Title: NUCLEOTIDE SEQUENCE ENCODING AN ALCOHOL DEHYDROGENASE FROM ARTEMISIA ANNUA AND USES THEREOF

Abstract: An isolated nucleic acid molecule cloned from *Artemisia annua* encodes an alcohol dehydrogenase (Adh1). *Artemisia annua* Adh1 enzymatically oxidizes artemisinic alcohol to artemisinic aldehyde. The nucleic acid molecule, and the enzyme encoded thereby, may be used in processes to produce artemisinic aldehyde, dihydroartemisinic aldehyde, artemisinic acid and/or dihydroartemisinic acid in a host cell. Artemisinic aldehyde, dihydroartemisinic aldehyde, artemisinic acid and/or dihydroartemisinic acid can be chemically converted to the antimalarial compound artemisinin.
NUCLEOTIDE SEQUENCE ENCODING AN ALCOHOL DEHYDROGENASE FROM ARTEMISIA ANNUA AND USES THEREOF

Cross-reference to Related Applications

This application claims the benefit of United States Provisional Patent Application Serial No. 61/137,701 filed August 1, 2008, the entire contents of which is herein incorporated by reference.

Field of the Invention

The present invention relates to production of plant-derived compounds of health and commercial interest. More particularly, the present invention relates to nucleotide sequences encoding an enzyme, to an enzyme encoded by a nucleotide sequence and to processes for producing artemisinic aldehyde and other precursors of artemisinin.

Background of the Invention

Plants, in general, contain a myriad of secondary metabolites often synthesized by unique biochemical processes operating only in exotic species. For plant-derived products such as drugs, the 1997 worldwide sales were US$ 10 billion (Rotheim 2002). In many cases the supply of the relevant plant material for these drugs is limited or variable. One approach to developing methods for producing these drugs is to apply the methods of biochemistry, molecular biology and genomics to elucidate the biosynthesis and relevant biosynthetic genes for compounds of value for human health.

With the realization that many of the enzymes involved in natural product biosynthesis represent variations within known classes of enzymes, expressed sequence tag (EST) analysis (combined with heterologous expression) provides a powerful means of identifying their corresponding genes (Cahoon et al. 1999; Gang et al. 2001; Lange et al. 2000; van de Loo, Turner, & Somerville 1995).

One area of interest is bioactive compounds of the tribe Anthemideae in the family Asteraceae (Compositae) (Torrell et al. 1999; Watson, Evans, & Boluarte 2000). Anthemideae (Asteraceae, subfamily Asteroideae) is a tribe of 109 genera which includes daisies, chrysanthemums, tarragon, chamomile, yarrow and sagebrushes (Watson, Evans, & Boluarte 2000). These plants are aromatic in nature resulting from
high concentrations of mono- and sesqui- terpenes. Many of the species in this tribe are valued for the health benefits or insecticidal properties.

Of particular interest is artemisinin from Artemisia annua or sweet wormwood. In 1972, Chinese scientists reported the isolation of the sesquiterpene lactone containing an endoperoxide group (see Fig. 1) from Artemisia and called it qinghaosu (van Agtmael, Eggelte, & van Boxtel 1999). Prior to this sweet wormwood or qinghao had been used in traditional Chinese medicine for centuries. Artemisinin has become very important for the treatment of malaria in Southeast Asia and elsewhere, particularly for multi-drug-resistant falciparum forms of the disease (O'Neil 2005; Rathore et al. 2005; Robert, Coppel, & Meunier 2002; Wilairatana et al. 2002; Wu 2002). Since the discovery of artemisinin, a number of semi-synthetic derivatives have been developed for specific applications in malaria treatment.

Malaria remains a serious health problem which affects over 400 million people, especially in Africa and Southeast Asia, causing the deaths in excess of 2 million each year. Increasing resistance of the malaria parasite, Plasmodium falciparum, towards current antimalarial drugs is a cause for concern. The future value of antimalarial drugs based on the artemisinin structure is illustrated by the development by Bayer AG of Artemisone, an artemisinin derivative reported to be 10-30 fold more active than artesunate, for which clinical trials are currently under way.

