thr Arg Glu Gln Leu Lys Ser Ser Met Thr His
180        185        190

```

-continued

---

Val Asp Ala Pro Leu Ala Lys Gln Ile Ala His Ala Leu Glu Val Pro  
 195 200 205  
 Ala His Lys Arg Ile Gln Arg Leu Glu Asn Ile Arg Tyr Leu Thr Ile  
 210 215 220  
 Tyr Gln Glu Glu Lys Gly Arg Asn Asp Val Leu Leu Glu Leu Ala Lys  
 225 230 235 240  
 Leu Asp Phe Asn Ile Leu Gln Gln Leu His Lys Lys Glu Leu Arg Asp  
 245 250 255  
 Leu Thr Lys Trp Trp Lys Asp Thr Asp Val Ala Gly Lys Leu Pro Phe  
 260 265 270  
 Ile Arg Asp Arg Leu Val Glu Cys Tyr Tyr Trp Ile Leu Gly Val Tyr  
 275 280 285  
 Tyr Glu Pro Glu Tyr Ser Arg Ala Arg Ile Phe Ser Thr Lys Met Thr  
 290 295 300  
 Ile Met Val Ser Val Val Asp Asp Ile Tyr Asp Val Tyr Ala Thr Glu  
 305 310 315 320  
 Asp Glu Leu Gln Leu Phe Thr Asp Ala Ile Tyr Arg Trp Asp Leu Glu  
 325 330 335  
 Gly Leu Asp Gln Leu Pro Gln Phe Leu Lys Asp Cys Phe Leu Val Leu  
 340 345 350  
 Tyr Asp Thr Val Lys Glu Leu Glu Asp Glu Leu Glu Pro Glu Gly Lys  
 355 360 365  
 Ser Tyr Arg Gly Tyr Tyr Val Lys Asp Ala Met Lys Val Leu Ala Arg  
 370 375 380  
 Asp Tyr Phe Val Glu His Lys Trp Tyr Asn Arg Asn Ile Val Pro Ser  
 385 390 395 400  
 Val Glu Glu Tyr Leu Arg Val Ser Cys Ile Ser Val Ala Val His Met  
 405 410 415  
 Ala Asn Val His Cys Cys Ala Gly Met Gly Asp Val Met Ser Lys Glu  
 420 425 430  
 Ala Phe Glu Trp Leu Lys Ser Glu Pro Lys Val Val Met Asp Ala Ser  
 435 440 445  
 Leu Ile Gly Arg Leu Leu Asp Asp Met Gln Ser Thr Glu Phe Glu Gln  
 450 455 460  
 Lys Arg Gly His Val Ala Ser Ala Val Gln Cys Tyr Met Asn Glu Tyr  
 465 470 475 480  
 Gly Val Thr Tyr Lys Glu Ala Cys Glu Lys Leu His Glu Met Ala Ala  
 485 490 495  
 Leu Ala Trp Lys Asp Val Asn Gln Ala Cys Leu Lys Pro Thr Val Phe  
 500 505 510  
 Pro Leu Pro Val Phe Met Pro Ala Ile Asn Leu Ala Arg Val Ala Glu  
 515 520 525  
 Val Ile Tyr Leu Arg Gly Asp Gly Tyr Thr His Ser Gly Gly Glu Thr  
 530 535 540  
 Lys Glu Asn Ile Thr Leu Met Leu Val Asn Pro Ile Ser Val  
 545 550 555

&lt;210&gt; SEQ ID NO 12

&lt;211&gt; LENGTH: 1677

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Artificial Sequence

&lt;220&gt; FEATURE:

&lt;223&gt; OTHER INFORMATION: Codon optimized DNA sequence of SCH51-998-28

&lt;400&gt; SEQUENCE: 12

-continued

---

atggatctga gcaccagtcg ggttctgagc agctcaccgc tgcgggttga agatggtaaa	60
aatccggcag ttcgctgtag cgcaggtttt catccgagca tttgggggtga tcattttctg	120
agctataccg aggatcacaa aaaactggat gcatggtcag aacgtaccca ggttctgaaa	180
