

(43) International Publication Date
25 November 2010 (25.11.2010)(10) International Publication Number
WO 2010/135674 A2(51) International Patent Classification:
C12P 7/16 (2006.01)(21) International Application Number:
PCT/US2010/035820(22) International Filing Date:
21 May 2010 (21.05.2010)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/180,757 22 May 2009 (22.05.2009) US(71) Applicant (for all designated States except US): **BOARD OF TRUSTEES OF MICHIGAN STATE UNIVERSITY** [US/US]; 450 Administration Building, East Lansing, Michigan 48824-1046 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **GRAY, Dennis** [US/US]; 1628 Indiana Avenue, Lansing, Michigan 48906 (US). **SHARKEY, Thomas** [US/US]; 1210 Prescott Drive, East Lansing, Michigan 48823 (US).(74) Agents: **PERDOK SHONKA, Monique M.** et al.; Schwegman, Lundberg & Woessner, P.A., P.O. Box 2938, Minneapolis, Minnesota 55402 (US).

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

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

Published:

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

(54) Title: METHYL BUTENOL SYNTHASE

Effect of Boiling and EDTA on Enzyme Activity

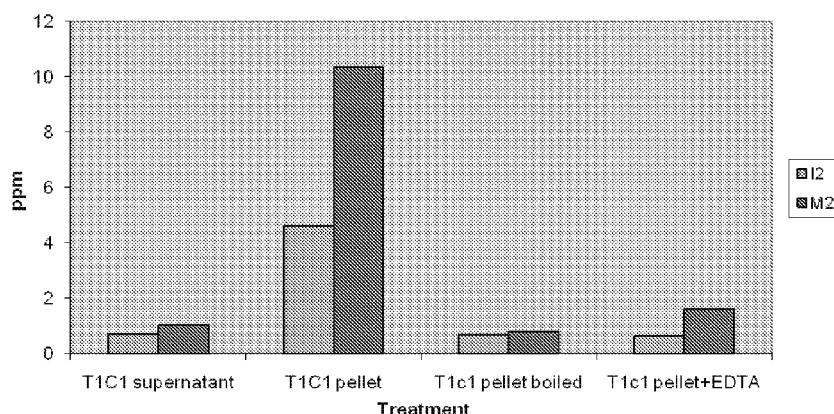


FIGURE 4

(57) Abstract: The present invention provides novels genes encoding methyl butenol (MBO) synthase, methy butenol synthases and their use in methyl butenol production.

METHYL BUTENOL SYNTHASE

Related Application

This application claims priority from U.S. Provisional Application Serial No.
5 61/180,757 filed May 22, 2009, which application is herein incorporated by reference.

Statement of Government Rights

This work was funded by NSF Grant No. IOS 0830225. The government has
certain rights to this invention.

10

Background of the Invention

In 2007, there were 1.8 million alternative fuel vehicles sold in the United States,
indicating an increasing popularity of alternative fuels. There is growing perceived
economic and political need for the development of alternative fuel sources due to general
environmental, economic, and geopolitical concerns of sustainability.

15

The major environmental concern is that most of the observed increase in globally
averaged temperatures since the mid-20th century is due to the observed increase in
greenhouse gas concentrations. Since burning fossil fuels is known to increase
greenhouse gas concentrations in the atmosphere, it is believed that they are a likely
contributor to global warming.

20

Although fossil fuels have become the dominant energy resource for the modern
world, alcohol has been used as a fuel throughout history. The first four aliphatic alcohols
(methanol, ethanol, propanol, and butanol) are of interest as fuels because they can be
synthesized and they have characteristics which allow them to be used in current engines.
Biobutanol has an energy density that is closer to gasoline than the other alcohols;

25

however, this advantage is outweighed by disadvantages (compared to ethanol and
methanol) concerning, for example, production.

Summary of the Invention

The present invention provides for the first time genes encoding methyl butenol
30 (MBO) synthase and its use in methyl butenol production. Thus, one embodiment
provides an isolated and purified methyl butenol (MBO) synthase nucleic acid molecule,
wherein the MBO synthase nucleic acid molecule comprises any one of SEQ ID NOs:1-
17 or a nucleic acid molecule having at least about 80% sequence identity thereof.

Another embodiment provides an expression vector comprising a MBO synthase nucleic acid molecule. One embodiment provides a prokaryotic or eukaryotic host cell transformed with a MBO synthase nucleic acid molecule (e.g., in an expression vector). In one embodiment the transformed host cells express the exogenous MBO synthase (mRNA or protein). In one embodiment, the host cells express MBO synthase and yield MBO.

One embodiment provides an isolated and purified methyl butenol (MBO) synthase polypeptide, wherein the MBO synthase polypeptide comprises any one of SEQ ID NOs:18-34 or a bioactive polypeptide having at least about 80% sequence identity thereof.

Another embodiment provides a method to produce methyl butenol comprising transforming a host cell with a nucleic acid molecule coding for a methyl butenol synthase and expressing said molecule in a host cell so as to yield methyl butenol. In one embodiment, the host cells are fermentative organisms, such as a bacteria, cyanobacteria or yeast (e.g., a bacteria or yeast that can break down sugar into alcohol) or eukaryotic micro algae. In one embodiment the fermentative organism is *Saccharomyces cerevisiae*, *Klebsiella oxytoca*, *Synechococcus sp.*, *Synechocystis sp.*, *Anabaena sp.*, *Chlorella sp.*, *Scenedesmus sp.*, *Bracteococcus sp.*, *Chlamydomonas sp.*, C5- or C6-fermentative organisms (including *Zymomonas* (e.g., *Zymomonas mobilis*)) or a combination thereof. In one embodiment, the transformed host cell is used in fermentation with a carbohydrate to yield MBO (similar to bioethanol production).

In one embodiment, the MBO synthase (in a purified or unpurified form) is used in combination with a carbohydrate to yield MBO. In one embodiment, the fermentative organism is bacteria, cyanobacteria or yeast (e.g., a bacteria or yeast that can break down sugar into alcohol) or eukaryotic micro algae. In one embodiment the fermentative organism is *Saccharomyces cerevisiae*, *Klebsiella oxytoca*, *Synechococcus sp.*, *Synechocystis sp.*, *Anabaena sp.*, *Chlorella sp.*, *Scenedesmus sp.*, *Bracteococcus sp.*, *Chlamydomonas sp.*, C5- or C6-fermentative organisms (including *Zymomonas* (e.g., *Zymomonas mobilis*)) or combinations thereof.

Brief Description of the Drawings

Figure 1 provides nucleic acid molecules coding for a methyl butenol synthase (SEQ ID NOs:1-17).

Figure 2 provides the sequence of several methyl butenol synthases. (SEQ ID NO:18-34).

Figures 3A-B demonstrate the methylbutenol and isoprene production from soluble and insoluble extract fractions.

5 Figure 4 demonstrates the effect of boiling and EDTA on enzyme activity.

Figure 5 demonstrates the effect of co-factor replacement on MBO synthase activity.

Detailed Description of the Invention

10 The present invention provides for the first time genes encoding methyl butenol (MBO) synthase and its use in methyl butenol production. In particular, the methyl butenol synthase gene was isolated from several species of pine. The gene is useful for the enzymatic production of methyl butenol, which can be used as an alternative fuel (e.g., a biofuel or gasoline replacement). Thus, the methods provided herein provide a
15 practical, non-polluting technology for producing methyl butenol using methyl butenol synthase proteins and genes. Exemplary cDNA sequences coding for methyl butenol synthases are provided in SEQ ID NOs:1-17, while exemplary amino acids sequences are provided in SEQ ID NOs:18-34 (see Figures 1 and 2).

Definitions

20 Certain terms used in the specification are defined and presented as follows:

2-methyl-3-buten-2-ol (molecular formula of $C_5H_{10}O$; also known as methyl butenol or methylbutenol (MBO)) is a natural 5-carbon alcohol that is produced and emitted from the foliage by several species of pine, e.g., species in the genus *Pinus*, with the following structure



25 **2-methyl-3-buten-2-ol**

Biosynthesis of MBO occurs in the chloroplast through the action of the enzyme MBO synthase, which uses DMADP derived from the MEP pathway as a substrate.

“Associated with/operably linked” refers to two DNA sequences that are related physically or functionally. For example, a promoter or regulatory DNA sequence is said
30 to be “associated with” a DNA sequence that codes for an RNA or a protein if the two

sequences are operably linked, or situated such that the regulator DNA sequence will affect the expression level of the coding or structural DNA sequence.

“Coding DNA sequence:” refers to a DNA sequence that is translated in an organism to produce a protein. “Expression” of a sequence includes the production of mRNA and/or protein.

“Isolated,” in the context of the present invention, an isolated nucleic acid molecule or an isolated polypeptide is a nucleic acid molecule or polypeptide that, by the hand of man, exists apart from its native environment and is therefore not a product of nature. Thus, the term isolated refers to a molecule (e.g., nucleic acid or protein) which is not associated with one or more nucleic acid molecules, proteins or one or more cellular components that are associated with the nucleic acid molecule or protein *in vivo*. An isolated nucleic acid molecule or an isolated protein may exist in a purified form (e.g., ranges of purity in samples comprising isolated nucleic acid molecules or an isolated protein are 50-55%, 55-60%, 60-65% and 60-70%; ranges of purity also include 70-75%, 75-80%, 80-85%; 85-90%, 90-95%, and 95-100%; however, samples with lower purity can also be useful, such as about <25%, 25-30%, 30-35%, 35-40%, 40-45% and 45-50%.) or may exist in a non-native environment such as, for example, a transgenic host cell.

A “cell” or “host cell” is a prokaryotic or eukaryotic cell.

Alternative fuels, also known as non-conventional fuels, are any materials or substances that can be used as fuels, other than conventional fuels. Conventional fuels include: fossil fuels (petroleum (oil), coal, propane, and natural gas), and nuclear materials such as uranium. Some well known alternative fuels include biodiesel, bioalcohol (methanol, ethanol, butanol), chemically stored electricity (batteries and fuel cells), hydrogen, non-fossil methane, non-fossil natural gas, vegetable oil and other biomass sources. For example, one alternative is alcohol fuel.

Alcohol fuels are usually of biological rather than petroleum sources. When obtained from biological sources, they are known as bioalcohols (e.g. bioethanol). As described herein 2-methyl-3-butene-2-ol (methyl butenol) can be manufactured in organisms carrying a synthase gene of the invention or by fermentative organisms in contact with a synthase protein of the invention and a carbohydrate source. It is a sustainable energy resource that can provide a more environmentally and economically friendly alternative to fossil fuels such as diesel and gasoline. It can be combined with gasoline at different percentages, or can be used in its pure form. Generally, methyl butenol is more advantageous than ethanol as a fuel due to its increased density and

relatively little or no water absorption as with ethanol. Unlike butanol, the presence of a branched carbon chain increases its octane rating and the presence of a double carbon-carbon bond can reduce toxicity.

5 The terms “comprises,” “comprising,” and the like can have the meaning ascribed to them in U.S. Patent Law and can mean “includes,” “including” and the like. As used herein, “including” or “includes” or the like means including, without limitation.

Methyl Butenol Synthase Genes/Proteins

10 Provided herein are methyl butenol synthase genes including those of SEQ ID NOs: 1-17 or a nucleic acid molecule which comprises a sequence that has at least about 50%, at least about 60%, at least about 70%, about 71%, about 72%, about 73%, about 74%, about 75%, about 76%, about 77%, about 78%, or about 79%, or at least about 80%, about 81%, about 82%, about 83%, about 84%, about 85%, about 86%, about 87%, about 88%, or about 89%, or at least about 90%, about 91%, about 92%, about 93%, or about 94%, or at least about 95%, about 96%, about 97%, about 98%, about 99%, or about 100% sequence identity compared to any one of SEQ ID NOs:1-17 using one of the alignment programs available in the art using standard parameters.

Also provided herein are methyl butenol synthase proteins including those of SEQ ID NOs:18-34 or a polypeptide which comprises a sequence that has at least about 50%, at least about 60%, at least about 70%, about 71%, about 72%, about 73%, about 74%, about 75%, about 76%, about 77%, about 78%, or about 79%, or at least about 80%, about 81%, about 82%, about 83%, about 84%, about 85%, about 86%, about 87%, about 88%, or about 89%, or at least about 90%, about 91%, about 92%, about 93%, or about 94%, or at least about 95%, about 96%, about 97%, about 98%, about 99%, or about 100% sequence identity compared to any one of SEQ ID NOs:18-34 using one of the alignment programs available in the art using standard parameters. In one embodiment, the differences in sequence are due to conservative amino acid changes. In one embodiment, the functional properties of the enzyme are improved through the use of molecular evolution and design studies.

25 Methods of alignment of sequences for comparison are available in the art. Thus, the determination of percent identity between any two sequences can be accomplished using a mathematical algorithm. Computer implementations of these mathematical algorithms can be utilized for comparison of sequences to determine sequence identity. Such implementations include, but are not limited to: CLUSTAL in the PC/Gene program (available from Intelligenetics, Mountain View, California); the ALIGN program

(Version 2.0) and GAP, BESTFIT, BLAST, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Version 8 (available from Genetics Computer Group (GCG), 575 Science Drive, Madison, Wisconsin, USA). Alignments using these programs can be performed using the default parameters.

5 Expression of Methyl butenol Synthase and Production of Methyl butenol (MBO)

Methyl butenol synthase can be introduced into any number of organisms to cause them to produce methyl butenol synthase and/or methyl butenol. For example, the genes can be cloned into appropriate expression vectors and expressed in bacteria or yeast. The creation of expression vectors comprising the nucleic acid molecules of the invention and
10 bacterial and yeast transformation techniques are described in detail in the literature and available to an art worker. Typically an expression vector contains (1) prokaryotic DNA elements coding for a bacterial replication origin and an antibiotic resistance gene to provide for the amplification and selection of the expression vector in a bacterial host or DNA elements to allow expression in a eukaryotic host cell (e.g., yeast); (2) regulatory
15 elements that control initiation of transcription such as a promoter; and (3) DNA elements that control the processing of transcripts such as introns, transcription termination/polyadenylation sequence.

Methods to introduce a nucleic acid molecule into a vector are well known in the art (Sambrook et al., 1989). As an example, a vector into which the nucleic acid segment is to
20 be inserted is treated with one or more restriction enzymes (restriction endonuclease) to produce a linearized vector having a blunt end, a "sticky" end with a 5' or a 3' overhang, or any combination thereof. The vector may also be treated with a restriction enzyme and subsequently treated with another modifying enzyme, such as a polymerase, an exonuclease, a phosphatase or a kinase, to create a linearized vector that has characteristics
25 useful for ligation of a nucleic acid molecule into the vector. The nucleic acid molecule that is to be inserted into the vector is treated with one or more restriction enzymes to create a linearized segment having a blunt end, a "sticky" end with a 5' or a 3' overhang, or any combination thereof. The nucleic acid molecule may also be treated with a restriction enzyme and subsequently treated with another DNA modifying enzyme. Such DNA
30 modifying enzymes include, but are not limited to, polymerase, exonuclease, phosphatase or a kinase, to create a polynucleic acid segment that has characteristics useful for ligation of a nucleic acid molecule into the vector.

The treated vector and nucleic acid molecule are then ligated together to form a construct containing a nucleic acid segment according to methods known in the art

5 An example of the preparation of an expression vector and expression of MBO
synthase in an organism is described in the examples below. However, any expression
vector and organism combination for expression available to an art worker may be
acceptable for use in the methods of the invention. It will be clear to one of ordinary skill
in the art which vector should be used depending on which cell type is used for a host
10 cell.

15 After preparation of the expression vector and introduction into an appropriate host cell, the cells can be grown in culture to produce methyl butenol synthase and/or methyl butenol, which can be collected (and optionally distilled or condensed).

20 C_nH_{2n}O_n (n is between 3 and 7), including a monosaccharide (e.g., glucose, dextrose or fructose) or a disaccharide (e.g., sucrose), starches (e.g., a carbohydrate consisting of a large number of glucose units joined together by glycosidic bonds, such as amylase or amylopectin) and/or cellulose (e.g., an organic compound with the formula (C₆H₁₀O₅)_n, a polysaccharide consisting of a linear chain of several hundred to over ten thousand
25 β(1→4) linked D-glucose units)) and a methyl butenol (MBO) synthase of the invention can be combined to yield MBO, which can be purified further (e.g., distilled or condensed from the head space).