Artemisinin is produced in relatively small amounts of 0.01 to 1.5% dry weight, making it and its derivatives relatively expensive (Gupta et al. 2002). Several studies describe the chemical synthesis of the sesquiterpene, but none are an economical alternative for isolation of artemisinin from the plant (Yadav, Babu, & Sabitha 2003). Therefore a higher concentration in the plant or production in an alternative host is desirable to make artemisinin available as economically as possible, especially for use in the Third World. Knowledge of the biosynthetic pathway and the genes involved should enable engineering of improved production of artemisinin. Alternatively, there is also the possibility of producing intermediates in the pathway to artemisinin which are of commercial value. For example, a compound 15 times more potent in vitro than artemisinin against Plasmodium falciparum has been synthesized from artemisinic alcohol (Jung, Lee, & Jung 2001).

There is evidence that artemisinin is localized to glandular trichomes on the surfaces of certain tissues of the plant (Duke et al. 1994; Duke & Paul 1993). The
number and even existence of these trichomes and the amount of artemisinin varies widely among bictypes.

Typically, compounds discovered in plants and found to be useful are produced commercially by i) chemical synthesis, where possible and economical, ii) extraction of cultivated or wild plants, or iii) cell or tissue culture (this is rarely economical). In those cases in which chemical synthesis is not economical, it makes sense to learn as much as possible about the biosynthesis of a natural product, such that it can be produced most efficiently in plants or cell/tissue culture. In the case of artemisinin, chemical synthesis is not commercially feasible. Since the compound is produced in small quantities in Artemisia, the drugs derived from artemisinin are relatively expensive, particularly for the Third World countries in which they are used. While the antimalarial drugs, chloroquine and sulfadoxine-pyrimethamine, cost as little as 20 cents for an adult treatment, artemisinin-derived compounds, by contrast, can be 100 times as expensive. Chloroquine resistance is prevalent and sulfadoxine-pyrimethamine resistance is increasing. The World Health Organization recently added the artemisinin-derived drug, artemether to their Model List of Essential Medicines, which are recommended to be available at all times in adequate amounts and in the appropriate dosage forms, and at a price that individuals and the community can afford. Consequently, it would be useful to be able to supply artemisinin-derived drugs more economically.

There are numerous patents relating to artemisinin and artemisinin derived drugs. These cover drug synthesis and formulation, Artemisia cultivation (Kumar 2002) and tissue culture and artemisinin extraction (Elferay 1990). In the past five years a reasonably clear picture of artemisinin biosynthesis has emerged as illustrated in Fig. 1 (Bertea et al. 2005). The identity of amorpha-4,11-diene as a biosynthetic intermediate was established, based on the presence of trace of amorpha-4,11-diene in Artemisia extracts and the cloning and expression of cDNAs representing amorpha-4,11-diene synthase, a terpene cyclase (Bouwmeester et al. 1999; Wallaart et al. 2001). A cytochrome P450 gene designated cyp71av1 was recently cloned and characterized (Teoh et al. 2006). The cyp71av1 gene encodes a hydroxylase that catalyzes the conversion of amorpha-4,11-diene to artemisinic alcohol. CYP71AV1 expressed in yeast is also capable of oxidizing artemisinic alcohol to artemisinic aldehyde and artemisinic aldehyde to artemisinic acid.

Commonly owned United States patent application 60/729,210 filed October 24, 2005, and now filed as a PCT patent application published May 3, 2007 under
publication number WO 2007/048235, the disclosures of which are herein incorporated by reference, discloses a gene encoding amorpha-4,11-diene hydroxylase, which catalyzes the first committed steps in artemisinin biosynthesis (Fig. 1).

Commonly owned international patent application PCT/CA2007/000614 filed April 4, 2007 and published on October 11, 2007 under publication number WO 2007/112596, the disclosure of which is herein incorporated by reference, discloses AaSAD (Aldh1), an aldehyde dehydrogenase. The AaSAD disclosed in this publication oxidizes artemisinic aldehyde to artemisinic acid.

Chang et al. (Chang 2007) reported the use of modified CYP71AV1 in E. coli and the production of artemisinic alcohol, artemisinic aldehyde and artemisinic acid. Given that artemisinic acid is more desirable for the commercial production of artemisinin, there is a need for an enzyme capable of efficient oxidation of artemisinic alcohol to artemisinic aldehyde. The aldehyde could then be oxidized to the acid either by CYP71AV1 or A. annua Aldh1

Summary of the Invention

The invention described herein addresses the production of artemisinin and artemisinin-related compounds, including precursors, of pharmaceutical and commercial interest.