gaagaagtgc gtcgtattct gattaatgca aaaggtagcc tggaagaact ggatctgctg	240
gatgcaatc agcgtctggg tgttaaatat cactttgaga aagaaatcga agaagcctg	300
catcatattt atgttgacaga aacctatgtg tcaaccgatg atctgtatag cgtagcctg	360
cgttttcgtc tgctgcgca gcagggttat aatgttagcg cagatgtgtt caaaaaattc	420
aaagatgaac gcggttaactt caaagcaagc ctgagcaccg atgcacgtgg tctgctgagc	480
ctgtatgaag cagcatttct gagcattcgt ggtgatgata ttctggatga agcaattacc	540
tttaccctg aacagctgaa aagcagcatg acctatgttg atgcaccgct ggcaaaaaca	600
attgcacatg cactggaagt tccggcacat aaacgtattc agcgcctgga aaatattcgc	660
tatctgacca tttaccaaga agaaaaaggt cgtaacgatg ttctgctgga actggccaaa	720
ctggatttta acattctgca gcagctgcat aaaaaagaac tgctgatct gaccaaatgg	780
tgaaagata ccgatgtgac aggtaaactg ccgtttattc gtgatcgtct ggttgaatgc	840
tattattgga ttctgggctg ttattatgag ccggaatata gccgtgcacg tatttttagc	900
acaaaaatga ccattatggt tagcgtgggtg gatgacatct atgatgttta tgcaaccgaa	960
gatgaaactgc agctgtttac cgatgcaatt taccgttggg atctggaagg tctggatcag	1020
ctgccgcagt tcctgaaaga ttgttttctg gttctgtatg ataccgtgaa agaactggaa	1080
gatgagctgg aaccggaagg taaaagctat cgtgggttatt atgttaaaga tgccatgaaa	1140
gttctggcac gcgattattt tgttgagcac aaatggtata accgcaatat tgttccgagc	1200
gtggaagaat atctgcgtgt tagctgtatt agcgttgacg ttcacatggc aaatgttcat	1260
tgtttgacg gtatgggtga tgtgatgagc aaagaagcat ttgaatggct gaaaagtgaa	1320
ccgaaagtgt ttatggatgc cagcctgatt ggtcgcctgc tggacgatat gcagagcacc	1380
gaatttgaac agaaacgtgg tcattgttca agcgcagttc agtgttatat gaatgaatat	1440
ggcgtgacct ataagagggc atgcgaaaaa ctgcatgaaa tggcagcact ggcatggaaa	1500
gatgttaatc aggcattgct gaaaccgacc gtttttccgc tgctgtttt tatgcctgca	1560
attaatctgg cacgtgttgc cgaagtattt tacctgcctg gggatgggta taccatagc	1620
gggtgtgaaa ccaagaaaaa cattaccctg atgctgggta atccgattag cgtttaa	1677

&lt;210&gt; SEQ ID NO 13

&lt;211&gt; LENGTH: 1680

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Drimys lanceolata

&lt;400&gt; SEQUENCE: 13

atggatgttc taattccctc ccctgtggct tccactctcc ctctgccga agatggaaac	60
ttggacgtcg ttcgcagatc cgcgggtttt catccgacgg tctggggcga tcacttctc	120
gcttactcgc ccgatccaac caaaatagat gcttggacta aaagagtga agagctgaag	180
caagaagtga agaggattct aagcaatgtg aaagggtcac tggaagagct gaacttgctt	240
gatgctatcc aacaccttgg gattgggttat cattttgaga aagagattga tgatgcttta	300
caactaatct ttgattccca tattgatgct tttcctactg atgatctata tgtggctgcc	360
ctccgattta gcctactaag gcgacaaggg cactgtgttt cttcagatgt attcaaaaaa	420

-continued

---