30 Examples

7

Example 1: Preparation of Methyl Butanol Synthase vectors and Assay for ActivityMaterials and MethodsRNA extraction

Total RNA was extracted from frozen pine needle tissue from *Pinus sabiniana* using a modified CTAB procedure adapted from (Chang et al. (1993) *Plant Molecular Biology Reporter*, 11(2):113-116.). Needle samples (250mg) were ground under liquid nitrogen, and extracted for 45 minutes at 70°C in 800ul volumes of extraction buffer containing 2% CTAB, 2% PVP-40, 100mM Tris pH8.0, 25mM EDTA, 2.0M NaCl, and 0.5g/L Spermidine with 4% β-mercaptoethanol, 4% PVPP, and an additional 4% PVP40 was added just before use. Following incubation at 70°C for 45 minutes, samples were extracted twice with 500μl volumes of 24:1 (v:v) chloroform:isoamyl alcohol solution. RNA was then precipitated by adding ¼ volume of 10M LiCl to the aqueous phase, gently mixing, and storing the mixture overnight at 4°C. Following centrifugation to collect the RNA pellet, the supernatant was discarded and the pellet was re-suspended in 500μl of SSE Buffer containing 1.0M NaCl, 0.5% SDS, 10mM Tris pH8.0, and 1mM EDTA. This solution was then extracted once with equal volumes of phenol, once with phenol:chloroform: IAA (24:24:1 v/v), and once with chloroform:IAA (24:1 v/v). To precipitate the RNA pellet, 2 volumes of cold (-20°C) ethanol were added to the supernatant followed by incubation for 2 hours at -20°C. Following centrifugation to collect the pellet, the RNA was air dried, re-suspended in distilled water, and stored at -20C until further use.

cDNA synthesis

First strand cDNA synthesis was performed using M-MLV reverse transcriptase obtained from Invitrogen (Carlsbad, CA) following the manufacturers instructions. Total RNA (2μg), and 1μg oligo-dT(17) primer were denatured by incubation at 70°C for 5 minutes, and then placed on ice for 5 min to anneal the oligo-dT primer to the mRNA polyA tail. Reagents were added to this mixture to achieve a 50μL volume containing 1x concentration M-MLV buffer, 0.5mM dNTPs, 0.1M DTT and 2.5 units of RNasin (Promega, Madison WI). The first strand synthesis reactions were performed by incubating for 60 min at 37°C, followed by incubating for 10 min at 42°C, and heat inactivation of the enzyme by incubating at 70°C for 15 min. Reactions were cooled on ice and stored at -20°C until further use.

Primer design.

Initially, degenerate PCR primers were designed to correspond to highly conserved regions of known pine monoterpene synthases. Sequencing of the PCR amplified products revealed 3 distinct sequences which showed high similarity to known pine monoterpene synthases. One sequence contained a phenylalanine residue 7 amino acids upstream of the highly conserved DDXXD motif of terpene synthases. Since this position corresponds to the location of a phenylalanine residue in isoprene synthases and is believed to limit the size of the substrate that can fit into the enzymes active site, this phenylalanine containing sequence was selected for further exploration. To obtain the 5' and 3' ends of this sequence 5' and 3' rapid amplification of cDNA ends (RACE) was employed. For these procedures gene specific forward and reverse primers were designed from the previously selected partial MBO synthase sequence.

PCR Reactions.

Initial PCR reactions were run using 1.25 units of Gotaq DNA polymerase in 30µl reaction volumes containing 1x Gotaq buffer, 20µM dNTPs, 1µM concentrations of each primer, and 0.2µg template cDNA. In preparation for cloning, full length cDNAs were amplified using either Pfu turbo (Stratagene, La Jolla CA) or Platinum Pfx (Invitrogen, Carlsbad CA) DNA polymerases according to the manufacturer instructions. To amplify full length cDNAs PCR reactions were run using an initial denaturation of 2 min at 95°C, followed by 35 repetitions of denaturation at 95°C for 30 sec, annealing at 52°C for 45 sec, and extension at 72°C for 3 min, followed by a final extension at 72°C for 9 min.

Rapid amplification of cDNA ends (RACE)

Rapid amplification of cDNA ends (RACE) procedures were employed to amplify the 5' and 3' ends of the partial MBO sequence obtained using degenerate primers.

RACE reactions were run using 1.25 units of Gotaq DNA polymerase in 30µl reaction volumes containing 1x Gotaq buffer, 20µM dNTPs, 1µM concentrations of each primer (gene specific, adapter, oligo dt-adapter), and 0.2µg template cDNA.

Cloning into the expression vector

The full length amino acid sequence of the MBO synthase gene contains an N-terminal region similar to the plastid transport sequences found in known conifer monoterpene synthases. These transport sequences have been found to interfere with protein expression in *E. coli* (Williams et al. (1998) *Biochemistry*, (37):12213-12220), hence primers were designed to amplify a fragment of the putative MBO synthase gene

lacking this chloroplast transport sequence. Hereafter this will be referred to as an expression length sequence.

Expression length sequences were first ligated into the vector p-GEMT easy (Promega, Madison WI) and transformed into the *E. coli* host strain DH5 α . pGEMT–
5 easy plasmids containing the MBO synthase gene (pGEMT-MBO) were extracted from the DH5 α host and digested with the restriction enzymes SphI and PstI to release an ~ 1800bp fragment containing the MBO synthase gene with a 5' SphI and 3' PstI splice site. This fragment was then gel purified and ligated into the expression vector pQE-31 (Qiagen) following digestion of pQE-31 vector with the restriction enzymes SphI and PstI
10 and gel purification of the 3.4kb digested product. Gel purification of restriction digests was done using a Qiagen gel purification kit (Qiagen). Following ligation of the expression length MBO sequence into the expression vector pQE-31, pQE-31-MBO was transformed into *E. coli* strain BL21-CodonPlus(DE3)-RIL (Stratagene, La Jolla CA). This strain of *E. coli* has been modified to correct for codon bias by increasing the
15 expression of several t-RNAs that match codons that are common in plant terpene synthases, but rare in *E. coli*. The MBO synthase gene contains several codons for which t-RNAs in *E. coli* are rare.

Functional Expression and Assay for Methyl Butenol production.

E. coli BL21-Codon Plus(DE3)-RIL containing expression plasmid pQE-31-MBO
20 were grown overnight at 37°C with shaking at 200rpm in 5ml Luri Broth supplemented with 200 μ g/ml ampicillin and 34 μ g/ml chloramphenicol. Aliquots of bacterial culture (0.5ml) were placed into 5ml vials and protein expression induced by adding 0.5mM IPTG. Vials were capped with a septum and incubated at 30°C with shaking at 200 rpm for 2 hours before analysis of headspace volatiles by gas chromatography.

25 To measure methyl butenol in vial headspace, 3ml aliquots of headspace were withdrawn from the vial and cryofocused by injection into a stainless steel sample loop immersed in liquid nitrogen. The cryofocused sample was then flash vaporized by immersing the sample loop in warm water and flushed onto a widebore capillary DB-1 GC column, 30m length, 0.32mm ID, 5micron film (JW Scientific, Folsom, CA) using
30 flow of Helium carrier gas, and analyzed using a Photoionization detector. Component separation was achieved using a temperature program of 30°C for 10 minutes, followed by a temperature ramp of 2°C per minute and a hold at 60°C for 5 minutes. The identity of methylbutenol measured in vial headspace was assessed by comparison of GC peak

retention times with those obtained from authentic MBO standards, and by GC-mass spectrometry.

Example II: Assays of Activity of MBO Synthase from *P. sabiniana*

The below assays were run at pH 8.0 in a 50 mM Hepes buffer with 10% glycerol,
 5 100 mM KCl, 20 mM MgCl₂, 5 mM MnCl₂, and 10 mM DMAPP. Assays were incubated for 2 hours at 37° C, and vial headspace was run on GC. Products were identified by comparing retention times with authentic standards.

A) Methylbutenol and Isoprene Production from Soluble and Insoluble Extract Fractions

10 Figure 3 demonstrates that expressed enzyme (with both N and C-terminal His tags (for purification) possess dual catalytic activity and can form both isoprene and methylbutenal from the substrate DMAPP (dimethylallyl pyrophosphate (or - diphosphate)). The enzyme partitions to the insoluble fraction during the extraction process.

15 B) Effect of Boiling and EDTA on Enzyme Activity

Figure 4 demonstrates that most of the activity is in the insoluble fraction. Furthermore, experiments determined that the conversion of DMAPP to isoprene and MBO is the result of the MBO synthase. For example, assays run with a boiled pellet (the most active extract) show little or no production of isoprene or MBO. Thus, boiling the
 20 extract inactivated the enzyme responsible for DMAPP conversion. Also, assays run with active enzyme in the presence of EDTA show little conversion of DMAPP to isoprene or MBO. This is significant because MBO synthase requires a divalent cation to function. EDTA chelates these cations and thus eliminated enzyme function.

C) Effect of Co-Factor Replacement on MBO Synthase Activity

25 Figure 5 demonstrates the effect of selectively removing the cation co-factors for MBO synthase function. The plant extract MBO synthase uses both Mn²⁺ and K⁺ as co-factors. Angiosperm isoprene synthases uses Mg²⁺. By washing the pellet containing the recombinant MBO synthase protein extracted from *E. coli*, the Mn²⁺ and K⁺ ions were removed. The pellets were then resuspended in isoprene synthase buffer containing Mg²⁺
 30 or MBO synthase buffer containing Mn²⁺ and K⁺. Assays run in isoprene synthase buffer (without Mn²⁺ and K⁺ cofactors) show little activity. Assays run in MBO synthase buffer (with both Mn²⁺ and K⁺) show activity producing isoprene and MBO with about 3 fold preference for forming MBO. Thus, the catalytic activity observed is from an MBO synthase (e.g., the cloned gene).

All publications, patents and patent applications are incorporated herein by reference. While in the foregoing specification this invention has been described in
5 relation to certain preferred embodiments thereof, and many details have been set forth for purposes of illustration, it will be apparent to those skilled in the art that the invention is susceptible to additional embodiments and that certain of the details described herein may be varied considerably without departing from the basic principles of the invention.

WHAT IS CLAIMED IS:

1. An isolated and purified methyl butenol (MBO) synthase nucleic acid molecule,
wherein the MBO synthase nucleic acid molecule comprises any one of SEQ ID
NOs:1-17 or a nucleic acid molecule having at least about 80% sequence identity
thereof.
2. An expression vector comprising a MBO synthase nucleic acid molecule of claim 1.
3. A prokaryotic or eukaryotic host cell transformed with a MBO synthase nucleic
acid molecule of claim 1.
4. A prokaryotic or eukaryotic host cell transformed with the vector of claim 2.
5. The host cell of claim 3 or 4, wherein the transformed cell expresses MBO
synthase.
6. An isolated and purified methyl butenol (MBO) synthase polypeptide, wherein the
MBO synthase polypeptide comprises any one of SEQ ID NOs:18-34 or a
polypeptide having at least about 80% sequence identity thereof.
7. A method to produce methyl butenol comprising transforming a host cell with a
nucleic acid molecule of claim 1 and expressing said molecule in a host cell so as
to yield methyl butenol.
8. The method of claim 7, wherein said host cell is a strain of yeast, bacteria,
cyanobacteria, eukaryotic micro algae or a combination thereof.
9. The method of claim 8, wherein said yeast is *Saccharomyces cerevisiae*.
10. The method of claim 8, wherein said bacteria is a C5- or C6-fermentative
organism.

11. The method of claim 10, wherein said C5-fermentative organism is a species of *Zymomonas*.
12. The method of claim 8, wherein said bacteria is *Klebsiella oxytoca*.
- 5 13. The method of claim 8, wherein said cyanobacteria is *Synechococcus sp.*, *Synechocystis sp.*, or *Anabaena sp.*
14. The method of claim 8, wherein said eukaryotic micro algae is *Chlorella sp.*
10 *Scenedesmus sp.*, *Bracteococcus sp.* or *Chlamydomonas sp.*
15. A method to produce methyl butenol (MBO) comprising contacting a methyl butenol (MBO) synthase with a substrate so as to yield MBO.
16. The method of claim 15, wherein the methyl butenol (MBO) synthase comprises
15 any one of SEQ ID NOs:18-34, a polypeptide having at least about 80% sequence identity.
17. The method of claim 15 or 16, wherein the substrate is a carbohydrate.
- 20 18. The method of claim 17, wherein the carbohydrate is a sugar, starch and/or cellulose.
19. An isolated and purified methyl butenol (MBO) synthase nucleic acid or protein molecule.

P._pseudostrobus_estevezii

TCTAGCAAGGTTAAGGTTGTCCGCAGAACGATCTCAACTTCCATCCGCATG
TGTCGGATAACCACTGTATCCGGTGAAGGCGTACAGAGACGCATAGCAAA
TCATCATTTCCAACCTCTGGGACGATAATTTTCATACAGTCCCTCTCAACCCC
TTATGGGGCAATTTTCGTACCATGAAAGTGCTCAGAAACTTATTGGAGAAG
TAAAAGAGATAATCAATTCAATCTCGGTTAAAGATGGAGAATTAATCACC
CCCTCCAATGATCTCCTTATGCGGCTCTCTATAGTCGATAGCATTGAACGT
TTGGGAATTGATAGGCATTTCAAAAGTGAAATAAAATCAGCTCTGGATTA
CGTTTACAGTTATTGGAACGAAAAAGGCATTGGGTGGGGAAGAGATAGTG
TTGTTGCCGATCTCAACTCAACTGCCTTGGGGCTTCGAACTCTACGACTAC
ACGGATACCCGGTGTCTTCAGATGTGTTACAACACTTCAAAGAACAAAAA
GGGCAGTTTGCATGTTTCGGCCATTCAAACAGAGGGAGAGATAAGAAGTGT
TCTCAACTTATTTTCGGGCTTCCCAAATTCCTTTCCGGGAGAGAAAGTTAT
GGAAGAGGCAGAAAGTCTTCTCTACAATATATTTAAAAGAAGCCATACTAA
AGCTTCCGGTCTGCGGTCTTTCACGAGAGATATCGTACGTTCTGGAATATG
GTTGGCATATAAATTTGCCAAGATTGGAAGCAAGGAACCTACATCGACGTA
TTTGACAGGACCCCATTTATTTGACGCCAAATATGAAGACCCAAAAACT
TCTAGAACTTGCAAAGTTGGAGTTCAATATGTTTCACTCTTTACAACAGCA
AGAGCTAAAGCTTCTCTCCAGATGGTGGAAAGATTTCGGGTTTCTCTCAAAT
GACCTTCCCTCGGCATCGTCACGTGGAATATTACACTTTGGCATCTTGCAT
TGATAGTGAACCTCAACATTCTTCGTTTCAGACTTGGATTTGCCAAAATCTT
TCATCTTGCCACGGTTCTTGACGATATTTACGACACCTTTGGCACGATGGA
TGAGCTAGAACTCTTCACGGCGGCAGTTAAGAGGTGGCATCCGTCTGCGA
CGGAGTGGCTTCCAGAATATATGAAAGGAGTATATATGGTGCTTTACGAA
ACCGTTAACGAAATGGCAGGAGAAGCAGAAAAGTCTCAAGGCCGAGACA
CGCTCAACTATGCCCCGAAATGCTTTGGAGGCTTATATTGATGCTTCTATGA
AAGAAGCGAAGTGGATTTTCAGTGGTTTTTTTGCCAACATTTGAGGAGTAC
CTGGATAACGGGAAAGTTAGTTTTCGGTTATAGCATTGGCACATTGCAACC
CATTCTGACGTTGGGCATTCCCTTTCTCATCACATCCTACAAGAAATAGA
CTTTCCTTCCAGGCTCAATGATGTGGCATGTTCCATTCTCCGACTAAAAGG
CGACATTACACTTACCAGGCTGAGAGGAGCCGTGGAGAAAAATCTTCGT
GTATATCATGTTATATGGAAGAGAATCCCGAGTCAACAGAGGAAGATGCA
ATCAATCATGTCAACTCCATGGTCGACAAATTACTCAAGGAACTAAATTG
GGAGTATCTGAGACCTGATAGCAATGTTCCAATCACTTCCAAGAAACATG
CATTTGACATTCTGAGAGCTTTCTACCATCTCTACAAATACCGAGATGGCT
TCAGCGTTGCGAACTATGAAATAAAGAATTTGGTCATGACAACCGTCATT
GAGCCTGTGCCTTTATAA- (SEQ ID NO:1)