There is provided an isolated nucleic acid molecule encoding an alcohol dehydrogenase capable of oxidizing artemisinic alcohol (Artemisia annua Adh1), the isolated nucleic acid molecule comprising a nucleotide sequence having at least 70% sequence identity to the nucleotide sequence as set forth in SEQ ID NO: 1.

There is provided a purified or partially purified alcohol dehydrogenase capable of oxidizing artemisinic alcohol comprising an amino acid sequence having at least 70% sequence identity to the amino acid sequence as set forth in SEQ ID NO: 2.

There is provided a purified or partially purified alcohol dehydrogenase capable of oxidizing artemisinic alcohol encoded by a nucleic acid molecule of the present invention.

There is provided a use of one or more purified or partially purified alcohol dehydrogenase capable of oxidizing artemisinic alcohol of the present invention in the production of dihydroartemisinic aldehyde.
There is provided a process for producing dihydroartemisinic aldehyde comprising expressing or overexpressing the isolated nucleic acid molecule of the present invention in a host cell.

There is provided a method of selecting or developing plants with altered artemisinic alcohol, artemisinic aldehyde, dihydroartemisinic aldehyde, dihydroartemisinic acid, artemisinic acid and/or artemisinin levels in a population of plants that naturally produces artemisinic alcohol, artemisinic aldehyde, dihydroartemisinic aldehyde, dihydroartemisinic acid, artemisinic acid and/or artemisinin comprising: detecting a target plant having altered levels of artemisinic alcohol, artemisinic aldehyde, dihydroartemisinic aldehyde, dihydroartemisinic acid, artemisinic acid and/or artemisinin compared to a control plant provided under similar conditions; isolating at least a portion of an alcohol dehydrogenase (Adh1) gene of the target plant and comparing the nucleotide sequence of said at least a portion to SEQ ID NO: 2 to detect a variation from SEQ ID NO: 2; detecting the variation in other plants; selectively breeding the plants with the variation to produce a population of plants having altered levels of artemisinic alcohol, artemisinic aldehyde, dihydroartemisinic aldehyde, dihydroartemisinic acid, artemisinic acid and/or artemisinin compared to a population of control plants produced under similar conditions.

There is provided a method of increasing artemisinic aldehyde, dihydroartemisinic aldehyde, dihydroartemisinic acid, artemisinic acid and/or artemisinin levels in a population of plants that naturally produces dihydroartemisinic aldehyde, dihydroartemisinic acid, artemisinic acid and/or artemisinin comprising: providing a population of mutated plants; detecting a target mutated plant within the population of mutated plants, the target mutated plant having an altered expression of an *Artemisia annua* Adh1 gene or altered activity of *A. annua* Adh1 enzyme compared to a control plant provided under similar conditions, said detecting comprising using primers developed from a nucleic acid molecule of the present invention to PCR amplify regions of the Adh1 gene from mutated plants in the population of mutated plants, identifying mismatches between the amplified regions and corresponding regions in the wild-type gene that lead to the altered expression or altered activity, and identifying the mutated plant that contains the mismatches; and, selectively breeding the target mutated plant to produce a population of plants having altered expression of the *A. annua* Adh1 gene or altered alcohol dehydrogenase enzyme activity compared to a population of control plants produced under similar conditions.
The nucleotide sequence may have at least 80%, at least 90%, at least 95% or at least 99% sequence identity to the nucleotide sequence as set forth in SEQ ID NO: 1. The amino acid sequence may have at least 80%, at least 90%, at least 95% or at least 99% sequence identity to the amino acid sequence as set forth in SEQ ID NO: 2.

The gene (nucleic acid molecule) of the present invention may be derived, for example cloned, from *Artemisia annua*. Obtaining other nucleic acid molecules of the present invention may be accomplished by well-known techniques in the art. Such techniques are disclosed in Sambrook et al. 2001 and Ausubel et al. eds. 2001. Generally, other nucleic acid molecules of the present invention may be obtained by:

a) identifying the existence of a homologous gene by techniques such as, for example, hybridization of a target gene to the complement of SEQ ID NO: 1, genome or transcriptome (cDNA) sequencing, or database searching in nucleic acid sequence databases such as Genbank; b) cloning the homologous gene and creating a plasmid for *E. coli* or yeast expression as described in Sambrook et al. 2001 or Ausubel et al. eds. 2001; and, c) then testing the gene product for artemisinic alcohol oxidation.