```

ttcaaagatg agcaggggaa tttcaaggca gagctgagca ccgatgcgaa aggtttgctg 480
agtctctatg acgcggcgta tctcagtgtg agaggggaag atatattgga tgaggccatt 540
cctttcacta gggagcacct taggacttgt attagccatg tagattctca tttggcagca 600
aaaattgagc attctctaga gcttccoctg catcatcgca taccaaggct agagaacagg 660
cactacatct cagtgtacga aggagagaag gaaaggaatg aagttgtact agagcttgcc 720
aaattagatt tcaatctgat tcaaatcttg caccaaagag agctgagggg catcacaacg 780
tggtggaatg agattgacct cgcagcaaag ctaccattta ttagggatag gttggtggag 840
tgctactatt ggatcatggg tgtctatttt gaaccaatat tctcaagggc tagagttttt 900
tcgacaaaaa tgacaatttt ggtctcagtt gtcgacgaca tatatgatgt ctacgctaca 960
gaggatgagc tccaactttt cactgacgca atctataggt gggatgccga ggacattgag 1020
cagcttccac agtacttgaa agattctttt cttgtactct ataacaccgt gaaggactta 1080
gaagaggagc tgaaccaga aggaaactca tategtggag actatgtaa agatgcgatg 1140
aaggttttgg caagagatta ctttgtggag cacaaatggt ataacagaaa aattgtaccg 1200
tcagtagagg actacctacg aatttcttgc attagtgttg ccgttcatat ggctacagtt 1260
cattgttgtg ctgggatgga tgaattgca accaaagagg cattcgaatg gttgaagacc 1320
gaacctaaac ttgttataga tgcatcactg attgggcgtc tcctcgatga catgcagtcc 1380
acctcgtttg agcaacagag aggtcatgtg tcatcggcgg tacagtgtta catgatccaa 1440
tatggcgtat cacacgaaga agcgtgtgag aagttgacag aaatggctgc aattgcatgg 1500
aaagatgtaa accaagcatg ccttaggcc actgttttcc caatgcctat tcttctgcct 1560
tcaatcaacc ttgcacgtgt ggcagaagtc atctacctgc gcggagatgg atatacat 1620
gctggtgggtg agacaaaaa acatatcacg gccatgcttg ttgaaccaat ccaagtctga 1680

```

&lt;210&gt; SEQ ID NO 14

&lt;211&gt; LENGTH: 559

&lt;212&gt; TYPE: PRT

<213> ORGANISM: *Drimys lanceolata*

&lt;400&gt; SEQUENCE: 14

```

Met Asp Val Leu Ile Pro Ser Pro Val Ala Ser Thr Leu Pro Leu Pro
1           5           10           15

Glu Asp Gly Asn Leu Asp Val Val Arg Arg Ser Ala Gly Phe His Pro
20          25          30

Thr Val Trp Gly Asp His Phe Leu Ala Tyr Ser Pro Asp Pro Thr Lys
35          40          45

Ile Asp Ala Trp Thr Lys Arg Val Glu Glu Leu Lys Gln Glu Val Lys
50          55          60

Arg Ile Leu Ser Asn Val Lys Gly Ser Leu Glu Glu Leu Asn Leu Leu
65          70          75          80

Asp Ala Ile Gln His Leu Gly Ile Gly Tyr His Phe Glu Lys Glu Ile
85          90          95

Asp Asp Ala Leu Gln Leu Ile Phe Asp Ser His Ile Asp Ala Phe Pro
100         105         110

Thr Asp Asp Leu Tyr Val Ala Ala Leu Arg Phe Ser Leu Leu Arg Arg
115         120         125

Gln Gly His Cys Val Ser Ser Asp Val Phe Lys Lys Phe Lys Asp Glu
130         135         140

Gln Gly Asn Phe Lys Ala Glu Leu Ser Thr Asp Ala Lys Gly Leu Leu
145         150         155         160

```



-continued

---

```

<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Codon optimized DNA sequence of SCH51-13163-6