P._cooperi_ornelasi_MBO

TCTAGCAAGGTTAAGGTTGTCCGCAGAACGATCTCAACTTCCATCCGCATG
TGTCGGATAACCACTGTATCCGGTGAAGGCGTACAGAGACGCATAGCAAA
TCATCATTTCCAACCTCTGGGACGATAATTTTCATACAGTCCCTCTCAACGCC
TTATGGGGCAATTTTCGTACCATGAAAGTGCTCAGAAACTTATTGGAGAAG
TAAAAGAGATGATCAATTCAATCTCGGTTAAAGATGGAGAATTAATCACC
CCCTCCAATGATCTCCTTATGCGGCTCTCTATAGTCGATAGCATTGAACGT
TTGGGAATCGATAGGCATTTCAAAAGTGAAATAAAATCAGCTCTGGATTA
TGTTTACAGTTATTGGAACGAAAAAGGCATTGGGTGGGGAAGAGATAGTG
TTGTTGCCGATCTCAACTCAACTGCCTTGGGGCTTCGAACTCTACGACTAC

FIG. 1A

2/20

ACGGATACCCGGTGTCTTCAGATGTGTTACAACACTTCAAAGAACAAAA
GGGCAGTTTGCATGTTTCGGCCATTCAAACAGAGGGAGAGATAAGAAGTGT
TCTCAACTTATTTTCGGGCTTCCCAAATTGCCTTTCCGGGAGAGAAAAGTTAT
GGAAGAGGCAGAAAGTCTTCTCTACAATATATTTAAAAGAAGCCATACTAA
AGCTTCCGGTCTGCGGTCTTTCACGAGAGATATCGTACGTTCTGGAATATG
GTTGGCATATAAATTTGCCAAGATTGGAAGCAAGGAACACTACATCGACGTA
TTTGGACAGGACCCCATTTATTTGACGCCAAATATGAAGACCCAAAAACT
TCTAGAAGTTGCAAAGTTGGAGTTCAATATGTTTCACTCTTTACAACAGCA
AGAGCTAAAGCTTCTCTCCAGATGGTGGAAAGATTTCGGGTTTCTCTCAAAT
GACCTTCCCTCGGCATCGTCACGTGGAATATTACACTTTGGCATCTTGCAT
TGATAGTGAACCTCAACATTCTTCGTTTCAGACTTGGATTTGCCAAAATCTT
TCATCTTGCCACGGTCTTGACGATATTTACGACACCTTTGGCACGATGGA
TGAGCTAGAAGCTCTTCACGGCGGCAGTTAAGAGGTGGCATCCGTCTGCGA
CGGAGTGGCTTCCAGAATATATGAAAGGAGTATATATGGTGCTTTACGAA
ACCGTTAACGAAATGGCAGGAGAAGCAGAAAAGTCTCAAGGCCGAGACA
CGCTCAACTATGCCCCGAAATGCTTTGGAGGCTTATATTGATGCTTCTATGA
AAGAAGCGAAGTGGATTTTCAGTGGTTTTTTGCCAACATTTGAGGAGTAC
CTGGATAACGGGAAAGTTAGTTTCGGTTATAGCATTGGCACATTGCAACC
CATCTGACGTTGGGCATTCCCTTTCCTCATCACATCCTACAAGAAATAGA
CTTTCCTTCCAGGCTCAATGATGTGGCATGTTCCATTCTCCGACTAAAAGG
CGACATTACACTTACCAGGCTGAGAGGAGCCGTGGAGAAAAATCTTCGT
GTATATCATGTTATATGGAAGAGAATCCCGAGTCAACAGAGGAAGATGCA
ATCAATCATATCAACTCCATGGTCGACAAATTAATCAAGGAACATAAATTG
GGAGTATCTGAGACCTGATAGCAATGTTCCAATCACTTCCAAGAAACATG
CATTTGACATTCTGACAGCTTCTACCATCTCTACAAATACCGAGATGGCT
TCAGCGTTGCGAACTATGAAATAAAGAATTTGGTCATGACAACCGTCATT
GAGCCTGTGCCTTTATAA-(SEQ ID NO:2)

P._hartwegii_MBO

TCTAGCAAGGTTAAGGTTGTCCGCAGAACGATCTCAACTTCCATCCGCATG
TGTCGGATAACCACTGTATCCGGTGAAGGCGTACAGAGACGCATAGCAAA
TCATCATTTCCAACCTCTGGGACGATAATTTTCATACAGTCCCTCTCAACGCC
TTATGGGGCAATTTTCGTACCATGAAAGTGCTCAGAAACTTATTGGAGAAG
TAAAAGAGATGATCAATTCAATCTCGGTTAAAGATGGAGAATTAATCACC
CCCTCCAATGATCTCCTTATGCGGCTCTCTATAGTCGATAGCATTGAACGT
TTGGGAATCGATAGGCATTTCAAAGTGAAATAAAATCAGCTCTGGATTA
TGTTTACAGTTATTGGAACGAAAAAGGCATTGGGTGGGGAAGAGATAGTG
TTGTTGCCGATCTCAACTCAACTGCCTTGGGGCTTCGAACTCTACGACTAC
ACGGATACCCGGTGTCTTCAGATGTGTTACAACACTTCAAAGAACAAAA
GGGCAGTTTGCATGTTTCGGCCATTCAAACAGAGGGAGAGATAAGAAGTGT
TCTCAACTTATTTTCGGGCTTCCCAAATTGCCTTTCCGGGAGAGAAAAGTTAT
GGAAGAGGCAGAAAGTCTTCTCTACAATATATTTAAAAGAAGCCATACTAA
AGCTTCCGGTCTGCGGTCTTTCACGAGAGATATCGTACGTTCTGGAATATG
GTTGGCATATAAATTTGCCAAGATTGGAAGCAAGGAACACTACATCGACGTA
TTTGGACAGGACCCCATTTATTTGACGCCAAATATGAAGACCCAAAAACT
TCTAGAAGTTGCAAAGTTGGAGTTCAATATGTTTCACTCTTTACAACAGCA
AGAGCTAAAGCTTCTCTCCAGATGGTGGAAAGATTTCGGGTTTCTCTCAAAT
GACCTTCCCTCGGCATCGTCACGTGGAATATTACACTTTGGCATCTTGCAT
TGATAGTGAACCTCAATATTCTTCGTTTCAGACTTGGATTTGCCAAAATCTT
TCATCTTGCCACGGTCTTGACGATATTTACGACACCTTTGGCACGATGGA

FIG. 1B

3/20

TGAGCTAGAACTCTTCACGGCGGCAGTTAAGAGGTGGCATCCGTCTGCGA
CGGAGTGGCTTCCAGAATATATGAAAGGAGTATATATGGTGCTTTACGAA
ACCGTTAACGAAATGGCAGGAGAAGCAGAAAAGTCTCAAGGCCGAGACA
CGCTCAACTATGCCCCGAAATGCTTTGGAGGCTTATATTGATGCTTCTATGA
AAGAAGCGAAGTGGATTTTCAGTGGTTTTTTTGCCAACATTTGAGGAGTAC
CTGGATAACGGGAAAGTTAGTTTTCGGTTATAGCATTGGCACATTGCAACC
CATTCTGACGTTGGGCATTCCCTTTCCTCATCACATCCTACAAGAAATAGA
CTTTCCTTCCAGGCTCAATGATGTGGCATGTTCCATTCTCCGACTAAAAGG
CGACATTCACACTTACCAGGCTGAGAGGAGCCGTGGAGAAAAATCTTCGT
GTATATCATGTTATATGGAAGAGAATCCCGAGTCAACAGAGGAAGATGCA
ATCAATCATATCAACTCCATGGTCGACAAATTACTCAAGGAACTAAATTG
GGAGTATCTGAGACCTGATAGCAATGTTCCAATCACTTCCAAGAAACATG
CATTTGACATTCTGAGAGCTTTCTACCATCTCTACAAATACCGAGATGGCT
TCAGCGTTGCGAACTATGAAATAAAGAATTTGGTCATGACAACCGTCATT
GAGCCTGTGCCTTTATAA-(SEQ ID NO:3)

*P. ponderosa*_MBO1

TCTAGCAAGGTTAAGGTTGTCCGCAGAACGATCTCAACTTCCATCCGCATG
TGTCGGATAACCACTGAATCCGGTGAAGGCGTACAGAGACGCATAGCAAA
TCATCATTTCCAACCTCTGGGACGATAATTTTCATACAGTCCCTCTCAACGCC
TTATGGGGCAATTTTCGTACCATGAAAGTGCTCAGAAACTTATTGGAGAAG
TAAAAGAGATGATCAATTCAATCTCGCTTAAAGATGGAGAATTAATCACC
CCCTCCAATGATCTCCTTATGCGGCTCTCTATAGTTGATAGCATTGAACGT
TTGGGAATCGATAGGCATTTCAAAAGTGAAATAAAATCAGCTCTGGATTA
TGTTTACAGTTATTGGAACGAAAAAGGCATTGGGTGGGGAAGAGATAGTG
TTGTTGCCGATCTCAACTCAACTGCCTTGGGGCTTCGAACTCTACGACTAC
ACGGATAACCCGGTGTCTTCAGATGTGTTACAACACTTCAAAGAACAAAAA
GGGCAGTTTGCATGTTTCGGCCATTCAAACAGAGGGAGAGATAAGAAGTGT
TCTCAACTTATTTTCGGGCTTCCCAAATTCCTTTCCGGGAGAGAAAGTTAT
GGAAGAGGCAGAAAGTCTTCTCTACAATATATTTAAAAGAAGCCATACTAA
AGCTTCCGGTCTGCGGTCTTTCACGAGAGATATCGTACGTTCTGGAATATG
GTTGGCATATAAATTTGCCAAGATTGGAAGCAAGGAACTACATCGACGTA
TTTGGACAGGACCCCATTTATTTGACGCCAAATATGAAGACCCAAAAACT
TCTAGAACTTGCAAAGTTGGAGTTCAATATGTTTCACTCTTTACAACAGCA
AGAGCTAAAGCTTCTCTCCAGATGGTGGAAAGATTTCGGGTTTCTCTCAAAT
GACCTTCCCTCGGCATCGTCACGTGGAATATTACACTTTGGCATCTTGCAT
TGATAGTGAACCTCAACATTCTTCGTTTCAGACTTGGATTTGCCAAAATCTT
TCATCTTGCCACGGTTCTTGACGATATTTACGACACCTTTGGCACGATGGA
TGAGCTAGAACTCTTCACGGCGGCAGTTAAGAGGTGGCATCCGTCTGCGA
CGGAGTGGCTTCCAGAATATATGAAAGGAGTATATATGGTGCTTTACGAA
ACCGTTAACGAAATGGCGAGAGAAGCAGAAAAGTCTCAAGGCCGAGACA
CGCTCAACTATGCCCCGAAATGCTTTGGAGGCTTATATTGATGCTTCTATGA
AAGAAGCGAAGTGGATTTTCAGTGGTTTTTTTGCCAACATTTGAGGAGTAC
CTGGATAACGGGAAAGTTAGTTTTCGGTTATAGCATTGGCACATTGCAACC
CATTCTGACGTTGGGCATTCCCTTTCCTCATCACATCCTACAAGAAATAGA
CTTTCCTTCCAGGCTCAATGATGTGGCATGTTCCATTCTCCGACTAAAAGG
CGACATTCACACTTACCAGGCTGAGAGGAGCCGTGGAGAAAAATCTTCGT
GTATATCATGTTATATGGAAGAGAATCCCGAGTCAACAGAGGAAGATGCA
ATCAATCATATCAACTCCATGGTCGACAAATTACTCAAGGAACTAAATTG
GGAGTATCTGAGACCTGATAGCAATGTTCCAATCACTTCCAAGAAACATG

FIG. 1C

4/20

CATTTGACATTCTGAGAACTTTCTACCATCTCTACAAATACCGAGATGGCT
TCAGCGTTGCGAACTATGAAATAAAGAATTTGGTCATGACAACCGTCATT
GAGCCTGTGCCTTTATAA-(SEQ ID NO:4)

*P._ponderosa*_MBO2

TCTAGCAAGGTTAAGGTTGTCCGCAGAACGATCTCAACTTCCATCCGCATG
TGTCGGATAAACCCTGAATCCGGTGAAGGCGTACAGAGACGCATAGCAAA
TCATCATTTCCAACCTCTGGGACGATAATTTTCATACAGTCCCTCTCAACGCC
TTATGGGGCAATTTTCGTACCATGAAAGTGCTCAGAAACTTATTGGAGAAG
TAAAAGAGATGATCAATTCAATCTCGCTTAAAGATGGAGAATTAATCACC
CCCTCCAATGATCTCCTTATGCGGCTCTCTATAGTTGATAGCATTGAACGT
TTGGGAATCGATAGGCATTTCAAAAGTGAAATAAAATCAGCTCTGGATTA
TGTTTACAGTTATTGGAACGAAAAAGGCATTGGGTGGGGAAGAGATAGTG
TTGTTGCCGATCTCAACTCAACTGCCTTGGGGCTTCGAACTCTACGACTAC
ACGGATAACCCGGTGTCTTCAGATGTGTTACAACACTTCAAAGAACAAAAA
GGGCAGTTTGCATGTTTCGGCCATTCAAACAGAGGGAGAGATAAGAAGTGT
TCTCAACTTATTTTCGGGCTTCCCAAATTGCCTTTCCGGGAGAGAAAGTTAT
GGAAGAGGCAGAAAGTCTTCTCTACAATATATTTAAAAGAAGCCATACTAA
AGCTTCCGGTCTGCGGTCTTTCACGAGAGATATCGTACGTCTTGGAAATATG
GTTGGCATATAAATTTGCCAAGATTGGAAGCAAGGAAGTACATCGACGTA
TTTGGACAGGACCCCATTTATTTGACGCCAAATATGAAGACCCAAAAAATT
TCTAGAACTTGCAAAGTTGGAGTTCAATATGTTTCACTCTTTACAACAGCA
AGAGCTAAAGCTTCTCTCCAGATGGTGGAAAGATTTCGGGTTTCTCTCAAAT
GACCTTCCCTCGGCATCGTCACGTGGAATATTACACTTTGGCATCTTGCAT
TGATAGTGAACCTCAACATTCTTCGTTCAGACTTGGATTTGCCAAAATCTT
TCATCTTGCCACGGTCTTGACGATATTTACGACACCTTTGGCACGATGGA
TGAGCTAGAACTCTTCACGGCGGCAGTTAAGAGGTGGCATCCGTCTGCGA
CGGAGTGGCTTCCAGAATATATGAAAGGAGTATATATGGTGCTTTACGAA
ACCGTTAACGAAATGGCGAGAGAAGCAGAAAAGTCTCAAGGCCGAGACA
CGCTCAACTATGCCCCGAAATGCTTTGGAGGCTTATATTGATGCTTCTATGA
AAGAAGCGAAGTGGATTTTCAGTGGTTTTTTGCCAACATTTGAGGAGTAC
CTGTATAACGGGAAAGTTAGTTTCGGTTATAGCATTGGCACATTGCAACCC
ATTCTGACGTTGGGCATTCCCTTTTCCTCATCACATCCTACAAGAAATAGAC
TTTCCTTCCAGGCTCAATGATGTGGCATGTTCCATTCTCCGACTAAAAGGC
GACATTACACTTACCAGGCTGAGAGGAGCCGTGGAGAAAAATCTTCGTG
TATATCATGTTATATGGAAGAGAATCCCGAGTCAACAGAGGAAGATGCAA
TCAATCATATCAACTCCATGGTTCGACAAATTACTCAAGGAACTAAATTGG
GAGTATCTGAGACCTGATAGCAATGTTCCAATCACTTCCAAGAAACATGC
ATTTGACATTCTGAGAACTTTCTACCATCTCTACAAATACCGAGATGGCTT
CAGCGTTGCGAACTATGAAATAAAGAATTTGGTCATGACAACCGTCATTG
AGCCTGTGCCTTTATAA-(SEQ ID NO:5)

*P._jeffreyi*_MBO

TCTAGCAAGGTTAAGGTTGTCCGCAGAACGATCTCAACTTCCATCCGCATG
TGTCGGATAAACCCTGAATCCGGTGAAGGCGTACAGAGACGCATAGCAAA
TCATCATTTCCAACCTCTGGGACGATAATTTTCATACAGTCCCTCTCAACGCC
TTATGGGGCAATTTTCGTACCATGAAAGTGCTCAGAAACTTATTGGAGAAG
TAAAAGAGATGATCAATTCAATCTCGCTTAAAGATGGAGAATTAATCACC
CCCTCCAATGATCTCCTTATGCGGCTCTCTATAGTTGATAGCATTGAACGT
TTGGGAATCGATAGGCATTTCAAAAGTGAAATAAAATCAGCTCTGGATTA