Database searching may employ commonly used computer programs such as BLASTX, BLASTP, TBLASTN and others.

Other nucleic acid molecules and proteins of the present invention may also be obtained by creating mutations of nucleic acid molecules and proteins already at hand. Such mutations may be accomplished by commonly known methods in the art as described in Sambrook et al. 2001 and Ausubel et al. eds. 2001.

Overexpression of one or more of the nucleic acid molecules or overproduction of the Adh1 enzyme may be done in *A. annua*. Expression of one or more of the nucleic acid molecules or expression of the alcohol dehydrogenase gene may be done in other hosts, for example plants, yeasts or bacteria. Overexpression or expression of one or more of the isolated nucleic acid molecules of the present invention may be done in combination with overexpression or expression of one or more other nucleic acid molecules involved in the biosynthesis of artemisinin, for example those encoding farnesyl diphosphate synthase, amorpha-4,11-diene synthase, amorpha-4,11-diene hydroxylase, double bond reductase, aldehyde dehydrogenase.

Part of the solution to the problem of producing artemisinin in an economical and timely fashion is the isolation and exploitation of genes involved in artemisinin biosynthesis. As in other examples of metabolic engineering, such genes can be used to enhance production by overexpression in the native plant (*A. annua*), a different
plant, or in micro-organisms such as bacteria or yeast. An example of this is the expression of the amorphaadiene synthase gene in *E. coli* to produce the artemisinin precursor amorphaadiene (Martin et al. 2003) and the production of artemisinic acid in yeast (Ro et al. 2006). One important step in the pathway to artemisinin *per se*, is thought to be the oxidation of artemisinic alcohol to artemisinic aldehyde. Consequently, the gene involved in this step may be used to produce artemisinic aldehyde in a host, alone or in combination with one or more of farnesyl diphosphate synthase, amorpha-4,11-diene synthase, and amorpha-4,11-diene hydroxylase. The resulting artemisinic aldehyde could then be converted to artemisinic acid, dihydroartemisinic aldehyde or dihydroartemisinic acid with other known genes or enzymes.

The resulting artemisinic acid or dihydroartemisinic acid could then be chemically converted to artemisinin or related compounds of commercial value. For example, dihydroartemisinic acid is a presumed late precursor of artemisinin, and its transformation to artemisinin has been shown to occur spontaneously through photooxidation, requiring no enzyme intervention (Sy & Brown 2002; Wallaart et al. 1999). Consequently, using dihydroartemisinic acid

The nucleic acid molecule of the present invention may also be used in the development of DNA markers and in targeted mutagenesis techniques (e.g. TILLING (Targeting Induced Local Lesions IN Genomes)).

A genetic marker (DNA marker) is a segment of DNA with an identifiable physical location on a chromosome and associated with a particular gene or trait and whose inheritance can be followed. A marker can be a gene, or it can be some section of DNA with no known function. Because DNA segments that lie near each other on a chromosome tend to be inherited together, markers are often used as indirect ways of tracking the inheritance pattern of a gene that has not yet been identified, but whose approximate location in the genome is known. Thus, markers can assist breeders in developing populations of organism having a particular trait of interest. Gene-specific markers can be used to detect genetic variation among individuals which is more likely to affect phenotypes relating to the function of a specific gene. For example, variation in a gene-specific marker based on *A. annua Adh1*, rather than variation in an anonymous DNA marker, would be more likely linked to variation in content of artemisinin or related compounds, by virtue of its association with the relevant biosynthetic pathway. In one embodiment, a DNA marker for *Adh1* could be developed by sequencing the polymerase chain reaction amplified *Adh1* gene
from a number of individual plants of *Artemisia annua*. Such sequencing would provide information about sequence polymorphisms within the gene. A range of methods available to those skilled in the art could be used to detect such polymorphisms, including cleaved amplified polymorphic sequences (CAPs) (Konieczny & Ausubel 1993).

The presence of such gene-specific polymorphisms could be correlated with levels of artemisinin or related compounds and used in a breeding program to select and/or develop lines of *Artemisia annua* with enhanced levels of artemisinin or related compounds. That is, the variation in genetic structure may be detected in other plants, and the plants with the variation selectively bred to produce a population of plants having increased levels of dihydroartemisinic aldehyde, dihydroartemisinic acid artemisinic acid and/or artemisinin compared to a population of control plants produced under similar conditions. Genetic markers are discussed in more detail in Bagge et al. 2007, Pfaff et al. 2003, Sandal et al. 2002 and Stone et al. 2002.