<400> SEQUENCE: 15
atggatgttc tgattccgag tccggttgca agcacccctgc cgctgccgga agatggtaat    60
ctggatgttg ttcgtcgtag cgcaggtttt catccgaccg tttgggggtga tcattttctg    120
gcatatagtc cggatccgac caaaattgat gcatggacca aacgtgttga ggaactgaaa    180
caagaagtga aacgtattct gagcaatgtg aaaggtagcc tggagaact gaatctgctg    240
gatgcaattc agcatctggg tattggttat cacttcgaga aagaaattga tgatgcactg    300
cagctgatct ttgatagcca tattgatgcc tttccgaccg atgatctgta tgttgacgca    360
ctgcgtttta gcctgctgcg tcgtcagggt cattgtgta gcagtgatgt tttcaaaaaa    420
ttcaagaagc agcagggcaa ctttaaagca gaactgagca cccgatgaaa aggtctgctg    480
agcctgtatg atgccgata tctgagcgtt cgtggtgaag atattctgga tgaagcaatt    540
ccggttaccc gtgaacatct gcgtacctgt attagccatg tggatagcca tctggcagca    600
aaaattgaac atagtctgga actgcctctg catcatcgta ttccgctctt ggaaaatcgt    660
cactatatta gcgtttatga aggcgaaaaa gaacgcaatg aagttgtgct ggaactggca    720
aaactggatt ttaacctgat tcagattctg catcagcgtg aactgcgtga tattaccacc    780
tggtggaatg aaattgacct ggcagccaaa ctgccgttta ttcgtgatcg tctggttgaa    840
tgctattatt ggattatggg cgtgtathtt gaaccgattt ttagccctgc acgtgtgttt    900
agcaccaaaa tgaccattct ggttagcgtg gtggatgata tctatgatgt ttatgcaacc    960
gaagatgagc tgcaactggt taccgatgcc atttatcgtt gggatgcaga agatattgaa   1020
cagctgcctc agtatctgaa agatagcttt ctggttctgt acaacaccgt gaaagatctg   1080
gaagaagaac tgaaccggga aggtaatagc tatcgtgggtg attatgttaa agacgccatg   1140
aaagttctgg cacgcgatta ttttgttgag cacaaatggt ataaccgcaa aattgttccg   1200
agcgtggaag attatctgcg tattagctgc attagcgttg cagttcacat ggcaaccgtt   1260
cattgttgty caggtatgga tgaattgca accaaagaag catttgagtg gctgaaaacc   1320
gaaccgaaac tggttattga tgcaagcctg attggtcgtc tgctggaaga tatgcagtca   1380
accagctttg aacagcagcg tggtcatggt agcagcgcag ttcagtgtta tatgattcag   1440
tatggtgtta gccatgaaga agcatgcgaa aaactgaccg aaatggcagc aattgcatgg   1500
aaagatgtta atcaggcatg tctgcgtccg accgtgtttc ctatgccgat tctgctgccg   1560
agcattaatc tggcacgtgt tgccgaagtt atctatctgc gtggtgatgg ttatacccat   1620
gccggtggty aaacccaaaa acatattacc gcaatgctgg tagaaccgat tcaggtttaa   1680

```

---

What is claimed is:

1. A method of producing drimenol comprising:
  - i) contacting farnesyl diphosphate (FPP) with a polypeptide having drimenol synthase activity and comprising an amino acid sequence having at least 85% sequence identity to a sequence selected from the group consisting of SEQ ID NO: 2 and SEQ ID NO: 5 to produce the drimenol; and
  - ii) optionally isolating the drimenol.
2. The method as recited in claim 1 wherein the drimenol is isolated.
3. The method as recited in claim 1 wherein the drimenol is produced with at least 30% selectivity.

4. The method as recited in claim 1 comprising the steps of transforming a host cell or non-human organism with a nucleic acid encoding a polypeptide comprising an amino acid sequence having at least 85% sequence identity of a sequence selected from the group consisting of SEQ ID NO: 2 and SEQ ID NO: 5 and culturing the host cell or organism under conditions that allow for the production of the polypeptide.
5. The method recited in claim 4 wherein the cell is a prokaryotic cell or a eukaryotic cell.
6. The method as recited in claim 4 wherein the cell is a bacterial cell.

7. The method as recited in claim 5 wherein the eukaryotic cell is a yeast cell or a plant cell.

8. The method as recited in claim 2 wherein the drimenol is produced with at least 30% selectivity.

\* \* \* \* \*