FIG. 1D

5/20

TGTTTACAGTTATTGGAACGAAAAAGGCATTGGGTGGGGAAGAGATAGTG
TTGTTGCCGATCTCAACTCAACTGCCTTGGGGCTTCGAACCTACGACTAC
ACGGATACCCGGTGTCTTCAGATGTGTTACAACACTTCAAAGAACAAAAA
GGGCAGTTTGCATGTTTCGGCCATTCAAACAGAGGGAGAGATAAGAAGTGT
TCTCAACTTATTTTCGGGCTTCCCAAATTGCCTTTCCGGGAGAGAAAGTTAT
GGAAGAGGCAGAAAGTCTTCTCTACAATATATTTAAAAGAAGCCATACTAA
AGCTTCCGGTCTGCGGTCTTTCACGAGAGATATCGTACGTTCTGGAATATG
GTTGGCATATAAAATTTGCCAAGATTGGAAGCAAGGAACCTACATCGACGTA
TTTGGACAGGACCCCATTTATTTGACGCCAAATATGAAGACCCAAAAACT
TCTAGAACTTGCAAAGTTGGAGTTCAATATGTTTCACTCTTTACAACAGCA
AGAGCTAAAGCTTCTCTCCAGATGGTGGAAAGATTTCGGGTTTCTCTCAAAT
GACCTTCCCTCGGCATCGTCACGTGGAATATTACACTTTGGCATCTTGCAT
TGATAGTGAACCTCAACATTCTTCGTTTCAGACTTGGATTTGCCAAAATCTT
TCATCTTGCCACGGTTCTTGACGATATTTACGACACCTTTGGCACGATGGA
TGAGCTAGAACTCTTCACGGCGGCAGTTAAGAGGTGGCATCCGTCTGCGA
CGGAGTGGCTTCCAGAATATATGAAAGGAGTATATATGGTGCTTTACGAA
ACCGTTAACGAAATGGCAGGAGAAGCAGAAAAGTCTCAAGGCCGAGACA
CGCTCAACTATGCCCCGAAATGCTTCGGAGGCTTATATTGATGCTTCTATGA
AAGAAGCGAAGTGGATTTTCAGTGGTTTTTTGCCAACATTTGAGGAGTAC
CTGGATAACGGGAAAGTTAGTTTTCGGTTATAGCATTGGCACATTGCAACC
CATTCTGACGTTGGGCATTCCCTTTCTCATCACATCCTACAAGAAATAGA
CTTTCCTTCCAGGCTCAATGATGTGGCATGTTCCATTCTCCGACTAAAAGG
CGACATTCACACTTACCAGGCTGAGAGGAGCCGTGGAGAAAAATCTTCGT
GTATATCATGTTATATGGAAGAGAATCCCGAGTCAACAGAGGAAGATGCA
ATCAATCATATCAACTCCATGGTCGACAAATTACTCAAGGAACATAATTG
GGAGTATCTGAGACCTGATAGCAATGTTCCAATCACTTCCAAGAAACATG
CATTTGACATTCTGAGAGCTTTCTACCATCTCTACAAATACCGAGATGGCT
TCAGCGTTGCGAACTATGAAATAAAGAATTTGGTCATGACAACCGTCATT
GAGCCTGTGCCTTTATAA-(SEQ ID NO:6)

*P._sabiniana*_MBO1

TCTAGCAAGGTTAAGGTTGTCCGCAGAACGATCTCAACTTCCATCCGCATG
TGTCGGATAACCACTGAATCCGGTGAAGGCGTACAGAGACGCATAGCAAA
TCATCATTTCCAACCTCTGGGACGATAATTTTCATACAGTCCCTCTCAACGCC
TTATGGGGCAATTTTCGTACCATGAAAGTGCTCAGAACTTATTGGAGAAG
TAAAAGAGATGATCAATTCAATCTCGGTTAAAGATGGAGAATTAATCACC
CCCTCCAATGATCTCCTTATGCGGCTCTCTATAGTTGATAGCATTGAACGT
TTAGGAATCGATAGGCATTTCAAAGTGAAATAAAATCAGCTCTGGATTA
TGTTTACAGTTATTGGAACGAAAAAGGCATTGGGTGGGGAAGAGATAGTG
TTGTTGCCGATCTCAACTCAACTGCCTTGGGGCTTCGAACCTACGACTAC
ACGGATACCCGGTGTCTTCAGATGTGTTACAACACTTCAAAGAACAAAAA
GGGCAGTTTGCATGTTTCGGCCATTCAAACAGAGGGAGAGATAAGAAGTGT
TCTCAACTTATTTTCGGGCTTCCCAAATTGCCTTTCCGGGAGAGAAAGTTAT
GGAAGAGGCAGAAAGTCTTCTCTACAATATATTTAAAAGAAGCCATACTAA
AGCTTCCGGTCTGCGGTCTTTCACGAGAGATATCGTACGTTCTGGAATATG
GTTGGCATATAAAATTTGCCAAGATTGGAAGCAAGGAACCTACATCGACGTA
TTTGGAGAGGACCCCATTTATTTGACGCCAAATATGAAGACCCAAAAACT
TCTAGAACTTGCAAAGTTGGAGTTCAATATGTTTCACTCTTTACAACAGCA
AGAGCTAAAGCTTCTCTCCAGATGGTGGAAAGATTTCGGGTTTCTCTCAAAT
GACCTTCCCTCGGCATCGTCACGTGGAATATTACACTTTGGCATCTTGCAT

FIG.1E

6/20

TGATAGTGAACCTCAACATTCTTCGTTTCAGACTTGGATTTGCCAAAATCTT
TCATCTTGCCACGGTTCCTTGACGATATTTACGACACCTTTGGCACGATGGA
TGAGCTAGAACTCTTCACGGCGGCAGTTAAGAGGTGGCATCCGTCTGCGA
CGGAGTGGCTTCCAGAATATATGAAAGGAGTATATATGGTGCTTTACGAA
ACCGTTAACGAAATGGCAGGAGAAGCAGAAAAGTCTCAAGGCCGAGACA
CGCTCAACTATGGCCGAAATGCTTTGGAGGCTTATATTGATGCTTCTATGG
AAGAAGCGAAGTGGATTTTCAGTGGTTTTTTTGCCAACATTTGAGGAGTAC
CTGGATAACGGGAAAGTTAGTTTCGGTTATGGCATTGGCACATTGCAACC
CATTCTGACGTTGGGCATTCCCTTTCCTCATCACATCCTACAAGAAATAGA
CTTTCCTTCCAGGCTCAATGATGTGGCATCTTCCATTCTCCGACTAAAAGG
CGACATTACACTTACCAGGCTGAGAGGAGCCGTGGAGAAAAATCTTCGT
GTATATCATGTTATATGGAAGAGAATCCCGAGTCAACAGAGGAAGATGCA
ATCAATCATATCAACTCCATGGTCGACAAATTACTCAAGGAACATAATTG
GGAGTATCTGAGACCTGATAGCAATGTTCCAATCACTTCCAAGAAACATG
CATTTGACATTCTGAGAGCTTTCTACCATCTCTACAAATACCGAGATGGCT
TCAGCGTTGCGAACTATGAAATAAAGAATTTGGTCATGACAACCGTCATT
GAGCCTG------(SEQ ID NO:7)

*P. sabiniana*_MBO2

TCTAGCAAGGTTAAGGTTGTCCGCAGAACGATCTCAACTTCCATCCGCATG
TGTCGGATAAACCCTGAATCCGGTGAAGGCGTACAGAGACGCATAGCAAA
TCATCATTTCCAACCTCTGGGACGATAATTTTCATACAGTCCCTCTCAACGCC
TTATGGGGCAATTTCTGTACCATGAAAGTGCTCAGAAACTTATTGGAGAAG
TAAAAGAGATGATCAATTCAATCTCGCTTAAAGATGGAGAATTAATCACC
CCCTCCAATGATCTCCTTATGCGGCTCTCTATAGTCGATAGCATTGAACGT
TTGGGAATCGATAGGCATTTCAAAAAGTGAAATAAAATCAGCTCTGGATTA
TGTTTACAGTTATTGGAACGAAAAAGGCATTGGGTGGGGAAGAGATAGTG
TTGTTGCCGATCTCAACTCAACTGCCTTGGGGCTTCGAACTCTACGACTAC
ACGGATAACCCGGTGTCTTCAGATGTGTTACAACACTTCAAAGAACAACAAA
GGGCAGTTTGCATGTTTCGGCCATTCAAACAGAGGGAGAGATAAGAAGTGT
TCTCAACTTATTTCTGGGCTTCCCAAATTGCCTTTCCGGGAGAGAAAGTTAT
GGAAGAGGCAGAAAGTCTTCTCTACAATATATTTAAAAGAAGCCATACTAA
AGCTTCCGGTCTGCGGTCTTTCACGAGAGATATCGTACGTTCTGGAATATG
GTTGGCATATAAATTTGCCAAGATTGGAAGCAAGGAACCTACATCGACGTA
TTTGGAGAGGACCCCATTTATTTGACGCCAAATATGAAGACCCAAAAACT
TCTAGAACTTGCAAAGTTGGAGTTCAATATGTTTCACTCTTTACAACAGCA
AGAGCTAAAGCTTCTCTCCAGATGGTGGAAAGATTCCGGTTTCTCTCAAAT
GACCTTCCCTCGGCATCGTCACGTGGAATATTACACTTTGGCATCTTGCAT
TGATAGTGAACCTCAACATTCTTCGTTTCAGACTTGGATTTGCCAAAATCTT
TCATCTTGCCACGGTTCCTTGACGATATTTACGACACCTTTGGCACGATGGA
TGAGCTAGAACTCTTCACGGCGGCAGTTAAGAGGTGGCATCCGTCTGCGA
CGGAGTGGCTTCCAGAATATATGAAAGGAGTATATATGGTGCTTTACGAA
ACCGTTAACGAAATGGCAGGAGAAGCAGAAAAGTCTCAAGGCCGAGACA
CGCTCAACTATGGCCGAAATGCTTTGGAGGCTTATATTGATGCTTCTATGG
AAGAAGCGAAGTGGATTTTCAGTGGTTTTTTTGCCAACATTTGAGGAGTAC
CTGGATAACGGGAAAGTTAGTTTCGGTTATGGCATTGGCACATTGCAACC
CATTCTGACGTTGGGCATTCCCTTTCCTCATCACATCCTACAAGAAATAGA
CTTTCCTTCCAGGCTCAATGATGTGGCATCTTCCATTCTCCGACTAAAAGG
CGACATTACACTTACCAGGCTGAGAGGAGCCGTGGAGAAAAATCTTCGT
GTATATCATGTTATATGGAAGAGAATCCCGAGTCAACAGAGGAAGATGCA

FIG. 1F

7/20

ATCAATCATATCAACTCCATGGTCGACAAATTACTCAAGGAACTAAATTG
GGAGTATCTGAGACCTGATAGCAATGTTCCAATCACTTCCAAGAAACATG
CATTTGACATTCTGAGAGCTTTCTACCATCTCTACAAATACCGAGATGGCT
TCAGCGTTGCGAACTATGAAATAAAGAATTTGGTCATGACAACCGTCATT
GAGCCTGTGCCTTTATAAA (SEQ ID NO:8)

*P. patula*_MBO

TCTAGCAAGGTTAAGGTTGTCCGCAGAACGATCTCAACTTCCATCCGCATG
TGTCGGATAACTACTGTATCCGGTGAAGGCGTACAGAGACGCATAGCAAA
TCATCATTTCCAACCTCTGGGACGATAATTTTCATACAGTCCCTCTCAACGCC
TTATGGGGCAATTTTCGTACCATGAAAGTGCTCAGAAACTTATTGGAGAAG
TAAAAGAGATGATCAATTCAATCTCGGTTAAAGATGGAGAATTAATCACC
CCCTCCAATGATCTCCTTATGCGGCTCTCTATAGTCGATAGCATTGAACGT
TTGGGAATCGATAGGCATTTCAAAAGTGAAATAAAATCAGCTCTGGATTA
TGTTTACAGTTATTGGAACGAAAAAGGCATTGGGTGGGGAAGAGATAGTG
TTGTTGCCGATCTCAACTCAACTGCCTTGGGGCTTCGAACTCTACGACTAC
ACGGATAACCCGGTGTCTTCAGATGTGTTACAACACTTCAAAGAACAAAAA
GGGCAGTTTGCATGTTTCGGCCATTCAAACAGAGGGAGAGATAAGAAGTGT
TCTCAACTTATTTTCGGGCTTCCCAAATTGCTTTTCCGGGAGAGAAAGTTAT
GGAAGAGGCAGAAAGTCTTCTCTACAATATATTTAAAAGAAGCCATACTAA
AGCTTCCGGTCTGCGGTCTTTCACGAGAGATATCGTACGTTCTGGAATATG
GTTGGCATATAAATTTGCCAAGATTGGAAGCAAGGAACTACATCGACGTA
TTTGACAGGACCCCATTTATTTGACGCCAAATATGAAGACCCAAAAACT
TCTAGAACTTGCAAAGTTGGAGTTCAATATGTTTCACTCTTTACAACAGCA
AGAGCTAAAGCTTCTCTCCAGATGGTGGAAAGATTTCGGGTTTCTCTCAAAT
GACCTTCCCTCGGCATCGTCACGTGGAATATTACACTTTGGCATCTTGCAT
TGATAGTGAACCTCAACATTCTTCGTTTCAGACTTGGATTTGCCAAAATATT
TCATCTTGCCACGGTTCTTGACGATATTTACGACACCTTTGGCACGATGGA
TGAGCTAGAACTCTTCACAGCGGCAGTTGAGAGGTGGCATCCGTCTGCGA
CGGAGTGGCTTCCAGAATATATGAAAGGAGTATATATGGTGCTTTACGAA
ACCGTTAACGAAATGGCAGGAGAAGCAGAAAAGTCTCAAGGCCGAGACA
CGCTCAACTATGCCCCGAAATGCTTTGGAGGCTTATATTGATGCTTCTATGA
AAGAAGCGAAGTGGATTTTCAGTGGTTTTTTTGCCAACATTTGAGGAGTAC
CTGGATAACGGGAAAGTTAGTTTTCGGTTATAGCATTGGCACATTGCAACC
CATTTCTGACGTTGGGCATTCCCTTTCTCATCACATCCTACAAGAAATAGA
CTTTCCTTCCAGGCTCAATGATGTGGCATGTTCCATTCTCCGACTAAAAGG
CGACATTCACACTTACCAGGCTGAGAGGAGCCGTGGAGAAAAATCTTCGT
GTATATCATGTTATATGGAAGAGAATCCCGAGTCAACAGAGGAAGATGCA
ATCGATCATATCAACTCCATGGTCGACAAATTACTCAAGGAACTAAATCG
GGAGTATCTGAGACCTGATAGCAATGTTCCAATCACTTCCAAGAAACATG
CATTTGACATTCTGAGAGCTTTCTACCATCTCTACAAATACCGAGATGGCT
TCAGCGTTGCGAACTATGAAATAAAGAATTTGGTCATGACAACCGTCATT
GAGCCTGTGCCTTTATAA-(SEQ ID NO:9)

*P. montezumae*_MBO

TCTAGCAAGGTTAAGGTTGTCCGCAGAACGATCTCAACTTCCATCCGCATG
TGTCGGATAACCACTGTATCCGGTGAAGGCGTACAGAGACGCATAGCAAA
TCATCATTTCCAACCTCTGGGACGATAATTTTCATACAGTCCCTCTCAACGCC
TTATGGGGCAATTTTCGTACCATGAAAGTGCTCAGAAACTTATTGGAGAAG
TAAAAGAGATGATCAATTCAATCTCGGTTAAAGATGGAGAATTAATCACC