TILLING (Bagge, Xia, & Lubberstedt 2007; Comai & Henikoff 2006; Henikoff, Till, & Comai 2004; Slade & Knauf 2005) involves treating seeds or individual cells with a mutagen to cause point mutations that are then discovered in genes of interest using a sensitive method for single-nucleotide mutation detection. Detection of desired mutations (e.g. mutations resulting in a change in expression of the gene product of interest) may be accomplished, for example, by PCR methods. For example, oligonucleotide primers derived from the gene (nucleic acid molecule) of interest, such as the nucleic acid molecules of the present invention, may be prepared and PCR may be used to amplify regions of the gene of interest from plants in the mutagenized population. Amplified mutant genes may be annealed to wild-type genes to find mismatches between the mutant genes and wild-type genes. Detected differences may be traced back to the plants which had the mutant gene thereby revealing which mutagenized plants will have the desired expression. These plants may then be selectively bred to produce a population having the desired expression.

Further features of the invention will be described or will become apparent in the course of the following detailed description.
**Brief Description of the Drawings**

In order that the invention may be more clearly understood, embodiments thereof will now be described in detail by way of example, with reference to the accompanying drawings, in which:

5 Fig. 1 depicts the proposed biosynthetic pathway for artemisinin biosynthesis.

Fig. 2 depicts the nucleotide sequence (SEQ ID NO: 1) of the *Artemisia annua Adh1* cDNA encoding alcohol dehydrogenase.

Fig. 3 depicts the predicted amino acid sequence (SEQ ID NO: 2) of the protein encoded by *Artemisia annua Adh1*.

10 Fig. 4 depicts a GC/MS analysis of alcohol dehydrogenase assays using artemisinic alcohol and dihydroartemisinic alcohol as substrates.

**Description of Preferred Embodiments**

**Materials and Methods:**

*Artemisinic aldehyde*

15 Artemisinic acid was isolated from dichloromethane extracts of *A. annua* flower buds and leaves (Teoh, Polichuk, Reed, Nowak, & Covello 2006) and was used to synthesize artemisinic aldehyde according to the method described by Chang et al. 2000, the disclosure of which is incorporated herein by reference.

*Dihydroartemisinic aldehyde*

20 Dihydroartemisinic aldehyde was synthesized from the isolated dihydroartemisinic acid (see above). The acid was converted to methyl dihydroartemisinate with excess diazomethane in diethyl ether at 0°C for 5 minutes. The ether and diazomethane were removed under a stream of nitrogen and the methyl ester was reduced to *(11R)-dihydroartemisinic alcohol* with excess 1.5 M diisobutyl aluminum hydride in toluene at room temperature for 10 min under nitrogen. With subsequent extraction, oxidation to the aldehyde with pyridinium chlorochromate (Corey & Suggs 1975) and purification by HPLC the *(11R)-dihydroartemisinic aldehyde* was produced at an overall yield of 48% with >90% purity according to GC analysis.
Plant materials

Artemisia annua L. seeds were obtained from Elixir Farm Botanicals, Bixey, Missouri, USA and from Pedro Melillo de Magalhães, State University of Campinas, Brazil (line 2/39). Seeds were germinated and grown in soil in a controlled environment chamber with 16 hour/25°C days and 8 hour/20°C nights. Plants that had reached the height of approximately 1.2 m (about 3 months) were transferred to flowering chamber with 12 hour/25°C days and 12 hour/20°C nights for the Elixir line and 8 hour/25°C days and 16 hour/20°C nights for line 2/39. Flower buds that developed after 19-21 days in the flowering chamber were harvested for total RNA isolation.

cDNA library construction and expressed sequence tag (EST) analysis

Total RNA was extracted and isolated from glandular trichomes and flower buds using a modified method described by Logeman, et al. 1987. cDNA synthesis from 1.5 micrograms of total RNA and construction of the trichome and flower bud cDNA library were carried out with Creator™ SMART™ cDNA Library Construction Kit (Clontech). A total of 6,239 clones and 2,208 clones for trichome and flower bud libraries, respectively were randomly picked and their DNA sequences determined. Sequencing was performed on an ABI3700 DNA sequencer using BigDye Terminator Cycle Sequencing Kit (Applied Biosystems) and the M13 reverse primer. DNA sequence traces were interpreted and vector and low quality sequences were eliminated using PHRED (Ewing et al. 1998) and LUCY (Chou & Holmes 2001). Clustering of the resulting EST dataset was done using STACKPACK (Miller et al. 1999) and sequence similarity was identified by BLAST (Altschul et al. 1990).