FIG. 1G

8/20

CCCTCCAATGATCTCCTTATGCGGCTCTCTATAGTCGATAGCATTGAACGT
TTGGGAATCGATAGGCATTTCAAAAGTGAAATAAAATCAGCTCTGGATTA
TGTTTACAGTTATTGGAACGAAAAAGGCATTGGGTGGGGAAGAGATAGTG
TTGTTGCCGATCTCAACTCAACTGCCTTGGGGCTTCGAACTCTACGACTAC
ACGGATAACCCGGTGTCTTCAGATGTGTTACAGCACTTCAAAGAACAAAAA
GGGCAGTTTGCATGTTTCGGCCATTCAAACAGAGGGAGAGATAAGAAGTGT
TCTCAACTTATTTTCGGGCTTCCCAAATTGCCTTTCCGGGAGAGAAAAGTTAT
GGAAGAGGCAGAAAGTCTTCTCTACAATATATTTAAAAGAAGCCATACTAA
AGCTTCCGGTCTGCGGTCTTTCACGAGAGATATCGTACGTTCTGGAATATG
GTTGGCATATAAATTTGCCAAGATTGGAAGCAAGGAACACTACATCGACGTA
TTTGGACAGGACCCCATTTATTTGACGCCAAATATGAAGACCCAAAAACT
TCTAGAACTTGCAAAGTTGGAGTTCAATATGTTTCACTCTTTACAACAGCA
AGAGCTAAAGCTTCTCTCCAGATGGTGGAAAGATTTCGGGTTTCTCTCAAAT
GACCTTCCCTCGGCATCGTCACGTGGAATATTACACTTTGGCATCTTGCAT
TGATAGTGAACCTCAACATTCTTCGTTTCAGACTTGGATTTGCCAAAATGTT
TCATCTTGCCACGGTCTTGACGATATTTACGACACCTTTGGCACGATGGA
TGAGCTAGAACTCTTCACGGCGGCAGTTAAGAGGTGGCATCCGTCTGCGA
CGGAGTGGCTTCCAGAATATATGAAAGGAGTATATATGGTGCTTTACGAA
ACCGTTAACGAAATGGCAGGAGAAGCAGAAAAGTCTCAAGGCCGAGACA
CGCTCAACTATGCCCCGAAATGCTTTGGAGGCTTATATTGATGCTTCTATGA
AAGAAGCGAAGTGGATTTTCAGTGGTTTTTTGCCAACATTTGAGGAGTAC
CTGGATAACGGGAAAGTTAGTTTTCGGTTATAGCATTGGCACATTGCAACC
CATTCTGACGTTGGGCATTCCCTTTCCTCATCACATCCTACAAGAAATAGA
CTTTCCTTCCAGGCTCAATGATGTGGCATGTTCCATTCTCCGACTAAAAGG
CGACATTCACACTTACGAGGCTGAGAGGAGCCGTGGAGAAAAATCTTCGT
GTATATCATGTTATATGGAAGAGAATCCCGAGTCAACAGAGGAAGATGCA
ATCAATCATATCAACTCCATGGTCGACAAATTACTCAAGGAACATAATTG
GGAGTATCTGAGACCTGATAGCAATGTTCCAATCACTTCCAAGAAACATG
CATTTGACATTCTGAGAGCTTTCTACCATCTCTACAAATACCGAGATGGCT
TCAGCGTTGCGAACTATGAAATAAAGAATTTGGTCATGACAACCGTCATT
GAGCCTGTGCCTTTATAA-(SEQ ID NO:10)

P._pseudostrobus_MBO-----
AGACGCAKAGSTGRTYWYCATTCCAACCTYTGGGACGATAATTTTCATACA
GTCCCTCTCAACGCCTTATGGGGCAATTTTCGTACCATGAAAGTGCTCAGAA
ACTTATTGGAGAAGTAAAAGAGATGATCAATTCAATCTCGGTAAAGATG
GAGAATTAATCACCCCTCCAATGATCTCCTTATGCGGCTCTCTATAGTCG
ATAGCATTGAACGTTTGGGAATCGATAGGCATTTCAAAAGTGAAATAAAA
TCAGCTCTGGATTATGTTTACAGTTATTGGAACGAAAAAGGCATTGGGTG
GGGAAGAGATAGTGTTGTTGCCGATCTCAACTCAACTGCCTTGGGGCTTC
GAACTCTACGACTACACGGATACCCGGTGTCTTCAGATGTGTTACAGCACT
TCAAAGAACAAAAAGGGCAGTTTGCATGTTTCGGCCATTCAAACAGAGGGA
GAGATAAGAAGTGTTCTCAACTTATTTTCGGGCTTCCCAAATTGCCTTTCCG
GGAGAGAAAGTTATGGAAGAGGCAGAAAGTCTTCTCTACAATATATTTAAA
AGAAGCCATACTAAAGCTTCCGGTCTGCGGTCTTTCACGAGAGATATCGT
ACGTTCTGGAATATGGTTGGCATATAAATTTGCCAAGATTGGAAGCAAGG
AACTACATCGACGTATTTGGACAGGACCCCATTTATTTGACGCCAAATATG
AAGACCCAAAAACTTCTAGAACTTGCAAAGTTGGAGTTCAATATGTTTCA
CTCTTTACAACAGCAAGAGCTAAAGCTTCTCTCCAGATGGTGGAAAGATT
CGGGTTTCTCTCAAATGACCTTCCCTCGGCATCGTCACGTGGAATATTACA

FIG. 1H

9/20

CTTTGGCATCTTGCATTGATAGTGAACCTCAACATTCTTCGTTTCAGACTTG
GATTTGCCAAAATGTTTCATCTTGCCACGGTTCTTGACGATATTTACGACA
CCTTTGGCACGATGGATGAGCTAGAACTCTTCACGGCGGCAGTTAAGAGG
TGGCATCCGTCTGCGACGGAGTGGCTTCCAGAATATATGAAAGGAGTATA
TATGGTGCTTTACGAAACCGTTAACGAAATGGCAGGAGAAGCAGAAAAAGT
CTCAAGGCCGAGACACGCTCAACTATGCCCCGAAATGCTTTGGAGGCTTAT
ATTGATGCTTCTATGAAAGAAGCGAAGTGGATTTTCAGTGGTTTTTTTGCCA
ACATTTGAGGAGTACCTGGATAACGGGAAAGTTAGTTTCGGTTATAGCAT
TGGCACATTGCAACCCATTCTGACGTTGGGCATTCCCTTTCCATCACAT
CCTACAAGAAATAGACTTTCCTTCCAGGCTCAATGATGTGGCATGTTCCAT
TCTCCGACTAAAAGGCGACATTCACACTTACGAGGCTGAGAGGAGCCGTG
GAGAAAAATCTTCGTGTATATCATGTTATATGGAAGAGAATCCCGAGTCA
ACAGAGGAAGATGCAATCAATCATATCAACTCCATGGTCGACAAATTACT
CAAGGAACTAAATTGGGAGTATCTGAGACCTGATAGCAATGTTCCAATCA
CTTCCAAGAAACATGCATTTGACATTCTGAGAGCTTTCTACCATCTCTACA
AATACCGAGATGGCTTCAGCGTTGCGAACTATGAAATAAAGAATTTGGTC
ATGACAACCGTCATTGAGCCTGTGCCTTTATAA- (SEQ ID NO:11)

P._coulteri_MBO-----
AGACGCATAGCTGGTCATCATTCCAACCTYTGGGACGATRATTTSATACAG
TCCCTYTCAACGCCTTATGGGGCAATTTTCGTACCATGAAAGTGCTCAGAA
ACTTATTGGAGAAGTAAAAGAGATGATCAATTCAATCTCGGTTAAAGATG
GAGAATTAATCACCCCCCTCCAATGATCTCCTTATGCGGCTCTCTATAGTTG
ATAGCATTGAACGTTTGGGAATCGATAGGCATTTCAAAAGTGAAATAAAA
TCAGCTCTGGATTATGTTTACAGTTATTGGAACGAAAAAGGCATTGGGTG
GGGAAGAGATAGTGTTGTTGCCGATCTCAACTCAACTGCCTTGGGGGCTTC
GAACTCTACGACTACACGGATACCCGGTGTCTTCAGATGTGTTACAACACT
TCAAAGAACAAAAAGGGCAGTTTGCATGTTCCGGCCATTCAAACAGAGGGA
GAGATAAGAAGTGTTCTCAACTTATTTTCGGGCTTCCCAAATTGCCTTTCCG
GGAGAGAAAGTTATGGAAGAGGCAGAAGTCTTCTCTACAATATATTTAAA
AGAAGCCATACTAAAGCTTCCGGTCTGCGGTCTTTACGAGAGATATCGT
ACGTTCTGGAATATGGTTGGCATATAAATTTGCCAAGATTGGAAGCAAGG
AACTACATCGACGTATTTGGAGAGGACCCCATTTATTTGACGCCAAATAT
GAAGACCCAAAAACTTCTAGAACTTGCAAAGTTGGAGTTCAATATGTTTC
ACTCTTTACAACAGCAAGAGCTAAAGCTTCTCTCCAGATGGTGGAAGAT
TCGGGTTTCTCTCAAATGACCTTCCCTCGGCATCGTCACGTGGAATATTAC
ACTTTGGCATCTTGCATTGATAGTGAACCTCAACATTCTTCGTTTCAGACTT
GGATTTGCCAAAATCTTTCATCTTGCCACGGTTCTTGACGATATTTACGAC
ACCTTTGGCACGATGGATGAGCTAGAACTCTTCACGGCGGCAGTTAAGAG
GTGGCATCCGTCTGCGACGGAGTGGCTTCCAGAATATATGAAAGGAGTAT
ATATGGTGCTTTACGAAACCGTTAACGAAATGGCAGGAGAAGCAGAAAA
GTCTCAAGGCCGAGACACGCTCAACTATGGCCGAAATGCTTTGGAGGCTT
ATATTGATGCTTCTATGGAAGAAGCGAAGTGGATTTTCAGTGGTTTTTTGC
CAACATTTGAGGAGTACCTGGATAACGGGAAAGTTAGTTTCGGTTATGGC
ATTGGCACATTGCAACCCATTCTGACGTTGGGCATTCCCTTTCCATCAC
ATCTTACAAGAAATAGACTTTCCTTCCAGGCTCAATGATGTGGCATCTTCC
ATTCTCCGACTAAAAGGCGACATTCACACTTACCAGGCTGAGAGGAGCCG
TGGAGAAAAATCTTCGTGTATATCATGTTATATGGAAGAGAATCCCGAGT
CAACAGAGGAAGATGCAATCAATCATATCAACTCCATGGTCGACAAATTA
CTCAAGGAACTAAATTGGGAGTATCTGAGACCTGATAGCAATGTTCCAAT

FIG. 1I

10/20

CACTTCCAAGAAACATGCATTTGACATTCTGAGAGCTTTCTACCATCTCTA
CAAATACCGAGATGGCTTCAGCGTTGCGAACTATGAAATAAAGAATTTGG
TCATGACAACCGTCATTGAGCCTGTGCCTTTATAA- (SEQ ID NO:12)

*P. torreyana*_MBO-----

AACCTCTGGGACGATGATTTGATACAGTCCCTTTCAACGCCTTATGGGGCA
ATTTTCGTACCATGAAAGTGCTCAGAACTTATTGGAGAAGTAAAAGAGAT
GATCAATTCAATCTCGGTTAAAGATGGAGAATTAATCACCCCCTCCAATG
ATCTCCTTATGCGGCTCTCTATAGTTGATAGCATTGAACGTTTGGGAATCG
ATAGGCATTTCAAAAGTGAAATAAAATCAGCTCTGGATTATGTTTACAGTT
ATTGGAACGAAAAAGGCATTGGGTGGGGAAGAGATAGTGTGTTGCCGAT
CTCAACTCAACTGCCTTGGGGCTTCGAACTCTACGACTACACGGATACCCG
GTGTCTTCAGATGTGTTACAACACTTCAAAGAACA AAAAAGGGCAGTTTGC
ATGTTTCGGCCATTCAAACAGAGGGGAGAGATAAGAAGTGTCTCAACTTAT
TTCGGGCTTCCCAAATTGCCTTTCCGGGAGAGAAAGTTATGGAAGAGGCA
GAAGTCTTCTCTACAATGTATTTAAAAGAAGCCATACTAAAGCTTCCGGTC
TGCGGTCTTTCACGAGAGATATCGTACGTTCTGGAATATGGTTGGCATATA
AATTTGCCAAGATTGGAAGCAAGGAACATCGACGTATTTGGACAGGA
CCCCATTTATTTGACGCCAAATATGAAGACCCAAAAACCTTCTAGA ACTTGC
AAAGTTGGAGTTCAATATGTTTCACTCTTTACAACAGCAAGAGCTAAAGC
TTCTCTCCAGATGGTGGAAAGATTCGGGTTTCTCTCAAATGACCTTCCCTC
GGCATCGTCACGTGGAATATTACACTTTGGCATCTTGCATTGATAGTGAAC
CTCAACATTTCTTCGTTTCACTTGGATTGTC AAAATCTTTCATCTTGGCAC
GGTTCTTGACGATATTTACGACACCTTTGGCACGATGGATGAGCTAGAACT
CTTCACGGCGGCAGTTAAGAGGTGGCATCCGTCTGCGACGGAGTGGCTTC
CAGAATATATGAAAGGAGTATATATGGTGCTTTACGAAACCGTTAACGAA
ATGGCAGGAGAAGCAGAAAAGTCTCAAGGCCGAGACACGCTCAACTATG
CCCGAAATGCTTTGGAGGCTTATATTGATGCTTCTATGGAAGAAGCGAAG
TGGATTTTCAGTGGTTTTTTTGCCAACATTTGAGGAGTACCTGGATAACGGG
AAAGTTAGTTTCGGTTATGGCATTTGGCACATTGCAACCCATTCTGACGTTG
GGCATTCCCTTTCCCTCATCACATCTACAAGAAATAGACTTTCCCTCCAGG
CTCAATGATGTGGCATCTTCCATTCTCCGACTAAAAGGCGACATTCACACT
TACCAGGCTGAGAGGAGCCGTGGAGAAAAATCTTCGTGTATATCATGTTA
TATGGAAGAGAATCCCGAGTCAACAGAGGAAGATGCAATCAATCATATCA
ACTCCATGGTCGACAAATTACTCAAGGAACTAAATTGGGAGTATCTGAGA
CCTGATAGCAATGTTCCAATCACTTCCAAGAAACATGCATTTGACATTCTG
AGAGCTTTCTACCATCTCTACAAATACCGAGATGGCTTCAGCGTTGCGAAC
TATGAAATAAAGAATTTGGTCATGACAACCGTCATTGAGCCTGTGCCTTTA
TAA- (SEQ ID NO:13)

*P. attenuata*_MBO-----

AGACGCAGAGGTGATTTCCATTCCAACCTCTGGGACGATAATTTCATACA
GTCCCTCTCAACGCCTTATGGGGCAATTTTCGTACCATGAAAGTGCTCAGAA
ACTTATTGGAGAAGTAAAAGAGATGATCAATTCAATCTCGGATAAAGATG
GAGAATTAATCACCCCCTCCAATGATCTCCTTATGCTGCTCTCTATAGTCG
ATAGCATTGAACGTTTGGGAATCGATAGGCATTTCAAAAGTGAAATAAAA
TCAGCTCTGGATTATGTTTACAGTTATTGGAACGAAAAAGGCATTGGGTG
GGGAAGAGATAGTGTGTTGCCGATCTCAACTCAACTGCCTTGGGGCTTC
GAACTCTACGACTACACGGATACCCGGTGTCTTCAGACGTGTTACAACAC
TTCAAAGAACA AAAATGGGCAGTTTGCATGTTTCGGCCATTCAAACAGAGGG

FIG. 1J

11/20

AGAGATAAGAAGTGTTCTCAACTTATTTTCGGGCTTCCCAAATTGCCTTTCC
GGGAGAGAAAGTTATGGAAGAGGCAGAAGTCTTCTCTACAAAATATTTAA
AAGAAGCCATACTAAAGCTTCCGGTCTGCGGTCTTTCACGAGAGATATCG
TACGTTCTGGAATATGGTTGGCATATGAATTTGCCAAGATTGGAAGCAAG
GAACTACATCGACGTATTTGGACAGGACCCCATTTATTTGACGCCAAATAT
GAAGACCCAAAACTTCTAGAACTTGCAAAGTTGGAGTTCAATATGTTTC
ACTCTTTACAACAGCAAGAGCTAAAGCTTCTCTCCAGATGGTGGAAAGAT
TCGGGTTTCTCTCAAATGACCTTCCCTCGGCATCGTCACGTGGAATATTAC
ACTTTGGCATCTTGCATTGATAGTGAACCTCAACATTCTTCGTTACAGACTT
GGATTTGCCAAAATATTTTCATCTTGCCACGGTTCTTGACGATATTTACGAC
ACCTTTGGCACGATGGATGAGCTAGAACTCTTCACAGCGGCAGTTAAGAG
GTGGCATCCGTCTGCGACGGAGTGGCTTCCAGAATATATGAAAGGAGTAT
ATATGGTGCTTTACGAAACCGTTAACGAAATGGCAGGAGAAGCAGAAAA
GTCTCAAGGCCGAGACACGCTCAACTATGCCCCGAAATGCTTTGGAGGCTT
ATATTGATGCTTCTATGGAAGAAGCGAAGTGGATTTTCAGTGGTTTTTTGC
CAACATTTGAGGAGTACCTGGATAACGGGAAAGTTAGTTTCGGTTATACC
ATTGGCACATTGCAACCCATTCTGACGTTGGGCATTCCCTTTCCTCATCAC
ATCCTACAAGAAATAGACTTTCCTTCCAGGCTCAATGATGTGGCATGTTCC
ATTCTCCGACTAAAAGGCCGACGTTACACTTACCAGCCTGAGAGGAGCCG
TGGAGAAGAATCTTCGTGTATATCATGTTATATTGAAGAGAATCCCGAGT
CAACAGAGGAAGATGCAATCAATCATATCAACTCCATGGTCGACAAATTA
CTCAAGGAACTAAATTGGGAGTATCTGAGACCTGATAGCAATGTTCCAAT
CACTTCCAAGAAACATGCATTTGACATTCTGAGAGCTTTGTACCATCTCTA
CAAATACCGAGATGGCTACAGCGTTGCGAACTATGAAATAAAGAATTTGG
TCATGACAACCGTCATTGAGCCTGTGCCTTTATAA- (SEQ ID NO:14)

P._radiata_MBO-----
AACCTCTGGGACGATGATTTGATACAGTCCCTTTCAACGCCTTATGGGGCA
ATTTTCGTACCATGAAAGTGCTCAGAACTTATTGGAGAAGTAAAAGAGAT
GATCAATTCAATCTCGGATAAAAATGGAGAATTAATCACCCCCCTCCAATG
ATCTCCTTATGCGGCTCTCTATAGTCGATAGCATTGAACGTTTGGGAATCG
ATAGGCATTTCAAAAAGTGAAATAAAAATCAGCTCTGGATTATGTTTACAGTT
ATTGGAACGAAAAAGGCATTGGGTGGGGAAGAGATAGTGTGTTGCCGAT
CTCAACTCAACTGCCTTGGGGCTTCGAACTCTACGACTACACGGATACCCG
GTGTCTTCAGACGTGTTACAACACTTCAAAGAACAAAATGGGCAGTTTGC
ATGTTCCGGCCATTCAAACAGAGGGAGAGATAAGAAGTGTTCTCAACTTAT
TTCGGGCTTCCCAAATTGCCTTTCGGGAGAGAGAAAGTTATGGAAGAGGCA
GAAGTCTTCTCTACAAAATATTTAAAAGAAGCCATACTAAAGCTTCCGGT
CTGCGGTCTTTCACGAGAGATATCGTACGTTCTGGAATATGGTTGGCATAT
GAATTTGCCAAGATTGGAAGCAAGGAACATCGACGTATTTGGACAGG
ACCCCATTTATTTGACGCCAAATATGAAGACCCAAAACTTCTAGAACTT
GCAAAGTTGGAGTTCAATATGTTTCACTCTTTACAACAGCAAGAGCTAAA
GCTTCTCTCCAGATGGTGGAAAGATTTCGGGTTTCTCTCAAATGACCTTCCC
TCGGCATCGTCACGTGGAATATTACACTTTGGCATCTTGCATTGATAGTGA
ACCTCAACATTCTTCGTTACAGACTTGGATTTGCCAAAATATTTTCATCTTGC
CACGGTTCTTGACGATATTTACGACACCTTTGGCACGATGGATGAGCTAG
AACTCTTCACAGCGGCAGTTAAGAGGTGGCATCCGTCTGCGACGGAGTGG
CTTCCAGAATATATGAAAGGAGTATATATGGTGCTTTACGAAACCGTTAA
CGAAATGGCAGGAGAAGCAGAAAAGTCTCAAGGCCGAGACACGCTCAAC
TATGCCCCGAAATGCTTTGGAGGCTTATATTGATGCTTCTATGGAAGAAGCG

FIG. 1K

SUBSTITUTE SHEET (RULE 26)

12/20

AAGTGGATTTTCAGTGGTTTTTTTGCCAACATTTGAGGAGTACCTGGATAAC
GGGAAAGTTAGTTTCGGTTATAACCATTTGGCACATTGCAACCCATTCTGACG
TTGGGCATTCCCTTTTCCTCATCACATCCTACAAGAAATAGACTTTCCTTCC
AGGCTCAATGATGTGGCATGTTCCATTCTCCGACTAAAAGGCGACGTTCA
CACTTACCAGCCTGAGAGGAGCCGTGGAGAAGAATCTTCGTGTATATCAT
GTTATATGGAAGAGAATCCCCGAGTCAACAGAGGAAGATGCAATCAATCAT
ATCAACTCCATGGTCGACAAATTACTCAAGGAACTAAATTGGGAGTATCT
GAGACCTGATAGCAATGTTCCAATCACTTCCAAGAAACATGCATTTGACA
TTCTGAGAGCTTTGTACCATCTCTACAAATACCGAGATGGCTACAGCGTTG
CGAACTATGAAATAAAGAATTTGGTCATGACAACTGTCATTGAGCCTGTG
CCTTTATAA- (SEQ ID NO:15)

P._contorta_MBO1

TCTAGCAAGGTTAAGGTTGTCCGCAGAACGATGTCAACTTCCATCCGCAT
GTGTCAGATAACCACTGTATCCGGTGAAGGCGTACAGAGACGCATAGCAA
ATCATCATTTCCAACCTCTGGGACGATAATTTTCATACAGTCCCTCTCAACGC
CTTATGGGGCAATTTTCGTACCATGAAAGTGCTCAGAACTTATTGGAGAA
GTAAAAGAGATGATCAATTCAATCTCGGTTAAAGATGGAGAATTAATCAC
CCCCITCCAATGATCTCCTTATGCGGCTCTCTATAGTCGATAGCATTGAACG
TTTGGGAATCGATAGGCATTTCAAAAGTGAAATAAAATCAGCTCTGGATT
ATGTTTACAGTTATTGGAACGAAAAAGGCATTGGGTGGGGAAGAGATAGT
GTTGTTGCCGATCTCAACTCAACTGCCTTGGGGCTTCGAACTCTACGACTA
CACGGATACCCGGTGTCTTCAGACGTGTTACAACACTTCAAAGAACAAAA
AGGGCAGTTTGCATGTTCCGGCCATTCAAACAGAGGGAGAGATAAGAAGTG
TTCTCAACTTATTTCCGGGCTTCCCAAATTGCCTTTCCGGGAGAGAAAGTTA
TGGAAGAGGCAGAAAGTCTTCTCTACAAAATATTTAAAAGAAGCCATACTA
AAGCTTCCGGTCTGCGGTCTTTACGAGAGATATCGTACGTTCTGGAATAT
GGTTGGCATATGAATTTGCCAAGATTGGAAGCAAGGAACTACATCGACGT
ATTTGGACAGGACCCCATTTATTTGACGCCAAATATGAAGACCCAAAAAC
TTCTAGAACTTGCTAAGTTGGAGTTCAATATGTTTCACTCTTTACAACAGC
AAGAGCTAAAGCTTCTCTCCAGATGGTGGAAGATTCCGGGTTTCTCTCAA
ATGACCTTCCCTCGGCATCGTCACGTGGAATATTACACTTTGGCATCTTGC
ATTGATAGTGAACCTCAACATTTCTTCGTTACAGACTTGGATTTGCCAAAATA
TTTCATCTTGCCACGGTTCTTGACGATATTTACGACACCTTTGGCACGATG
GATGAGCTAGAACTCTTCACAGCGGCAGTTAAGAGGTGGCATCCGTCTGC
GACGGAGTGGCTTCCAGAATATATGAAAGGAGTATATATGGTGCTTTACG
AAACCGTTAACGAAATGGCAGGAGAAGCAGAAAAGTCTCAAGGCCGAGA
CACGCTCAACTATGCCCCGAAATGCTTTGGAGGCTTATATTGATGCTTCTAT
GGAAGAAGCGAAGTGGATTTTCAGTGGTTTTTTTGCCAACATTTGAGGAGT
ACCTGGATAACGGGAAAGTTAGTTTCGGTTATAACCATTTGGCACATTGCAA
CCCATTTCTGACGTTGGGCATTCCCTTTTCCTCATCACATCCTACAAGAAATA
GACTTTTCCTTCCAGGCTCAATGATGTGGCATGTTCCATTCTCCGACTAAAA
GGCGACATTCACACTTACCAGCCTGAGAGGAGCCGTGGAGAAGAATCTTC
GTGTATATCATGTTATATGGAAGATAATCCCGAGTCAACAGAGGAAGATG
CAATCAATCATATCAACTCCATGGTCGACAAATTACTCAAGGAACTAAAT
TGGGAGTATCTGAGACCTGATAGCAATGTTCCAATCACTTCCAAGAAACA
TGCATTTGACATTCTGAGAGCTTTCTACCATCTCTACAAATACCGAGATGG
CTACAGCGTTGCGAACTATGAAATAAAGAATTTGGTCATGACAACTGTCA
TTGAGCCTGTGCCTTTATAA- (SEQ ID NO:16)

FIG. 1L

P._contorta_MBO2

TCTAGCAAGGTTAAGGTTGTCCGCAGAACGATGTCAACTTCCATCCGCAT
GTGTCAGATAACCACTGTATCCGGTGAAGGCGTACAGAGACGCATAGCAA
ATCATCATTCCAACCTCTGGGACGATAATTTTCATACAGTCCCTCTCAACGC
CTTATGGGGCAATTTTCGTACCATGAAAGTGCTCAGAAACTTATTGGAGAA
GTAAAAGAGATGATCAATTCAATCTCGGTAAAGATGGAGAATTAATCAC
CCCGTCCAATGATCTCCTTATGCGGCTCTCTATAGTCGATAGCATTGAACG
TTTGGGAATCGATAGGCATTTCAAAAAGTGAAATAAAAATCAGCTCTGGATT
ATGTTTACAGTTATTGGAACGAAAAAGGCATTGGGTGGGGAAGAGATAGT
GTTGTTGCCGATCTCAACTCAACTGCCTTGGGGCTTCGAACTCTACGACTA
CACGGATACCCGGTGTCTTCAGACGTGTTACAACACTTCAAAGAACAAAA
AGGGCAGTTTGCATGTTCCGGCCATTCAAACAGAGGGAGAGATAAGAAGTG
TTCTCAACTTATTTTCGGGCTTCCCAAATTGCCTTTCCGGGAGAGAAAGTTA
TGGAAGAGGCAGAAAGTCTTCTCTACAAAATATTTAAAAGAAGCCATACTA
AAGCTTCCGGTCTGCGGTCTTTCACGAGAGATATCGTACGTTCTGGAATAT
GGTTGGCATATAAAATTTGCCAAGATTGGAAGCAAGGAACCTACATCGACGT
ATTTGGACAGGACCCCATTTATTTGACGCCAAATATGAAGACCCAAAAAC
TTCTAGAACTTGCAAAGTTGGAGTTCAATATGTTTCACTCTTTACAACAGC
AAGAGCTAAAGCTTCTCTCCAGATGGTGGAAAGATTCCGGGTTTCTCTCAA
ATGACCTTCCCTCGGCATCGTCACGTGGAATATTACACTTTGGCATCTTGC
ATTGATAGTGAACCTCAACATTCTTCGTTTCAGACTTGGATTGCCAAAATA
TTTCATCTTGCCACGGTTCTTGACGATATTTACGACACCTTTGGCACGATG
GATGAGCTAGAACTCTTCACAGCGGCAGTTAAGAGGTGGCATCCGTCTGC
GACGGAGTGGCTTCCAGAATATATGAAAGGAGTATATATGGTGCTTTACG
AAACCGTTAACGAAATGGCAGGAGAAGCAGAAAAGTCTCAAGGCCGAGA
CACGCTCAACTATGCCCGAAATGCTTTGGAGGCTTATATTGATGCTTCTAT
GGAAGAAGCGAAGTGGATTTTCAGTGGTTTTTTTGCCAACATTTGAGGAGT
ACCTGGATAACGGGAAAGTTAGTTTCGGTTATACCATTGGCACATTGCAA
CCCATTTCTGACGTTGGGCATTCCCTTTTCCTCATCACATCCTACAAGAAATA
GACTTTTCCTTCCAGGCTCAATGATGTGGCATGTTCCATTCTCCGACTAAAA
GGCGACGTTACACTTACCAGGCTGAGAGGAGCCGTGGAGAAGAATCTTC
GTGTATATCATGTTATATGGAAGATAATCCCGAGTCAACAGAGGAAGATG
CAATCAATCATATCAACTCCATGGTTCGACAAATTACTCAAGGAACATAAT
TGGGAGTATCTGAGACCTGATAGCAATGTTCCAATCACTTCCAAGAAACA
TGCATTTGACATTCTGAGAGCTTTCTACCATCTCTACAAATACCGAGATGG
CTACAGCGTTGCGAACTATGAAATAAAGAATTTGGTCATGACAACTGTCA
TTGAGCCTGTGCCTTTATAA-(SEQ ID NO:17)

FIG. 1M

P._patula_MBO_translation_frame_+1

SSKVKVVRTISTSIRMCRIITVSGEGVQRRRIANHHSNLWDDNFIQSLSTPYGA
ISYHESAQKLIGEVKEMINSISVKDGELITPSNDLLMRLSIVDSIERLGIDRHFKS
EIKSALDYVYSYWNEKGIGWGRDSVVADLNSTALGLRTLRLHGYPVSSDVL
QHFKEQKGQFACSAIQTEGEIRSVNLFRASQIAFPGEKVMEEAEVFSTIYLKE
AILKLPVCGLSREISYVLEYGWHINLPRLEARNYIDVFGQDPIYLTPNMKTQKL
LELAKLEFNMFHSLQQELKLLSRWWKDSGFSQMTFPRHRHVEYYTLASCID
SEPQHSSFRLGFAKIFHLATVLDDIYDTFGTMDELELFTA AVERWHPSATEWL
PEYMKGVYMVLYETVNEMAGEAEKSQGRDTLNYARNALEAYIDASMKEAK
WIFSGFLPTFEEYLDNGKVSFGYSIGTLQPILTLGIPFPHHILQEIDFPSRLNDVA
CSILRLKGDIHTYQAERSRGEKSSCISCYMEENPESTEEDAIDHINSMVDKLLK
ELNREYLRPDSNPITSKKHAFDILRAFYLHYKYRDGFSVANYEIKNLVMTTV
IEPVPL (SEQ ID NO:18)

P._pseudostrobus_estevzii_translation_frame_+1

SSKVKVVRTISTSIRMCRIITVSGEGVQRRRIANHHSNLWDDNFIQSLSTPYGA
ISYHESAQKLIGEVKEIINSISVKDGELITPSNDLLMRLSIVDSIERLGIDRHFKE
IKSALDYVYSYWNEKGIGWGRDSVVADLNSTALGLRTLRLHGYPVSSDVLQ
HFKEQKGQFACSAIQTEGEIRSVNLFRASQIAFPGEKVMEEAEVFSTIYLKEA
ILKLPVCGLSREISYVLEYGWHINLPRLEARNYIDVFGQDPIYLTPNMKTQKLL
ELAKLEFNMFHSLQQELKLLSRWWKDSGFSQMTFPRHRHVEYYTLASCIDS
EPQHSSFRLGFAKIFHLATVLDDIYDTFGTMDELELFTA AAVKRWHPSATEWLP
EYMKGVYMVLYETVNEMAGEAEKSQGRDTLNYARNALEAYIDASMKEAK
WIFSGFLPTFEEYLDNGKVSFGYSIGTLQPILTLGIPFPHHILQEIDFPSRLNDVA
CSILRLKGDIHTYQAERSRGEKSSCISCYMEENPESTEEDAINHVNSMVDKLL
KELNWEYLRPDSNPITSKKHAFDILRAFYLHYKYRDGFSVANYEIKNLVMT
TVIEPVPL (SEQ ID NO:19)