Isolation of full-length Artemisia annua Adh1 cDNA

The gene corresponding to an Artemisia annua cDNA clone was identified by sequence similarity searches as showing similarity to other nucleotide sequences encoding alcohol dehydrogenases was designated Artemisia annua Adh1. The nucleotide sequence of the open reading frame of Adh1 is given by SEQ ID NO: 1.

Expression of Adh1

The Gateway vector pDEST17 had the Adh1 gene inserted and was subsequently transformed into a competent expression E. coli strain Rosetta (DE3). These cells were grown overnight in 3 ml of LB broth amended with 100 μg/ml ampicillin. An aliquot of 1 ml overnight culture was sub cultured to 100 ml of fresh LB
broth also amended with 100 µg/ml of ampicillin. The culture was incubated at 37°C until the optical density reached 0.8 at 600 nm. A final concentration of 1 mM IPTG was then added to the culture and the flask was incubated at 30°C for another 4 hours. Cells were harvested by centrifugation at 10,000 rcf for 10 minutes. Pelleted cells were resuspended in 20 mM sodium phosphate pH 7.4 and 500 mM NaCl. Cell walls were disrupted by sonication on ice at intervals of 0.7 seconds for 1 minute, followed by a 1 minute cool down this was repeated three times. The sonicated cells were centrifuged at 10,000 rcf for 10 minutes; the supernatant which contained the soluble ADH1 was concentrated and desalted by spin dialysis (using Amicon Ultra-15 devices; Millipore, MA), and stored in a new tube at 4°C. This method of recombination and expression was also used for a control plasmid pDEST17-gus.

Specific ADH1 Activity

Enzyme assay was prepared to determine the specific activity of ADH1. The assay contained a final concentration of 1 mM nicotinamide adenine dinucleotide (NAD⁺), 0.1 µg/µl artemisinic alcohol or dihydroartemisinic alcohol (AAOH or DHAAOH) and approximately 90 µg of crude protein. The assays were incubated at 30°C for 30 minutes at 500 rpm, followed by an ethyl acetate extraction. The upper organic phase was transferred to a GC vial and the sample was run on an Agilent 6890N – Network GC System coupled to an Agilent 5973 Network Mass Selective Detector. The column present in the GC was an Innowax column (30 m, I.D. 0.250 mm, Film 0.25 µm), with a 50:1 split injector, the temperature gradient began at 125°C and increased at 5°C/minute to a maximum temperature of 300°C.

Results:

Specific ADH1 Activity

Enzyme assays were prepared to determine whether ADH1 acted specifically on artemisinic alcohol or if there was also activity observed with dihydroartemisinic alcohol as a substrate. These substrates were tested with cells that were expressing the ADH1 gene as well as a control gus gene. It was observed that ADH1 acts on AAOH and not DHAAOH (Fig. 4). Fig. 4A shows GC/MS analysis of crude Adh1 incubated with artemisinic alcohol (retention time ≈ 20.55 min), and the presence of product artemisinic aldehyde (14.37 min). Peaks were confirmed by mass spectrometry. Fig. 4B shows GC/MS analysis of crude Adh1 incubated with
dihydroartemisinic alcohol (19.05 min), and no observed product. No product was observed with extracts of *E. coli* transformed with the control plasmid pDEST17-gus.

Sequences:

SEQ ID NO: 1 – *Artemisia annua* (alcohol dehydrogenase gene) (1137 bp)

```
5  atggccccaa aagctcccgg tgctcatcct tgaagccg cgggtttgtatg
ggattcaaat gtccctcagga gtctcagga gataaaggggt gcacagcga
aagcatactg aagctcccggt tggtaagttt acctttcatca tctctctctct
ctgctcaggt ctacagcagctg aacattttta cgggctcgtcgt
aaccttttct cctcttcatt tataacctta ataaatctatcttt aacagatcv
```