P._cooperi_ornelasi_MBO_translation_frame_+1

SSKVKVVRTISTSIRMCRIITVSGEGVQRRRIANHHSNLWDDNFIQSLSTPYGA
ISYHESAQKLIGEVKEMINSISVKDGELITPSNDLLMRLSIVDSIERLGIDRHFKS
EIKSALDYVYSYWNEKGIGWGRDSVVADLNSTALGLRTLRLHGYPVSSDVL
QHFKEQKGQFACSAIQTEGEIRSVNLFRASQIAFPGEKVMEEAEVFSTIYLKE
AILKLPVCGLSREISYVLEYGWHINLPRLEARNYIDVFGQDPIYLTPNMKTQKL
LELAKLEFNMFHSLQQELKLLSRWWKDSGFSQMTFPRHRHVEYYTLASCID
SEPQHSSFRLGFAKIFHLATVLDDIYDTFGTMDELELFTA AAVKRWHPSATEWL
PEYMKGVYMVLYETVNEMAGEAEKSQGRDTLNYARNALEAYIDASMKEAK
WIFSGFLPTFEEYLDNGKVSFGYSIGTLQPILTLGIPFPHHILQEIDFPSRLNDVA
CSILRLKGDIHTYQAERSRGEKSSCISCYMEENPESTEEDAINHINSMVDKLLK
ELNWEYLRPDSNPITSKKHAFDILTAFYLHYKYRDGFSVANYEIKNLVMTT
VIEPVPL (SEQ ID NO:20)

P._hartwegii_MBO_translation_frame_+1

SSKVKVVRTISTSIRMCRIITVSGEGVQRRRIANHHSNLWDDNFIQSLSTPYGA
ISYHESAQKLIGEVKEMINSISVKDGELITPSNDLLMRLSIVDSIERLGIDRHFKE
EIKSALDYVYSYWNEKGIGWGRDSVVADLNSTALGLRTLRLHGYPVSSDVL
QHFKEQKGQFACSAIQTEGEIRSVNLFRASQIAFPGEKVMEEAEVFSTIYLKE

FIG. 2A

15/20

AILKL PVCGLSREISYVLEYGWHINLPRLEARNYIDVFGQDPIYLTPNMKTQKL
 LELAKLEFNMFHSLQQQELKLLSRWWKDSGFSQMTFPRHRHVEYYTLASCID
 SEPQYSSFR LGFAKIFHLATVLDDIYDTFGTMDELELFTA AVKRWHP SATEWL
 PEYMKGVYMVLYETVNEMAGEAEKSQGRDTLNYARNALEAYIDASMKEAK
 WIFSGFLPTFEEYLDNGKVSFGYSIGTLQPILTLGIPFPHHILQEIDFPSRLNDVA
 CSILRLKGDIHTYQAERSRGEKSSCISCYMEENPESTEEDAINHINS MVDKLLK
 ELNWEYL RPDSNPITSSKKHAFDILRAFYLHYKYRDGFSVANYEIKNLVMTT
 VIEPVPL (SEQ ID NO:21)

P._ponderosa_MBO1_translation_frame_+1

SSKVKVVRTISTIRMC RITTESGEGVQRR IANHH SNLWDDNFIQSLSTPYGA
 ISYHESAQKLIGEVKEMINSISLKDGELITPSNDLLMRLSIVDSIERLGIDRHFKS
 EIKSALDYVYSYWNEKGIGWGRDSV VADLNSTALGLRTLRLHGYPVSSDVL
 QHFKEQKGQFACSAIQTEGEIRSVNLFRASQIAFPGEKVMEEAEVFSTIYLKE
 AILKL PVCGLSREISYVLEYGWHINLPRLEARNYIDVFGQDPIYLTPNMKTQKL
 LELAKLEFNMFHSLQQQELKLLSRWWKDSGFSQMTFPRHRHVEYYTLASCID
 SEPHSSFR LGFAKIFHLATVLDDIYDTFGTMDELELFTA AVKRWHP SATEWL
 PEYMKGVYMVLYETVNEMAREAEKSQGRDTLNYARNALEAYIDASMKEAK
 WIFSGFLPTFEEYLDNGKVSFGYSIGTLQPILTLGIPFPHHILQEIDFPSRLNDVA
 CSILRLKGDIHTYQAERSRGEKSSCISCYMEENPESTEEDAINHINS MVDKLLK
 ELNWEYL RPDSNPITSSKKHAFDILRTFYHLYKYRDGFSVANYEIKNLVMTT
 VIEPVPL (SEQ ID NO:22)

P._ponderosa_MBO2_translation_frame_+1

SSKVKVVRTISTIRMC RITTESGEGVQRR IANHH SNLWDDNFIQSLSTPYGA
 ISYHESAQKLIGEVKEMINSISLKDGELITPSNDLLMRLSIVDSIERLGIDRHFKS
 EIKSALDYVYSYWNEKGIGWGRDSV VADLNSTALGLRTLRLHGYPVSSDVL
 QHFKEQKGQFACSAIQTEGEIRSVNLFRASQIAFPGEKVMEEAEVFSTIYLKE
 AILKL PVCGLSREISYVLEYGWHINLPRLEARNYIDVFGQDPIYLTPNMKTQKF
 LELAKLEFNMFHSLQQQELKLLSRWWKDSGFSQMTFPRHRHVEYYTLASCID
 SEPHSSFR LGFAKIFHLATVLDDIYDTFGTMDELELFTA AVKRWHP SATEWL
 PEYMKGVYMVLYETVNEMAREAEKSQGRDTLNYARNALEAYIDASMKEAK
 WIFSGFLPTFEEYLYNGKVSFGYSIGTLQPILTLGIPFPHHILQEIDFPSRLNDVA
 CSILRLKGDIHTYQAERSRGEKSSCISCYMEENPESTEEDAINHINS MVDKLLK
 ELNWEYL RPDSNPITSSKKHAFDILRTFYHLYKYRDGFSVANYEIKNLVMTT
 VIEPVPL (SEQ ID NO:23)

P._jeffreyi_MBO_translation_frame_+1

SSKVKVVRTISTIRMC RITTESGEGVQRR IANHH SNLWDDNFIQSLSTPYGA
 ISYHESAQKLIGEVKEMINSISLKDGELITPSNDLLMRLSIVDSIERLGIDRHFKS
 EIKSALDYVYSYWNEKGIGWGRDSV VADLNSTALGLRTLRLHGYPVSSDVL
 QHFKEQKGQFACSAIQTEGEIRSVNLFRASQIAFPGEKVMEEAEVFSTIYLKE
 AILKL PVCGLSREISYVLEYGWHINLPRLEARNYIDVFGQDPIYLTPNMKTQKL
 LELAKLEFNMFHSLQQQELKLLSRWWKDSGFSQMTFPRHRHVEYYTLASCID
 SEPHSSFR LGFAKIFHLATVLDDIYDTFGTMDELELFTA AVKRWHP SATEWL
 PEYMKGVYMVLYETVNEMAGEAEKSQGRDTLNYARNALEAYIDASMKEAK
 WIFSGFLPTFEEYLDNGKVSFGYSIGTLQPILTLGIPFPHHILQEIDFPSRLNDVA
 CSILRLKGDIHTYQAERSRGEKSSCISCYMEENPESTEEDAINHINS MVDKLLK
 ELNWEYL RPDSNPITSSKKHAFDILRAFYLHYKYRDGFSVANYEIKNLVMTT
 VIEPVPL (SEQ ID NO:24)

FIG. 2B

P._montezumae_MBO_translation_frame_+1

SSKVKVVRTISTSRMCRITTVSGEGVQRRRIANHHSNLWDDNFIQSLSTPYGA
ISYHESAQKLIGEVKEMINSISVKDGELITPSNDLLMRLSIVDSIERLGIDRHFKS
EIKSALDYVYSYWNEKGIGWGRDSVVADLNSTALGLRTLRLHGYVPVSSDVL
QHFKEQKGQFACSAIQTEGEIRSVNLFRASQIAFPGEKVMEEAEVFSTIYLKE
AILKLPVCGLSREISYVLEYGWHINLPRLEARNYIDVFGQDPIYLTPTNMKTQKL
LELAKLEFNMFHSLQQQELKLLSRWWKDSGFSQMTFPRHRHVEYYTLASCID
SEPQHSSFRLGFAKMFHLATVLDIYDTFGTMDELELFTA AVKRWHP SATEW
LPEYMKGVYVMVLYETVNEMAGEAEKSQGRDTLNYARNALEAYIDASMKEA
KWIFSGFLPTFEEYLDNGKVSFGYSIGTLQPILTLGIPFPHILQEIDFPSRLNDV
ACSILRLKGDIIHTYEAERSRGEKSSCISCYMEENPESTEEDAINHINSMDVKLL
KELNWEYLRPDSNPITSSKKHAFDILRAFYLHYKYRDGFSVANYEIKNLVMT
TVIEPVPL (SEQ ID NO:25)

P._pseudostrobus_MBO_translation_frame_+1

RRXXXXHNSNLWDDNFIQSLSTPYGAISYHESAQKLIGEVKEMINSISVKDGELI
TPSNDLLMRLSIVDSIERLGIDRHFKSEIKSALDYVYSYWNEKGIGWGRDSVV
ADLNSTALGLRTLRLHGYVPVSSDVLQHFKEQKGQFACSAIQTEGEIRSVNLFR
RASQIAFPGEKVMEEAEVFSTIYLKEAILKLPVCGLSREISYVLEYGWHINLPR
LEARNYIDVFGQDPIYLTPTNMKTQKLLELAKLEFNMFHSLQQQELKLLSRWW
KDSGFSQMTFPRHRHVEYYTLASCIDSEPQHSSFRLGFAKMFHLATVLDIYD
TFGTMDLELFTA AVKRWHP SATEWLPEYMKGVYVMVLYETVNEMAGEAEK
SQGRDTLNYARNALEAYIDASMKEAKWIFSGFLPTFEEYLDNGKVSFGYSIGT
LQPILTLGIPFPHILQEIDFPSRLNDVACSILRLKGDIIHTYEAERSRGEKSSCISC
YMEENPESTEEDAINHINSMDVKLLKELNWEYLRPDSNPITSSKKHAFDILRA
FYHLYKYRDGFSVANYEIKNLVMTTVIEPVPL (SEQ ID NO:26)

P._sabiniana_MBO1_translation_frame_+1

SSKVKVVRTISTSRMCRITTESGEGVQRRRIANHHSNLWDDNFIQSLSTPYGA
ISYHESAQKLIGEVKEMINSISVKDGELITPSNDLLMRLSIVDSIERLGIDRHFKS
EIKSALDYVYSYWNEKGIGWGRDSVVADLNSTALGLRTLRLHGYVPVSSDVL
QHFKEQKGQFACSAIQTEGEIRSVNLFRASQIAFPGEKVMEEAEVFSTIYLKE
AILKLPVCGLSREISYVLEYGWHINLPRLEARNYIDVFGEDPIYLTPTNMKTQKL
LELAKLEFNMFHSLQQQELKLLSRWWKDSGFSQMTFPRHRHVEYYTLASCID
SEPQHSSFRLGFAKIFHLATVLDIYDTFGTMDELELFTA AVKRWHP SATEWL
PEYMKGVYVMVLYETVNEMAGEAEKSQGRDTLNYGRNALEAYIDASMEEAK
WIFSGFLPTFEEYLDNGKVSFGYGIGTLQPILTLGIPFPHILQEIDFPSRLNDVA
SSILRLKGDIIHTYQAERSRGEKSSCISCYMEENPESTEEDAINHINSMDVKLLK
ELNWEYLRPDSNPITSSKKHAFDILRAFYLHYKYRDGFSVANYEIKNLVMTT
VIEP--- (SEQ ID NO:27)

P._sabiniana_MBO2_translation_frame_+1

SSKVKVVRTISTSRMCRITTESGEGVQRRRIANHHSNLWDDNFIQSLSTPYGA
ISYHESAQKLIGEVKEMINSISLKDGEITPSNDLLMRLSIVDSIERLGIDRHFKS
EIKSALDYVYSYWNEKGIGWGRDSVVADLNSTALGLRTLRLHGYVPVSSDVL
QHFKEQKGQFACSAIQTEGEIRSVNLFRASQIAFPGEKVMEEAEVFSTIYLKE
AILKLPVCGLSREISYVLEYGWHINLPRLEARNYIDVFGEDPIYLTPTNMKTQKL
LELAKLEFNMFHSLQQQELKLLSRWWKDSGFSQMTFPRHRHVEYYTLASCID
SEPQHSSFRLGFAKIFHLATVLDIYDTFGTMDELELFTA AVKRWHP SATEWL

FIG. 2C

17/20

PEYMKGVYMVLYETVNMAGEAEKSQGRDTLNYGRNALEAYIDASMEEAK
WIFSGFLPTFEEYLDNGKVSFGYGIGTLQPILTLGIPFPHHILQEIDFPSRLNDVA
SSILRLKGDIIHTYQAERSRGEKSSCISCYMEENPESTEEDAINHINSMDVKLLK
ELNWEYLRPDSNPITSSKKHAFDILRAFYLKYRDLGFSVANYEIKNLVMTT
VIEPVPL (SEQ ID NO:28)

P._coulteri_MBO_translation_frame_+1

RRIAGHHSNLWDDDBXIQLSTPYGAISYHESAQKLIGEVKEMINSISVKDGELI
TPSNDLLMRLSIVDSIERLGIDRHFSEIKSALDYVYSYWNEKGIGWGRDSVV
ADLNSTALGLRTLRLHGYPVSSDVLQHFKEQKQGFACSAIQTEGEIRSVLNLF
RASQIAFPGEKVMEEAEVFSTIYLKEAILKLPVCGLSREISYVLEYGWHINLPR
LEARNYIDVFGEDPIYLTNPMKTQKLLELAKLEFNMFHSLQQQELKLLSRWW
KDSGFSQMTFPRHRHVEYYTLASCIDSEPQHSSFRIGFAKIFHLATVLDIYDT
FGTMDELELFTA AVKRWHPSATEWLPEYMKGVYMVLYETVNMAGEAEKS
QGRDTLNYGRNALEAYIDASMEEAKWIFSGFLPTFEEYLDNGKVSFGYGIGTL
QPILTLGIPFPHHILQEIDFPSRLNDVASSILRLKGDIIHTYQAERSRGEKSSCISC
YMEENPESTEEDAINHINSMDVKLLKELNWEYLRPDSNPITSSKKHAFDILRA
FYHLYKYRDLGFSVANYEIKNLVMTTVIEPVPL (SEQ ID NO:29)

P._torreyana_MBO_translation_frame_+1

NLWDDDLIQLSTPYGAISYHESAQKLIGEVKEMINSISVKDGELITPSNDLLM
RLSIVDSIERLGIDRHFSEIKSALDYVYSYWNEKGIGWGRDSVVADLNSTAL
GLRTLRLHGYPVSSDVLQHFKEQKQGFACSAIQTEGEIRSVLNLFASQIAFPG
EKMEEAEVFSTMYLKEAILKLPVCGLSREISYVLEYGWHINLPRLEARNYID
VFGQDPIYLTNPMKTQKLLELAKLEFNMFHSLQQQELKLLSRWWKDSGFSQ
MTFPRHRHVEYYTLASCIDSEPQHSSFRIGFAKIFHLATVLDIYDTFGTMDE
LELFTA AVKRWHPSATEWLPEYMKGVYMVLYETVNMAGEAEKSQGRDTL
NYARNALEAYIDASMEEAKWIFSGFLPTFEEYLDNGKVSFGYGIGTLQPILTL
GIPFPHHILQEIDFPSRLNDVASSILRLKGDIIHTYQAERSRGEKSSCISCYMEEN
PESTEEDAINHINSMDVKLLKELNWEYLRPDSNPITSSKKHAFDILRAFYLKY
RDLGFSVANYEIKNLVMTTVIEPVPL (SEQ ID NO:30)

P._attenuata_MBO_translation_frame_+1

RRRGDFHSNLWDDNFIQLSTPYGAISYHESAQKLIGEVKEMINSISDKDGELI
TPSNDLLMLLSIVDSIERLGIDRHFSEIKSALDYVYSYWNEKGIGWGRDSVV
ADLNSTALGLRTLRLHGYPVSSDVLQHFKEQNGQGFACSAIQTEGEIRSVLNLF
RASQIAFPGEKVMEEAEVFSTKYLKEAILKLPVCGLSREISYVLEYGWHMNLP
RLEARNYIDVFGQDPIYLTNPMKTQKLLELAKLEFNMFHSLQQQELKLLSRW
WKDSGFSQMTFPRHRHVEYYTLASCIDSEPQHSSFRIGFAKIFHLATVLDIY
DTFGTMDELELFTA AVKRWHPSATEWLPEYMKGVYMVLYETVNMAGEAE
KSQGRDTLNYARNALEAYIDASMEEAKWIFSGFLPTFEEYLDNGKVSFGYTIG
TLQPILTLGIPFPHHILQEIDFPSRLNDVACSILRLKGDVHTYQPERSRGEESSCI
SCYIEENPESTEEDAINHINSMDVKLLKELNWEYLRPDSNPITSSKKHAFDILR
ALYHLYKYRDLGYSVANYEIKNLVMTTVIEPVPL (SEQ ID NO:31)

P._radiata_MBO_translation_frame_+1

NLWDDDLIQLSTPYGAISYHESAQKLIGEVKEMINSISDKNGELITPSNDLLM
RLSIVDSIERLGIDRHFSEIKSALDYVYSYWNEKGIGWGRDSVVADLNSTAL
GLRTLRLHGYPVSSDVLQHFKEQNGQGFACSAIQTEGEIRSVLNLFASQIAFPG
EKMEEAEVFSTKYLKEAILKLPVCGLSREISYVLEYGWHMNLPRLARNYI

FIG. 2D

18/20

DVFGQDPIYLTPNMKTQKLELAKLEFNMFHSLQQQELKLLSRWWKDSGFS
 QMTFPRHRHVEYYTLASCIDSEPHSSFRLGFAKIFHLATVLDDIYDTFGTMD
 ELELFTA AVKRWHPSATEWLPEYMKGVYMLYETVNEMAGEAEKSQGRDT
 LNYARNALEAYIDASMEEAKWIFSGFLPTFEEYLDNGKVSFGYTIGTLQPILTL
 GIPFPHHILQEIDFPSRLNDVACSILRLKGDVHTYQPERSRGEESSCISCYMEEN
 PESTEEDAINHINSMDVKLLKELNWEYLRPDSNPITSKKHAFDILRALYHLY
 KYRDGYSVANYEIKNLVMTTVIEPVPL (SEQ ID NO:32)

P._contorta_MBO1_translation_frame_+1

SSKVKVVRRTMSTSIRMCQITTVSGEGVQRRRIANHHSNLWDDNFIQSLSTPYG
 AISYHESAQKLIGEVKEMINSISVKDGELITPSNDLLMRLSIVDSIERLGIDRHF
 KSEIKSALDYVYSYWNEKGIGWGRDSVVADLNSTALGLRTLRLHGYPVSSDV
 LQHFKEQKQGQFACSAIQTEGEIRSVLNLFRASQIAFPGEKVMEEAEVFSTKYL
 KEAILKLPVCGLSREISYVLEYGWHMNLPRLEARNYIDVFGQDPIYLTPNMKT
 QKLELAKLEFNMFHSLQQQELKLLSRWWKDSGFSQMTFPRHRHVEYYTLA
 SCIDSEPHSSFRLGFAKIFHLATVLDDIYDTFGTMDLELELFTA AVKRWHPSAT
 EWLPEYMKGVYMLYETVNEMAGEAEKSQGRDTLNYARNALEAYIDASME
 EAKWIFSGFLPTFEEYLDNGKVSFGYTIGTLQPILTLGIPFPHHILQEIDFPSRLN
 DVACSILRLKGDVHTYQPERSRGEESSCISCYMEDNPESTEEDAINHINSMDV
 KLLKELNWEYLRPDSNPITSKKHAFDILRAFYLHYKYRDGYSVANYEIKNLV
 MTTVIEPVPL (SEQ ID NO:33)

P._contorta_MBO2_translation_frame_+1

SSKVKVVRRTMSTSIRMCQITTVSGEGVQRRRIANHHSNLWDDNFIQSLSTPYG
 AISYHESAQKLIGEVKEMINSISVKDGELITPSNDLLMRLSIVDSIERLGIDRHF
 KSEIKSALDYVYSYWNEKGIGWGRDSVVADLNSTALGLRTLRLHGYPVSSDV
 LQHFKEQKQGQFACSAIQTEGEIRSVLNLFRASQIAFPGEKVMEEAEVFSTKYL
 KEAILKLPVCGLSREISYVLEYGWHINLPRLEARNYIDVFGQDPIYLTPNMKTQ
 KLELAKLEFNMFHSLQQQELKLLSRWWKDSGFSQMTFPRHRHVEYYTLAS
 CIDSEPHSSFRLGFAKIFHLATVLDDIYDTFGTMDLELELFTA AVKRWHPSAT
 EWLPEYMKGVYMLYETVNEMAGEAEKSQGRDTLNYARNALEAYIDASME
 EAKWIFSGFLPTFEEYLDNGKVSFGYTIGTLQPILTLGIPFPHHILQEIDFPSRLN
 DVACSILRLKGDVHTYQAERSRGEESSCISCYMEDNPESTEEDAINHINSMDV
 KLLKELNWEYLRPDSNPITSKKHAFDILRAFYLHYKYRDGYSVANYEIKNL
 VMTTVIEPVPL (SEQ ID NO:34)

FIG. 2E

19/20

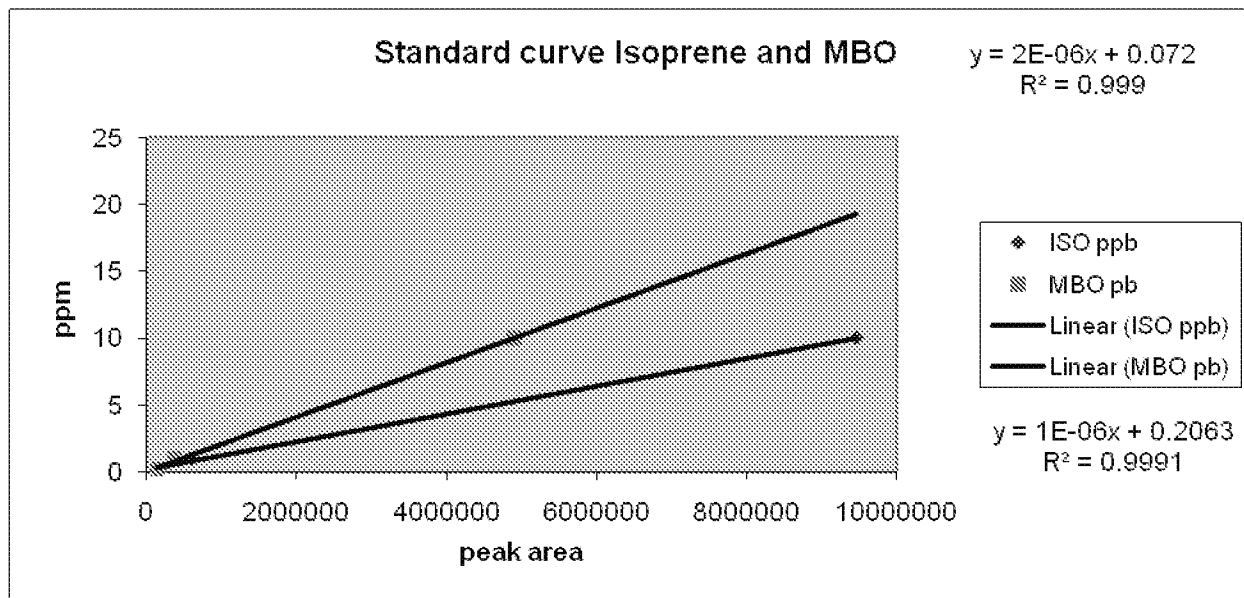


FIG. 3A

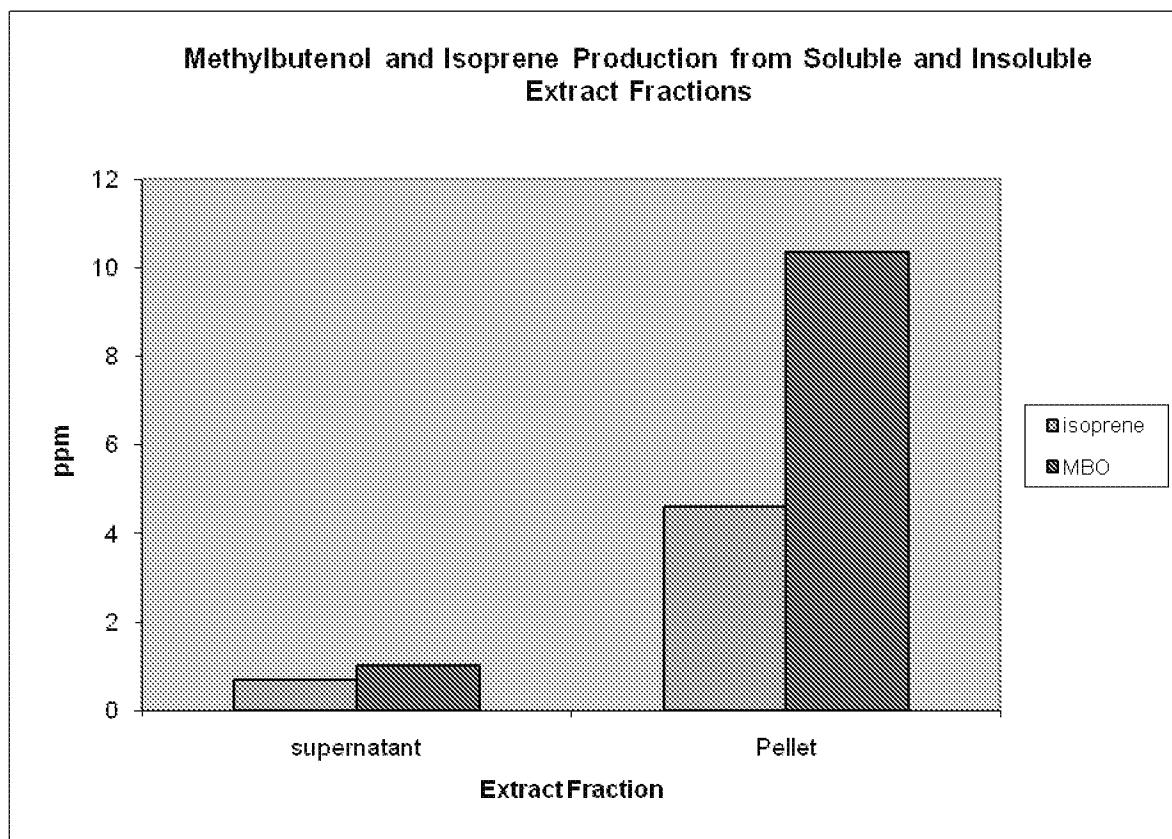


FIG. 3B

20/20

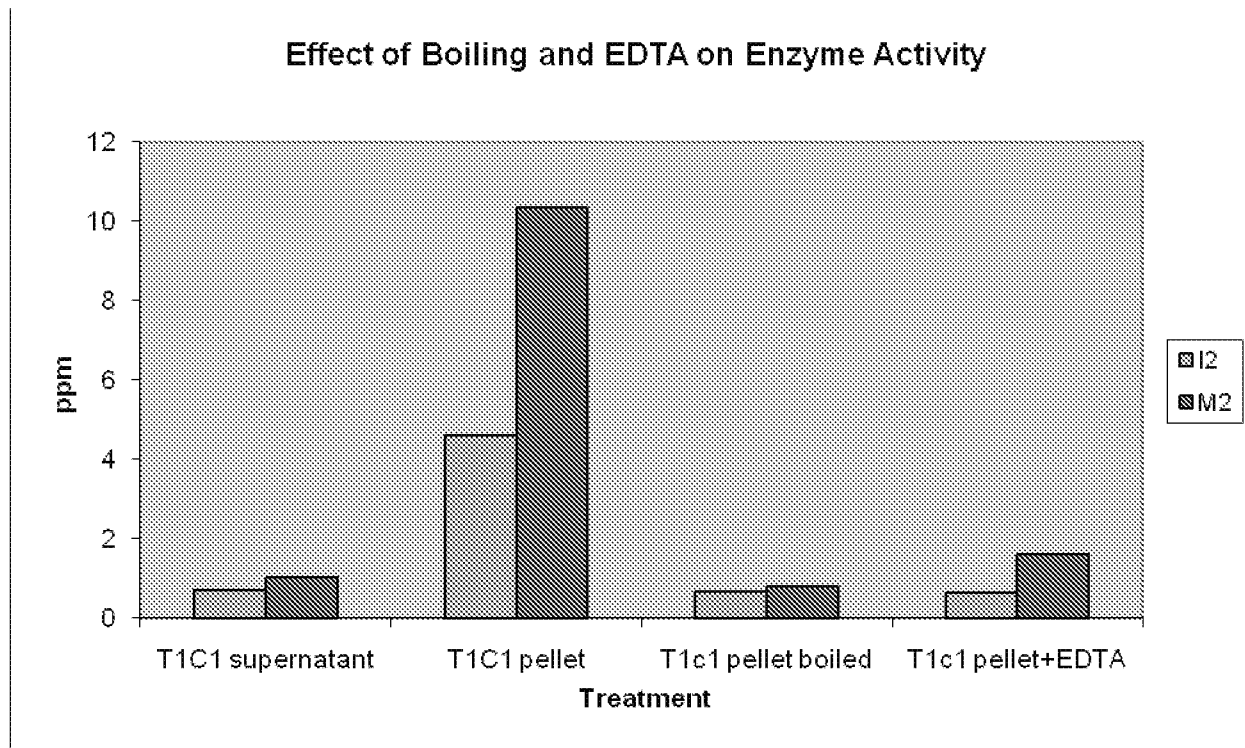


FIGURE 4

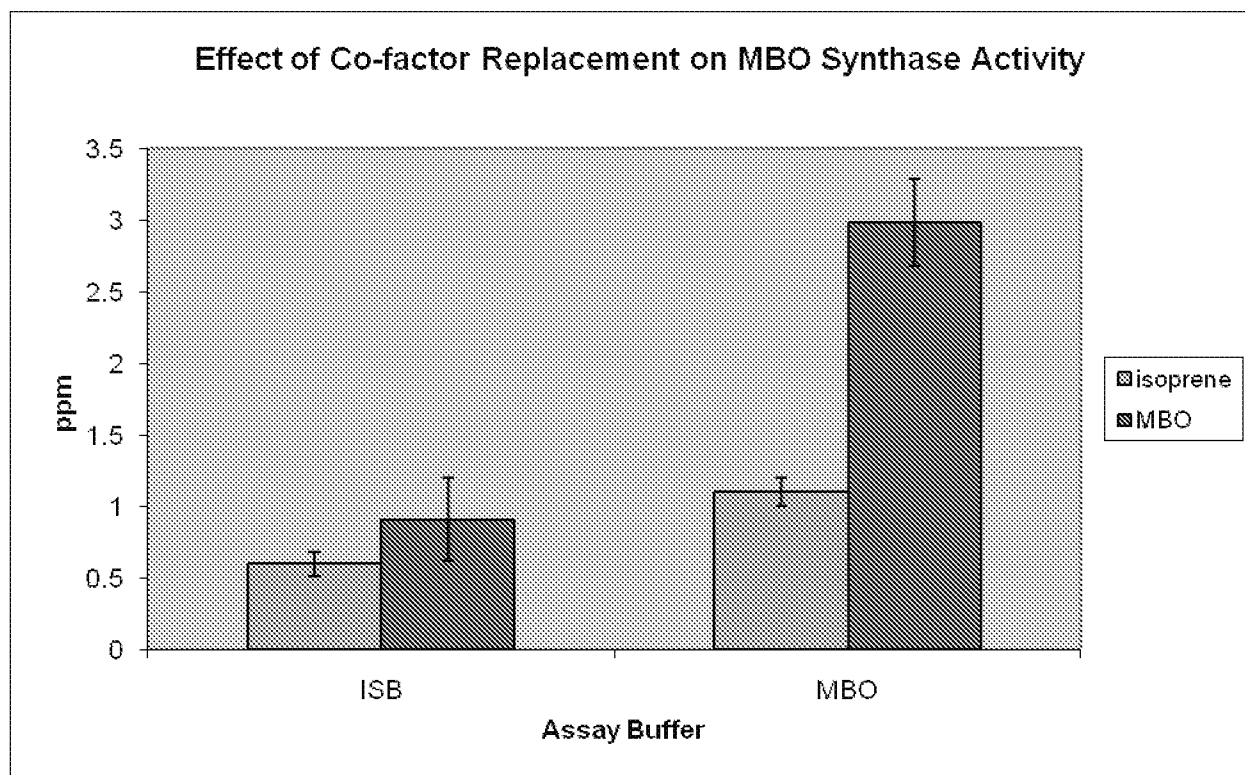


FIGURE 5