SEQ ID NO: 2 – *Artemisia annua* (alcohol dehydrogenase) (378 aa)

```
30  MAQKAPGVITCAAVWESSGPVLEERIVDPKASEVRKMLCASLCHTDVLCTKGFIPLFP
    RIPGHEGVGYESTGKDALKLPDGIVNYLGECCQCLNCKTGLHFTNGHNYPFSGLMNHDGT
    SRMSIARTGEISYHAFASCTRTEYAVADCNYVRKNIKSYKASFLSCGFTTGGATWRETQV
    SKGSSVAVFGITVGLGVIQGLQGASKIIIGDVNYQAKGKVFEMDFINPRKDPKCSVSE
    LVKELTHCLGDPCFECTCVPSLNEALEASKIGGTVVPIGAGGEEAVANSLSILFSGRTLKF
    TAFGGVRTQSDLPIIHKCNKXQDLDELETIHLNIQAEFEILKKFPDCVHILK
```

**References:** The disclosures of the following references are incorporated herein by reference in their entirety.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990), "Basic local alignment search tool", *Journal of Molecular Biology*, vol. 215, pp. 403-410.


Kumar, S. (2002), Method for maximization of artemisinin production by the plant Artemisia annua, U.S. patent 6,393,763.


"A genetic linkage map of the model legume *Lotus japonicus* and strategies for fast mapping of new loci", *Genetics*, vol. 161, no. 4, pp. 1673-1683.


Other advantages that are inherent to the structure are obvious to one skilled in the art. The embodiments are described herein illustratively and are not meant to limit the scope of the invention as claimed. Variations of the foregoing embodiments will be evident to a person of ordinary skill and are intended by the inventor to be encompassed by the following claims.
Claims:

1. Isolated nucleic acid molecule comprising a nucleotide sequence having at least 70% nucleotide sequence identity to the nucleotide sequence as set forth in SEQ ID NO: 1, and encoding an alcohol dehydrogenase.

2. Isolated nucleic acid molecule comprising a nucleotide sequence encoding an alcohol dehydrogenase having an amino acid sequence with at least 70% amino acid sequence identity to SEQ ID NO: 2.

3. The nucleic acid molecule according to claim 2, wherein the amino acid sequence identity is at least 90%.

4. The nucleic acid molecule according to claim 2, wherein the amino acid sequence identity is at least 95%.

5. The isolated nucleic acid molecule according to any one of claims 1 to 4 having a nucleotide sequence comprising SEQ ID NO: 1.

6. The isolated nucleic acid molecule of any one of claims 1 to 5 derived from Artemisia annua.

7. Purified or partially purified alcohol dehydrogenase having an amino acid sequence with at least 70% amino acid sequence identity to SEQ ID NO: 2.

8. The dehydrogenase according to claim 7, wherein the amino acid sequence identity is at least 90%.

9. The dehydrogenase according to claim 7, wherein the amino acid sequence identity is at least 95%.

10. The dehydrogenase according to claim 7 having an amino acid sequence comprising SEQ ID NO: 2.

11. Purified or partially purified alcohol dehydrogenase encoded by the isolated nucleic acid molecule as defined in any one of claims 1 to 6.

12. Use of one or more isolated nucleic acid molecules as defined in any one of claims 1 to 6 in the production of artemisinic aldehyde.
13. Use of one or more purified or partially purified alcohol dehydrogenases as defined in any one of claims 7 to 11 in the production of artemisinic aldehyde.

14. Process for producing artemisinic aldehyde comprising expressing or overexpressing one or more isolated nucleic acid molecules as defined in any one of claims 1 to 6 in a host cell.

15. The process according to claim 14, further comprising expressing or overexpressing in the host cell one or more nucleic acid molecules encoding farnesyl diphosphate synthase, amorpha-4,11-diene synthase, amorpha-4,11-diene hydroxylase, double bond reductase, aldehyde dehydrogenase, or any combination thereof.

16. Process for producing artemisinic aldehyde, dihydroartemisinic acid and/or artemisinic acid comprising producing or overproducing in a host cell one or more alcohol dehydrogenases as defined in any one of claims 7 to 11.

17. The process according to any one of claims 14 to 16, wherein the host cell is a plant cell.

18. The process according to any one of claims 14 to 16, wherein the host cell is a yeast cell.

19. The process according to any one of claims 14 to 16, wherein the host cell is a bacterial cell.

20. Method of selecting or developing plants with altered artemisinic aldehyde, dihydroartemisinic aldehyde, dihydroartemisinic acid, artemisinic acid and/or artemisinin levels in a population of plants that naturally produces artemisinic aldehyde, dihydroartemisinic aldehyde, dihydroartemisinic acid, artemisinic acid and/or artemisinin comprising:

   detecting a target plant having altered levels of artemisinic aldehyde, dihydroartemisinic aldehyde, dihydroartemisinic acid, artemisinic acid and/or artemisinin compared to a control plant provided under similar conditions;

   isolating at least a portion of an alcohol dehydrogenase gene of the target plant and comparing the nucleotide sequence of said at least a portion to SEQ ID NO: 1 to detect a variation from SEQ ID NO: 1;
detecting the variation in other plants;

selectively breeding the plants with the variation to produce a population of plants having altered levels of artemisinic aldehyde, dihydroartemisinic aldehyde, dihydroartemisinic acid, artemisinic acid and/or artemisinin compared to a population of control plants produced under similar conditions.

21. Method of increasing artemisinic aldehyde, dihydroartemisinic aldehyde, dihydroartemisinic acid, artemisinic acid and/or artemisinin levels in a population of plants that naturally produces dihydroartemisinic aldehyde, dihydroartemisinic acid, artemisinic acid and/or artemisinin comprising:

providing a population of mutated plants;

detecting a target mutated plant within the population of mutated plants, the target mutated plant having an altered expression of an alcohol dehydrogenase gene or altered activity of an alcohol dehydrogenase enzyme compared to a control plant provided under similar conditions,

said detecting comprising using primers developed from a nucleic acid molecule as defined in any one of claims 1 to 6 to PCR amplify regions of the alcohol dehydrogenase gene from mutated plants in the population of mutated plants, identifying mismatches between the amplified regions and corresponding regions in wild-type gene that lead to the altered expression or altered activity, and identifying the mutated plant that contains the mismatches; and,

selectively breeding the target mutated plant to produce a population of plants having altered expression of the alcohol dehydrogenase gene or altered activity of the alcohol dehydrogenase enzyme compared to a population of control plants produced under similar conditions.
Fig. 1
FIG. 2

FIG. 3
INTERNATIONAL SEARCH REPORT

International application No.
PCT/CA2009/000966

A. CLASSIFICATION OF SUBJECT MATTER
IPC: C12N 15/53 (2006.01), A01H 1/00 (2006.01), A01H 1/04 (2006.01), C12N 15/63 (2006.01), C12N 9/04 (2006.01), C12P 17/18 (2006.01), C12Q 1/68 (2006.01)
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED
Minimum documentation searched (classification system followed by classification symbols)
IPC: C12N 15/53 (2006.01), A01H 1/00 (2006.01), A01H 1/04 (2006.01), C12N 15/63 (2006.01), C12N 9/04 (2006.01), C12P 17/18 (2006.01), C12Q 1/68 (2006.01)

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic database(s) consulted during the international search (name of database(s) and, where practicable, search terms used)
Databases: Canadian Patent, Delphion, Pubmed, Scopus, CAPplus, Esp@cenet, GenPept, GenBank
Keywords: Anthelmintic, dihydroartemisinic aldehyde or acid, artemisinic acid, artemisinin, Artemisia annua, sweet wormwood, SEQ ID NOs: 1-2

C. DOCUMENTS CONSIDERED TO BE RELEVANT

<table>
<thead>
<tr>
<th>Category</th>
<th>Citation of document, with indication, where appropriate, of the relevant passages</th>
<th>Relevant to claim No.</th>
</tr>
</thead>
</table>

[ ] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

Date of the actual completion of the international search
6 August 2009 (06-08-2009)

Date of mailing of the international search report
7 October 2009 (07-10-2009)

Authorized officer
Damiano Conte 819-934-7931

Form PCT/ISA/210 (second sheet) (July 2008)
1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of:

   a. type of material
      * [X] a sequence listing
      * [ ] table(s) related to the sequence listing

   b. format of material
      * [X] on paper
      * [X] in electronic form

   c. time of filing/furnishing
      * [X] contained in the international application as filed.
      * [X] filed together with the international application in electronic form
      * [ ] furnished subsequently to this Authority for the purposes of search.

2. [ ] In addition, in the case that more than one version or copy of a sequence listing and/or table relating thereto has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that in the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments: