2-27-2007

Loline Alkaloid Gene Clusters of the Fungal Endophyte Neotyphodium Uncinatum

Christopher L. Schardl
University of Kentucky, schardl@uky.edu

Heather H. Wilkinson
University of Kentucky

Martin J. Spiering
University of Kentucky

Click here to let us know how access to this document benefits you.

Follow this and additional works at: https://uknowledge.uky.edu/plantpath_patents

Part of the Plant Pathology Commons

Recommended Citation
Scharld, Christopher L.; Wilkinson, Heather H.; and Spiering, Martin J., "Loline Alkaloid Gene Clusters of the Fungal Endophyte Neotyphodium Uncinatum" (2007). Plant Pathology Faculty Patents. 5.
https://uknowledge.uky.edu/plantpath_patents/5

This Patent is brought to you for free and open access by the Plant Pathology at UKnowledge. It has been accepted for inclusion in Plant Pathology Faculty Patents by an authorized administrator of UKnowledge. For more information, please contact UKnowledge@Isr.uky.edu.
LOLINE ALKALOID GENE CLUSTERS OF THE FUNGAL ENDOPHYTE NEOTYPHIDIUM UNCINATUM

Inventors: Christopher L. Scharld, Lexington, KY (US); Heather H. Wilkinson, College Station, TX (US); Martin J. Spiering, Lexington, KY (US)

Assignee: University of Kentucky Research Foundation, Lexington, KY (US)

Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 475 days.

Appl. No.: 10/601,700
Filed: Jun. 24, 2003

Prior Publication Data

Related U.S. Application Data
Provisional application No. 60/390,446, filed on Jun. 24, 2002.

Int. Cl.
C12P 17/14 (2006.01)
C07H 21/04 (2006.01)
A61B 69/36 (2006.01)

U.S. Cl. .......... 435/252.3; 435/120; 435/320.1; 435/252.3; 435/252.8; 435/254.11; 536/23.2

Field of Classification Search .......... 435/251.2, 435/252.2, 435/252.3, 252.8, 254.11, 252.33, 320.1; 536/23.2

See application file for complete search history.

References Cited

OTHER PUBLICATIONS

Primary Examiner—Rebecca E. Prouty
Assistant Examiner—Kagneg Gebreyesus
Attorney, Agent or Firm—McDermott Will & Emery LLP

ABSTRACT

Loline alkaloids (LA), which are 1-aminopyrrolizidines with an oxygen bridge, are produced by Epichloë (anamorph=Neotyphodium) species, endophytes of grasses. LA are insecticidal, thus helping protect host plants from insect herbivory. Suppression subtractive hybridization PCR was used to isolate transcripts up-regulated during loline alkaloid production in cultures of Neotyphodium uncinatum. Subtracted cDNAs were cloned, and a X-phage cDNA library from an LA-expressing N. uncinatum culture was screened with subtracted cDNA. In BLAST searches, several cDNAs identified had sequence similarities to aspartate kinases, and another with O-acetylhomoserine-thiolylase. Differential expression of these two genes in LA-producing cultures of N. uncinatum was confirmed, and in a survey of 23 isolates from 21 Neotyphodium and Epichloë species these two genes strictly correlated with LA production. Two nucleic acid molecules encoding two loline alkaloid gene clusters have been identified.

7 Claims, 6 Drawing Sheets
OTHER PUBLICATIONS


<table>
<thead>
<tr>
<th>Compound</th>
<th>R</th>
<th>R'</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loline</td>
<td>CH₃</td>
<td>H</td>
</tr>
<tr>
<td>Norloline</td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td>N-Methylloline</td>
<td>CH₃</td>
<td>CH₃</td>
</tr>
<tr>
<td>N-Formylloline</td>
<td>CH₃</td>
<td>CHO</td>
</tr>
<tr>
<td>N-Formylnorloline</td>
<td>H</td>
<td>CHO</td>
</tr>
<tr>
<td>N-Acetylloline</td>
<td>CH₃</td>
<td>COCH₃</td>
</tr>
<tr>
<td>N-Acetylnorloline</td>
<td>H</td>
<td>COCH₃</td>
</tr>
</tbody>
</table>

**Figure 1**
Figure 2
Figure 5
Fig. XX. Loline alkaloid biosynthesis genes
LOLINE ALKALOID GENE CLUSTERS OF THE FUNGAL ENDOPHYTE NEOTYPHODIUM UNCINATUM

RELATED APPLICATIONS

The present application claims the benefit of the priority date of provisional patent application No. 60/390,446, filed Jun. 24, 2002. The complete disclosure of the earlier filed application is incorporated by reference herein.

GOVERNMENT SUPPORT

The present invention was supported by the United States National Science Foundation Integrative Plant Biology Program 9808554 to C.S. and the United States Department of Agriculture NRICGP grant 9901343.

FIELD OF THE INVENTION

The present invention relates generally to alkaloids and alkaloid biosynthesis. In particular, the invention pertains to the nucleic acids encoding loline alkaloid synthesis genes and the tailoring enzymes of loline alkaloid biosynthesis, and to recombinant vectors and host cells containing such genes, and to the recombinant production of alkaloids and uses thereof.

BACKGROUND OF THE INVENTION

Loline alkaloids (LA; saturated 1-aminopyrrolizidine alkaloids with an ether bridge, FIG. 1), are produced in a number of associations of grasses with endophytes of the genus Epichloë and their asexual descendants, Neotyphodium spp. In addition, LA are reported from the plants Adenocarpus spp. and Aryegeria mollys of the families Fabaceae and Convolvulaceae, respectively. LA produced in grass-endophyte symbioses have strong insecticidal and feeding-deterrent properties (Riedell et al., 1999, J. Entomol. Sci. 26: 122–129; Wilkinson et al., 2000, Mol. Plant-Microbe Interact. 13: 1027–1033). Moreover, grasses infected by LA-producing endophytes, such as Neotyphodium coenophialum and N. uncinatum, have greater tolerance to drought conditions (Arechavalea et al., 1989, Agron. J. 81: 83–90; Bacon, 1993, Agric. Ecos. Environ. 44: 123–141) than grasses infected by closely related endophytes, such as N. loli, that do not produce LA (Burker et al., 1997, Agric. Eos. Environ. 44: 123–141; Cheplick et al., 2000, Mycol. Res. 97: 1083–1092). Growth suppression (allelopathy) of neighboring plants by meadow fescue (Lolium pratense) infected with N. uncinatum may indicate a potential for additional beneficial roles of these alkaloids in grass plant competitiveness and persistence.

LA can accumulate to extremely high levels in grass tissues, occasionally reaching more than 2% of the plant’s dry mass (Craven et al., 2001, Sydowia 53: 44–73). These quantities far exceed the biomass of the fungus and the amounts of other alkaloids, such as ergot alkaloids, indole-diterpenoids, and peramine, also produced in some of the endophyte-grass symbiota. However, despite their exceptional levels in the grass and importance of LA in grass survival, little is known about LA biosynthesis. This is in contrast to some of the other endophyte-associated alkaloids, such as ergopeptides and indole-diterpenoids, for which much of the biosynthetic pathways have been elucidated and key enzymes identified.

It was previously unknown whether LA are of fungal or plant origin, or produced by both symbiotic partners together, but a recent study has established that N. uncinatum can produce LA in axenic culture (Blankenship et al., 2001, Phytochemistry 58: 395–401). This finding presents opportunities to identify genes involved in LA biosynthesis. Knowledge of the LA biosynthesis genes would allow more detailed studies on the roles of LA in plant persistence, in particular on possible contributions to abiotic stress tolerance, as well as the cloning and use of these genes to generate genetically engineered plants.

SUMMARY OF THE INVENTION

In one aspect, the present invention provides isolated nucleic acid compounds comprising at least a sequence identical or complementary to all or part of a coding sequence for the loline alkaloid biosynthetic gene cluster from Neotyphodium uncinatum (SEQ ID NO. 15, and SEQ ID NO. 16). It appears that SEQ ID NO. 17 may be linked to the 5′ end of SEQ ID NO. 16. Preferably, a part of said coding sequence is an open reading frame (ORF) selected from the group consisting of ORF1, ORF2, ORF3, ORF4, ORF5, ORF6, ORF7, ORF8, ORF9, ORF1′, ORF2′, ORF3′, ORF4′, ORF5′, ORF6′, ORF7′, ORF8′, ORF9′ or ORF10′. More preferably, a part of said coding sequence is an ORF selected from the group consisting of ORF1, ORF2, ORF3, ORF4, ORF5, ORF6, ORF7, ORF8, ORF9, ORF1′, ORF2′, ORF3′, ORF4′, ORF5′, ORF6′, ORF7′, and ORF8′.

In one embodiment, the present invention provides an isolated nucleic acid strand that encodes a loline alkaloid gene cluster or subunit thereof comprising a nucleotide sequence identical or complementary to, or an amino acid sequence encoded by a nucleotide sequence identical or complementary to, all or part of a coding sequence for loline alkaloid biosynthetic gene cluster of SEQ ID NO. 15 or SEQ ID NO. 16. Preferably, the gene cluster encodes a functional gene cluster and optionally, selected tailoring enzymes. The gene cluster may be derived from a single species or may be hybrid in nature. In certain embodiments, the gene cluster is a replacement gene cluster. The replacement gene cluster may be a variant, hybrid, mutant, analog or derivative thereof.

In another embodiment, the invention provides an isolated nucleic acid that encodes three or more ORFs comprising a sequence identical or complementary to all or part of a coding sequence for enzymes performing the biosynthesis of loline alkaloids from Neotyphodium uncinatum. Preferably, the ORFs encode a functional gene cluster and optionally, selected tailoring enzymes. In certain embodiments, an ORF may be derived from a single species or may be hybrid in nature. In certain embodiments at least one of the ORFs is native to the loline alkaloid gene cluster of SEQ ID NO. 15 or SEQ ID NO. 16. In certain other embodiments, at least one of the ORFs is native to SEQ ID NO: 17. In still other embodiments, at least one ORF is derived from a non-loline alkaloid producing Neotyphodium strain, or is hybrid in nature. In yet other embodiments, at least one ORF is a variant, mutant, analog or derivative of the native coding sequence of SEQ ID NO: 15, SEQ ID NO: 16 or SEQ ID NO: 17.

In still another embodiment, the present invention provides isolated nucleic acid compounds comprising three or more genes of the coding sequence for the biosynthesis of loline alkaloids. Preferably, the mixture of genes encode a functional gene cluster and optionally, selected tailoring enzymes. In certain embodiments, a gene may be derived
from a single species or may be hybrid in nature. In certain embodiments at least one gene is derived from a loline alkaloid biosynthetic gene cluster. In other embodiments, at least one gene is derived from a non-loline alkaloid producing *Neotyphodium* strain, or is hybrid in nature. Non-limiting exemplary non-*Neotyphodium* biosynthetic genes are preferably subunits of the *Neotyphodium caudatum* (see e.g., E. S. Han, C. E. Hurst, K. J. Van Etten, and J. M. Konttinen, *Opinion* 40, 32-33, 2001) gene cluster. In yet other embodiments, at least one gene may be a variant, mutant, analog, or derivative of the native coding sequence of SEQ ID NO: 15, SEQ ID NO: 16 or SEQ ID NO: 17. It is also preferred that the encoded activity of the gene is that of, for example and without limitation, an epoxidase, α-type pyridoxal phosphate (PLP) associated enzymes, including, by example, class-V aminotransferase, cytochromes P450, aspartate kinase allosteric amino acid binding domain, oxidoreductase, ornithine decarboxylase, gamma-type PLP enzyme, FAO-containing monooxygenase, and cyclohexanone oxidase.

In another aspect, the present invention provides recombinant expression vectors encoding a loline alkaloid gene cluster, or variants, hybrids, mutants, analogs or derivatives thereof. In certain embodiments, vectors encode one or more subunit of a loline alkaloid gene cluster, or variants, hybrids, mutants, analogs or derivatives thereof.

In another aspect, the present invention provides a host cell transformed with a recombinant expression vector described herein.

In still another aspect, the invention provides a method of preparing loline alkaloid, said method comprising introducing a recombinant vector that encodes a loline alkaloid gene cluster or subunit thereof into a host cell, culturing said host cell under conditions such that loline alkaloid is produced or expressed, and isolating the loline alkaloid from the host cell. In one embodiment, the method is practiced with an *E. coli* host cell. The gene cluster may be a replacement gene cluster and preferably a functional gene cluster. In certain embodiments, the invention provides methods for preparing new alkaloid-type compounds, preferably, loline-type alkaloids. The loline-type alkaloid produced may be loline alkaloid or loline alkaloid variants, hybrids, mutants, analogs or derivatives thereof. Such alkaloids are useful as an insecticide.

These and other embodiments and aspects of the invention will be more fully understood after consideration of the attached Drawings and their brief description below, together with the detailed description, example, and claims that follow.

**BRIEF DESCRIPTION OF THE DRAWINGS**

FIG. 1 shows the structures of the loline alkaloids found in certain grass—Epicolea/Neotyphodium symbionts. N-Formyllooline and N-acetyllooline were abundant in *N. uncinatum* grown in LA-inducing medium.

FIG. 2 is an autoradiograph showing expression of transcripts isolated in the suppression subtractive hybridization in loline-producing (+) and suppressed (−) cultures. In each lane was loaded 0.5 µg of total cDNA synthesized from total RNA and probed with subtracted cDNA: molecular sizes indicated (in kilobases) correspond to molecular marker (HindIII/EcoRI-cut λDNA) run on the same gel.

FIG. 3 is an autoradiograph showing expression of IolA and IolC genes in LA-producing (+) and suppressed (−) cultures. In each lane was loaded 0.5 µg of total cDNA synthesized from total RNA. cDNAs were probed with a mixture of a labeled 523 bp fragment from IolA and a labeled 1427 bp fragment from IolC. Identities of the hybridizing bands were confirmed in separate experiments with the individual probes (data not shown). Bottom panel shows expression of the IolB2 as a control. Molecular sizes (in kilobases) are indicated, and correspond to bands of a DNA-size maker (HindIII/EcoRI-cut λDNA) run in the same gel.

FIG. 4 is a Southern blot of HindIII-digested genomic DNAs probed for IolA (panel A), IolC (panel B), and IolB2 (panel C). Genomic DNAs were from *N. lolii* 138 (lane 1), *E. festucae* CBS 102477 (lane 2), *E. festucae* CBS 102475 (lane 3), and *N. uncinatum* CBS 102646 (lane 4). Numbers adjacent to each lane indicate band sizes (in kilobases) of the molecular marker run in the same gel. For LA phenotype of each species/isolate see Table 3.

FIG. 5 demonstrates the presence of the IolA and IolC genes in endophyte species and isolates differing in LA production. Shown are electropherograms with 2 μl of PCR product loaded in each lane. The multiplex PCR generated a 523 bp product from IolA and a 461 bp product from IolC. The control PCR generated a 726 bp product from IolB2. Numbers above each lane indicate species or isolate listed under the same number in Table 3; lanes B were PCR blanks run without added template DNA; lanes M are molecular size markers (sizes indicated in bp).

FIG. 6 illustrates the *N. uncinatum* Iol clusters 1 (IOL1) (upper bar) and 2 (IOL2) (lower bar). It appears that the IolL2 allele and IolM are linked to IOL2.

**DETAILED DESCRIPTION OF THE INVENTION**

Given the valuable agricultural properties of loline alkaloids, there is a need for methods and reagents for producing large quantities of loline-type alkaloids, for producing loline-type alkaloids in host cells that do not produce loline alkaloids naturally, and for producing novel loline-type alkaloids not found in nature. The present invention provides the protein encoding nucleic acids and methods that produce loline-type alkaloids, with particular application to methods for producing the loline alkaloids and variants, hybrids, mutants, analogs, derivatives and novel compounds related through structure or genetics to loline alkaloid.


All publications, patents and patent applications cited herein, whether supra or infra, are hereby incorporated by reference in their entirety.

**Definitions**

As used herein, the term “alkaloid-type compound” refers to a compound or molecule that is encoded by at least one native alkaloid subunit, or variant, hybrid, mutant, analog, or derivative thereof, including, for example, without limitation, loline-type alkaloid.

As used herein, the term “allele” refers to one of two or more alternate forms of a gene occupying the same locus in
a particular chromosome or linkage structure and differing from other alleles of the locus at one or more mutational sites. Non-limiting types of alleles include, neutral, amorph, hypomorphs, hypermorphs, antimorph, neomorph, isoalleles and unstable alleles.

As used herein the term "coding sequence" or a sequence which "encodes", a particular protein, is a nucleic acid sequence which is transcribed (in the case of DNA) and translated (in the case of mRNA) into a polypeptide in vitro or in vivo when placed under the control of appropriate regulatory sequences. The boundaries of the coding sequence are determined by a start codon at the 5' (amino terminus and a translation stop codon at the 3' (carboxy) terminus. A coding sequence can include, but is not limited to, cDNA from procaryotic or eucaryotic mRNA, genomic DNA sequences from procaryotic or eucaryotic DNA, and even synthetic DNA sequences. A transcription termination sequence will usually be located 3' to the coding sequence.

As used herein the term DNA "control sequences" refers collectively to promoter sequences, ribosome binding sites, polyadenylation signals, transcription termination sequences, upstream regulatory domains, enhancers and the like, which collectively provide for the transcription and translation of a coding sequence in a host cell. Not all of these control sequences need always be present in a recombinant vector so long as the desired gene is capable of being transcribed and translated.

As used herein the term "functional gene cluster" refers to a set of genes (e.g., three or more) or subunits of a biosynthesis gene cluster, which catalyzes the synthesis of an active or functional alkaloid.

As used herein the term "gene" refers to a segment of DNA or its complement that is involved in producing a polypeptide chain, including regions preceding (leader) and following (trailer) the coding sequence as well as intervening sequences (introns) between individual coding sequences (exons). A "loline alkaloid gene" refers to at least any of the ORFs of SEQ ID NO. 15 and SEQ ID NO. 16.

As used herein the term "cluster" refers to a set of (e.g., three or more) closely related genes that code for the same or similar proteins and which are usually grouped together on the same chromosome. A "loline alkaloid gene cluster" refers to a set of genes (e.g., three or more) encoded by at least any of the ORFs of SEQ ID NO. 15 or SEQ ID NO. 16.

As used herein the term "genetically engineered host cell" is meant a host cell where the native gene cluster or subunits thereof has/have been deleted using recombinant DNA techniques. Thus, the term would not encompass mutational events occurring in nature. A "host cell" is a cell derived from a procaryotic microorganism or an eucaryotic cell line cultured as a unicellular entity, which can be, or has been, used as a recipient for recombinant vectors bearing the alkaloid gene clusters of the invention. The term includes the progeny of the original cell which has been transfected. It is understood that the progeny of a single parental cell may not necessarily be completely identical in morphology or in genomic or total DNA complement to the original parent, due to accidental or deliberate mutation. Progeny of the parental cell, which are sufficiently similar to the parent to be characterized by the relevant property, such as the presence of a nucleotide sequence encoding desired biosynthetic enzymes, are included in the definition, and are covered by the above terms.

As used herein the term "loline alkaloid analog" or "analog" refers to a compound or molecule that resembles a loline alkaloid and that contains one or more structural differences relative to the loline alkaloid. Preferably, the loline analog has a desired activity of loline alkaloid although a loline analog may have enhanced or the same activity than products of the loline alkaloid gene cluster. For example, the degree of saturation of at least one bond in the loline alkaloid structure can be changed (e.g., a single bond can be changed to a double or triple bond, or the converse), a bond can be removed, one or more carbon, oxygen or hydrogen atoms can be replaced with a different atom or a chemical moiety (e.g., a halogen, oxygen, nitrogen, sulfur, hydroxy, methoxy, alkyl, aryl, cycloalkyl, heterocycle, amine, amide, ketone, aldehyde, etc.) and the like. Also other peripheral groups, such as OH groups, methyl groups, O-methyl groups, halogenate etc. can be added, modified or removed. Other types of derivatives of loline that would be encompassed by the term "loline alkaloid analog" are known in the art. Non-limiting examples are norloline, N-methyloliline, N-formylolfiline, N-formylolornoline, N-acetylolfiline and N-acetylnololine.

As used herein the term "loline alkaloid derivative" or "derivative" refers to a compound or molecule, that may be produced from loline in one or more steps or with few chemical or moiety modifications.

As used herein the term "loline-type alkaloid" refers to a compound or molecule that is encoded by one or more native gene of, or a variant, hybrid, mutant, analog or derivative thereof, at least SEQ ID NO. 15 or SEQ ID NO. 16.

As used herein, the term "modification enzyme" or" tailoring enzyme" refers to a protein or enzyme that is involved in modifying an alkaloid after its core has been synthesized by the necessary components to catalyze the production of an active or functional alkaloid. Exemplary modification enzymes involved in loline-type alkaloid synthesis include, without limitation, oxidoreductases, dioxygenases and N-methyltransferase.

As used herein, the term "modification step" or "tailoring step" refers to an action or actions taken by a protein or enzyme to modify an alkaloid after its core has been synthesized by the necessary components to catalyze the production of an active or functional alkaloid.

As used herein the term "mutant" refers to a nucleic acid compound, protein, molecule, vector or cell resulting from mutation of the native wild type coding sequence or subunits thereof.

As used herein the term "mutation" refers to any change that alters a native coding sequence either by displacement, addition, deletion, insertion, cross-linking, or other destruction or substitution of one or more nucleotides of the native coding sequence. Techniques for modifying nucleotide sequences, such as site-directed mutagenesis, are also known to those skilled in the art.

As used herein the term "nucleic acid" sequence can include, but is not limited to, procaryotic sequences, eucaryotic mRNA, cDNA from eucaryotic mRNA, genomic DNA sequences from eucaryotic (e.g., mammalian) DNA, and even synthetic DNA sequences, and complements thereof. The term also captures sequences that include any of the known base analogs of DNA and RNA such as, but not limited to 4-acetylcytosine, 8-hydroxy-N6-methyladenosine, aziridinylcytosine, pseudoisocytosine, 5-(carboxyhydroxymethyl)uracil, 5-fluouracil, 5-bromouracil, 5-carboxymethylaminomethyl-2-thiouracil, 5-carboxymethylaminomethyl-2-thiouracil, 5-carboxymethyldihydrouracil, 5-carboxymethyluracil, 5-methyluracil, 5-methylcytosine, N6-isopentenyldenedene, 1-methyladenine, 1-methylguanine, 1-methylguanine, 2,2-dimethylguanine, 2-methylguanine, 2-methylcytosine, 2-methylcytosine, 2-methylguanine, N6-methyladenine,
7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqucosine, 5'-methoxycarbonylmethyluracil, 5-methoxuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxoacyclic acid methylster, uracil-5-oxoacyclic acid, oxybutoxosine, pseudouracil, qucosine, 2-thiocytosine, 5-methyl-2-thiocar-2-thiouracil, 4-thiouracil, 5-thiouracil, N-uracil-5-oxoacyclic acid methylster, uracil-5-oxoacyclic acid, pseudouracil, qucosine, 2-thiocytosine, and 2,6-diaminopurine. A transcription termination sequence will usually be located 3' to the coding sequence.

As used here the term "open reading frame" or "ORF" refers to a region of a nucleic acid molecule that contains a series of triplet bases coding for amino acids without any termination codons. An "open reading frame" includes any start codons.

As used herein the term "replacement gene cluster" is meant any set of genes (e.g., three or more), optionally including genes encoding modification or tailoring enzymes, capable of producing a functional gene cluster when under the direction of one or more compatible control elements, as defined above, in a host cell transformed therewith. The term "replacement gene cluster" encompasses three or more genes encoding for the various proteins necessary to catalyze the production of an alkaid. A "replacement gene cluster" need not include all of the genes found in the corresponding cluster in nature. Rather, the gene cluster need only encode, but is not limited to, the necessary components to catalyze the production of an active alkaid. For example, if the gene cluster includes, for example, eight genes in its native state and only three of these genes are necessary to provide an active alkaid, only these three genes need be present, and a variety of the non-necessary genes may optionally be present. The term, "replacement gene cluster" may also contain genes coding for modification or tailoring enzymes or tailoring enzymes to the core alkaid catalyzed by the necessary components to catalyze the production of an active or functional alkaid. Furthermore, a replacement gene cluster can include genes derived from a single species, or may be hybrid in nature with, e.g., a gene derived from a cluster for the synthesis of a particular alkaid replaced with a corresponding gene from a cluster for the synthesis of another alkaid. Hybrid clusters can include genes derived from different species. The genes included in the replacement gene cluster need not be the native genes, but can be variants, mutants or analogs thereof. Variants are prepared by methods known in the art (see Maniatis et al. Molecular Cloning: A Laboratory Manual (Current Edition)). Mutants or analogs may be prepared by the deletion, insertion or substitution of one or more nucleotides of the coding sequence. Techniques for modifying nucleotide sequences, such as site-directed mutagenesis, are described in the literature. The genes included in the replacement gene cluster need not be on the same plasmid or if present on the same plasmid, can be controlled by the same or different control sequences.

As used herein, the term "subunit" refers to a part of a gene cluster including, for example, a module, domain, gene, or open reading frame, and parts thereof. A "subunit" may comprise for example, a gene or genes derived from a single species or may be hybrid in nature (e.g., a gene derived from a cluster for the synthesis of a particular alkaid replaced with a corresponding gene from a cluster for the synthesis of another alkaid). A "subunit" may comprise variants, mutants, analogs or derivatives of the native gene(s). Variants, mutants, analogs or derivatives thereof may be prepared by techniques known to those of skill in the art, including, without limitation, the displacement, addition, deletion, insertion, cross-linking, or other destruction or substitution of one or more nucleotides of the coding sequence. Techniques for modifying nucleotide sequences, such as site-directed mutagenesis, are also known to those skilled in the art.

As used herein the term "alkaid variant" or "variant" refers to a nucleic acid sequence that hybridizes to an isolated nucleic acid sequence under high stringency conditions and has a desired or enhanced activity of the complement. Variants may include alleles, mutants, hybrids, derivatives, or analogs. Variants also include the polypeptides coded for by these hybridizable nucleic acids.

Identification of loIA and loIC

Production of LA in N. uncinatum can be regulated by culture conditions, such as carbon and nitrogen source and pH in the culture medium, and is completely suppressed in a complex medium (Blankenship et al., 2001, Phytochemistry 58: 395-401), suggesting differential expression of genes involved in LA biosynthesis. Isolation of the genes up-regulated during LA production is a first step in identifying possible enzymes in the biosynthesis of the LA. Different methods are now available for the isolation of differentially expressed genes (Diatczenko et al., 1996, Proc. Natl. Acad. Sci. USA 93: 6025-6030; Liang and Pardee, 1992, Science 257: 967-971), incorporated herein. One such method, suppression subtractive hybridization (Diatczenko et al., 1996; Diatczenko et al., 1999, Meth. Enzymol. 303: 349-380) (incorporated herein), has been particularly useful for identifying differentially expressed genes. This technique was used herein to identify genes up-regulated in N. uncinatum during LA production.

Culture conditions inducing or suppressing LA alkaloid accumulation in the fungus N. uncinatum (Blankenship et al., 2001) were used in combination with suppression subtractive hybridization for isolation of gene transcripts that are up-regulated during LA production. This approach was highly effective in enriching cDNAs differentially expressed in LA-producing cultures: subtracted cDNAs hybridized much more strongly with cDNAs from LA-producing cultures than with cDNAs from LA-suppressed cultures (See FIG. 2). However, a few weak hybridizing bands were present in the total cDNA from the suppressed cultures, which was expected because very low LA levels accumulated in the suppressed cultures. Success of the approach was further indicated by the identification of loIA and loIC, genes that were present only in the species and isolates that produced LA, and, in the case of loIC, related to a sequence previously found to cosegregate with the LA-producing phenotype in Mendelian analysis of F. vesca (Wilkinson et al., 2000, Mol. Plant-Microbe Interact. 13: 1027-1033). The relationships of loIA and loIC to known biosynthetic enzymes further suggested that this approach identified transcripts of LA biosynthesis genes.

The subtracted cDNAs comprised approximately 6-7% of all transcripts present in LA-producing cultures. This estimate appears reliable, since the number of independent clones in the cDNA library from LA-producing cultures (4.1x10^6) greatly exceeds the number of clones (1.0x10^6) estimated to be required for a library representing the complexity of the original mRNA population (Ausubel et al., 2001).

A relatively small number of the subtracted cDNAs and cDNA library clones were sequenced. Rather than conducting extensive sequencing, we focused on some of the cDNAs sequenced in this smaller survey, like loIA and loIC,
giving significant similarity to known genes in amino acid biosynthesis/conversion, to further test their association with LA production. These cDNAs appeared promising candidates, since it has been hypothesized earlier that LA have polyamines as precursors, which in turn are products from amino acid metabolism.

Among cDNAs isolated by the subtraction five independent clones from two alleles of genes designated lolA were identified. However, N. uncinatum has at least two copies of lolA and lolC. The lolA alleles encode predicted proteins significantly similar to aspartate kinases, the first step in biosynthesis of methionine, threonine, and isoleucine from aspartate. In addition, one cDNA clone of a gene, lolC, with similarity to fungal enzymes in methionine biosynthesis was identified. Expression of lolA and lolC was clearly up-regulated in the LA-producing cultures compared to expression in the suppressed cultures. Further evidence for involvement of lolA and lolC in LA production was the distribution of these genes among the Neotyphodium and Epichloë species surveyed, of which eight species produce LA. 12 do not, and one (E. festucae) is polymorphic for this phenotype. Restriction of lolA and lolC to LA-producing endophytes indicated that both genes are either involved or physically linked to genes involved in the LA producing phenotype. This observation, coupled with the observed up-regulation of lolA and lolC in the LA-producing cultures, lent support to an involvement of both genes in LA production.

Generation of knock-outs of lolA and/or lolC will provide further evidence of their roles in LA production. However, preliminary evidence indicates that N. uncinatum has at least two alleles of lolA and the possibility of more than one allele of lolC. Thus, different approaches will be necessary to generate complete knock-outs, one of which could be disruption of the putative lol genes in N. coesophialum for which procedures for knock-outs and double knock-outs have recently been developed.

The ORFs of the lolA alleles in N. uncinatum predicted proteins with lengths of approximately 210 amino acids, much shorter than the sizes of known aspartate kinases (for example, aspartate kinase of Sc. pombe, GenBank accession T39822, has a length of 519 amino acids). Potential reasons for this disparity could include truncation in the RT-PCR due to incomplete extension by the reverse transcriptase, or incorrect annealing of the 5’ and 3’ end-specific cDNA primers to internal gene sequences. cDNA-based northern analysis (see FIG. 3), indicated a strong band of the expected size for an mRNA encoding 210 amino acids, whereas incomplete extension would probably have resulted in multiple bands or smear in the total cDNA. Moreover, despite being very close in size, the two alleles lolA cDNAs varied in the lengths of their 5’ and 3’ terminal sequences (not shown). Because of this difference, truncation due to incorrect primer annealing also appears unlikely, leaving the possibility that the lolA gene encodes a protein much shorter than known aspartate kinases. The predicted lolA amino acid sequences have similarity only to the C-terminal region of aspartate kinase, but not to the N-terminal regions, containing regions for substrate affinity and the active center (Arévalo-Rodríguez et al., 1999). A search of the PROSITE database further indicated that the predicted lolA sequences do not have an N-terminal consensus sequence typical of aspartate kinases. The C-terminal region of aspartate kinases, to which the predicted lolA products have similarity, is thought to be involved in allosteric response of the enzyme (Arévalo-Rodríguez et al., 1999). It is thus possible that the predicted lolA proteins may have a binding site for an allosteric modulator similar to the modulators acting on aspartate kinase, which are normally allosterically regulated by the amino acids lysine, threonine, or isoleucine.

Multiple steps have been identified for the biosynthesis of the more common plant pyrrolizidine, the senecio alkaloids. Senecio alkaloids are synthesized from polyamines, such as putrescine (derived from decarboxylated ornithine) and spermidine. In part because of their structural similarities with senecio alkaloids, a pathway from polyamines has been proposed for LA (Bush et al., 1993). Relative positions of carbon and nitrogen atoms in the 1-aminopyrrolizidine structure (see FIG. 1) would be consistent with spermidine or a related compound as precursor, and spermidine is ultimately derived from the amino acids ornithine and methionine. Aspartate kinase and homocysteine synthase (or related enzymes) are steps in biosynthesis of methionine, which in turn is a precursor to decarboxylated S-adenosylmethionine, the source of the aminopropyl moiety of spermidine. The association of lolA and lolC with LA production indeed suggests possible LA-biosynthesis from aspartate via methionine. However, the substantial differences between predicted lolA and known aspartate kinases may cast doubt on this possibility. Nevertheless, we have observed specific incorporation of 4-[13C]-Asp into LA, indicating that aspartate is a precursor, although the exact sequence of biosynthetic steps remains to be established. Moreover, lolC also had similarity to an enzyme in the biosynthesis of rhizobiotine, a bacterial product which enhances nodulation. The activity of this enzyme encompasses formation of serinol and dihydrorhizobiotine biosynthesis, thus synthesis of compounds different from methionine precursors, further indicating that LA biosynthesis could differ from common amino acid and/or polyamine biosynthesis.

Another cDNA obtained with the subtraction had similarity to a putative zinc-finger transcription factor. Interestingly, in fungi such as Fusarium sporotrichioides and Emmidulans, transcriptional regulators can be part of secondary metabolite pathway clusters, raising the possibility that a specific transcriptional regulator also exists for LA genes. The probable transcription factor found here has similarities to C2H2 zinc-finger transcription factors. A C2H2-like transcription factor was found to be involved in the control of genes in the biosynthesis of trichothecene, a secondary metabolite produced by F. sporotrichioides. In our study, however, detection of the C2H2-like gene did not correlate with LA production in endophytes. Therefore, it likely that this putative transcription factor might be specifically expressed in N. uncinatum under the culture condition used to induce LA production, but may not be a specific regulator for LA biosynthesis genes. Another possibility, however, is that this factor regulates LA genes in N. uncinatum, but different factors regulate the orthologous genes in other endophyte species. In fact, loline alkaloids are not produced by other endophyte species in these culture conditions despite the presence of lolA and lolC, and despite their production of LA when symbiotic with plants. Therefore, the possibility of a unique regulator of LA synthesis in N. uncinatum warrants further investigation.

Other genes up-regulated during LA production that gave significant matches with known genes or sequences were a putative homing endonuclease, generally associated with unusual DNA splicing and incorporation events, and significant matches of cDNAs to sequences in the Neurospora crassa genome. However, for none of these genes do we currently have direct evidence for involvement in LA production. One sequence identified in four clones (K5, K8, C37, D5) was also detectable (by Southern blot) in at least
EXAMPLE 1

All chemicals (including antibiotics) and reagents used in the experiments described in the examples below were obtained from Sigma Corp. (St. Louis, Mo., USA), unless indicated otherwise. Growth media were from Difco Laboratories (Detroit, Mich., USA). Agarose for DNA and RNA gel electrophoreses was supplied by Bio Whittaker Molecular Applications (Rockland, Me., USA). For routine PCR of templates <1.0 kb, AmpliTaq Gold (Applied Biosystems, Foster City, Calif., USA) was used. PCRs for cDNA synthesis, suppression subtractive hybridization, and templates >1.0 kb were performed with the Advantage cDNA PCR Kit (Clontech, Palo Alto, Calif., USA).

Fungal Cultures and Analyses of Loline Alkaloids.

Mycelium of Neotyphodium uncinatum (voucher specimen CBS 102646 at Centraalbureau Voor Schimmelcultures, Utrecht, The Netherlands) was isolated from grass leaf tissues [meadow fescue (Lolium pratense—Festuca pratensis), plant 167 in our plant collection] on potato dextrose agar as previously described (Blankenship et al., 2001). The following procedures were carried out as described by Blankenship et al. (2001) with modifications. After 21 days of growth at 22 °C on PDA plates, 10 fungal colonies were transferred to, and homogenized in, 20 mL of LA-inducing medium (Blankenship et al., 2001) with 15 mM asparagine and 20 mM sucrose as the nitrogen and carbon sources, respectively. Ten ml of the homogenate was added to a 500-ml Erlenmeyer flask with 100 ml of fresh LA-inducing medium, and the culture incubated at 22 °C with rotary shaking (100 rpm). After five days of growth, mycelium was harvested in 50-ml tubes (Falcon, distributed by Becton Dickinson Labware, Lincoln Park, N.J., USA) by centrifugation (2000 g, 10 min), and the mycelium homogenized in 20 mL LA-inducing medium as described. To initiate main cultures for LA production, 1 mL of homogenized mycelium was added to 25 mL of LA-inducing medium and cultures were incubated as described above. To suppress LA production in cultures, but maintain growth conditions similar to the minimal medium, potato dextrose broth was added to give half-strength final concentration in the medium, and asparagine and sucrose were added to 7.5 mM and 10 mM final concentration, respectively. Except for this variation in medium composition, all growth conditions and source of inoculum for LA-suppressed cultures were the same as for LA-induced cultures. Cultures of N. uncinatum were grown under the conditions inducing or suppressing LA accumulation, and harvested during early accumulation when LA levels in the producing medium were <20 µg mL⁻¹. (Levels in similar cultures later reached >1000 µg mL⁻¹ in producing, but <10 µg mL⁻¹ in suppressed cultures.)

LA extraction from freeze-dried culture filtrates or plant tissues, and quantitation by gas chromatography (GC) analysis, were performed as described by Blankenship et al. (2001).

RNA Extraction, DNase Treatment, and Analysis.

Mycelium was harvested by vacuum filtration through Whattman No. 1 filter paper (Whatman International Ltd., Maidstone, England, UK) and total RNA was extracted from 0.2-0.3 g (fresh weight) mycelium with the RNeasy Plant Mini Kit (Qiagen Inc, Valencia, Calif., USA). Co-purified DNA was removed with the DNA-free™ kit (Ambion, Austin, Tex., USA) by treating the extracts (50 µl) with 2 units of DNase I for 30 min at 37 °C, whereupon DNase activity was stopped with DNase Inactivation Reagent (Ambion). Purified RNA was quantified by measuring absorbance at 260 nm and 280 nm in a GeneQuant spectrophotometer (Amersham Pharmacia Biotech, Piscataway, N.J., USA). Integrity of the total RNA was checked by electrophoresis in 1.2% formaldehyde agarose gels.

cDNA Synthesis and Suppression Subtractive Hybridization.

Total RNA was extracted from LA-producing and LA-suppressed cultures. However, low mycelial biomass resulted in low RNA yields. To obtain enough cDNA for subtractive hybridization and expression analysis (cDNA-based Northemns; Endge et al., 1999, BioTechniques 26: 542-548), cDNA was synthesized and amplified with the SMART™ PCR cDNA Synthesis Kit (Clontech). Three µl of RNA solution (300 ng/µl) was reverse transcribed with Superscript™ II following the instructions of the manufacturer (Gibco BRL, Grand Island, N.Y., USA). The reverse-transcription reaction was diluted with TE buffer to a total volume of 50 µl. Amplification of cDNA by long-distance PCR was carried out according to the protocol of the SMART™ PCR cDNA Synthesis Kit (Clontech) in a GeneAmp PCR System 2400 thermocycler (Perkin Elmer Inc., Boston, Mass.). One µl of the diluted reverse-transcription reaction was used, and the number of PCR cycles required for optimum amplification of cDNA was determined according to the manufacturer’s protocol (Clontech). The amplification step allows bulking up on cDNA, while likely maintaining the complexity of the original RNA population.

Suppression subtractive hybridization (Diatchenko et al., 1996; Daitchenko et al., 1999) was performed with the PCR-Select™ cDNA Subtraction Kit (Clontech) essentially as described in the Clontech PCR-Select™ manual. The PCR-Select procedure consists of RsaI digestion of cDNA, ligation of digested tester DNA (containing differentially expressed genes of interest) to two adaptors (1 and 2R, specified in the manual), and two rounds of hybridization with driver DNA used to subtract out cDNAs not differentially expressed in the tester, followed by amplification of the subtracted cDNA by PCR with primers specific to the adaptors. Primary PCR is followed by secondary PCR with nested primers. Only DNA fragments carrying different adaptors at each end tend to amplify exponentially.

cDNA previously amplified with the cDNA Synthesis Kit was purified and digested with RsaI. The digested cDNA was cleaned up with the PCR Purification Kit (Qiagen), eluted into 50 µl of elution buffer, and ethanol precipitated, and adaptors ligated to the tester DNA. In the first hybridization, 13 ng of adaptor-ligated tester was mixed with 147
ng of driver in two separate reactions (each reaction with adaptor 1 and 2R, respectively) and, after denaturation (98°C for 1.5 min), were allowed to anneal for 9 hr at 68°C. After this first hybridization, the two reactions were combined in the presence of 98 ng of denatured fresh driver and a second hybridization performed for 16 hr at 68°C. Amplification of tester-tester hybrids was performed as described in PCR Purification Kit manual. Efficiency of the ligation to the adaptors and of the subtraction was tested and confirmed as for the protocol, using two primers (5'-GGTAGATCTTCCAGAATTCGTCGAG-3' (SEQ ID NO. 1)) and 5'-GGTTTGCCTGGATTCTCAGAC-3' (SEQ ID NO. 2) specific to the β-tubulin gene (tub2).

Upon completion of suppression subtractive PCR, a portion of the product mixture was ligated into pCR®Blunt-TOPO®E, using the Zero Blunt TOPO PCR Cloning Kit (Invitrogen, Carlsbad, Calif., USA), and electroporated into TOP10 cells provided with the kit, to obtain a substracted expressed sequence tag (EST) bank. Another portion was used to generate hybridization probe PCR-labeled with digoxigenin (DIG) following the protocol of the manufacturer (Roche-Boehringer, Indianapolis, Ind., USA).

cDNA Library Construction.
cDNA synthesis and library construction were performed with the SMART™ cDNA Library Construction Kit (Clontech) according to the manufacturer’s instructions. First-strand cDNA was synthesized with the same amount of RNA as used in the cDNA synthesis for the subtraction, and 2 μl undiluted first-strand reaction was used as template to amplify the cDNA. The amplified cDNA was digested with SfiI, size fractionated for removal of low-molecular-size (<0.1 kb) cDNA, and ligated into λTriplEx2 vector (Clontech). cDNA ligated into vector was added to λ phase Gigapack III Gold packaging extract (Stratagene, La Jolla, Calif., USA), and titered in E. coli strain XL1-Blue as specified by the manufacturer. For cDNA library amplification, overnight cultures of XL1-Blue were inoculated with an amount of packaged phage suspension to yield 1×10⁹ pfu per 150 mm plate (Falcon); in total, 15 plates were prepared, so the amplified library was derived from 1.5×10⁹ primary clones. After incubation overnight at 37°C, each plate was added 12 ml of λ dilution buffer (100 mM NaCl, 10 mM MgSO₄, 35 mM Tris-HCl, pH 7.5, 0.01% gelatin), followed by 24 hr incubation at 4°C. The phage suspensions were then titered for each plate. Since differences in titer between plates would affect representation of cDNA clones in the final amplified library, the appropriate volume of each suspension was determined so that, when combined, each plate contributed equally to the total number of pfu in the pooled library. After pooling, the titer of the amplified library was 5.4×10⁶ pfu ml⁻¹.

Southern Blot and PCR Analysis of Genomic DNA.
Fungal genomic DNA was isolated by the method of Al-Samarrin & Schmid, 2000, Lett. Appl. Microbiol, 30: 53-56. Because Neotyphodium occulans does not grow autonomously in culture, DNA from the Lolium multiflorum-N. occulans symbiotum was isolated by the method of Doyle and Doyle, 1990, Focus 12: 13–15 for PCR analysis. Probes for Southern-blot, dot-blot and cDNA-based northern-blot hybridizations were labeled with DIG as described above. Total subtracted cDNA was labeled by using the primary PCR product in the subtraction as template and the nested PCR primers supplied with the PCR-Select™ cDNA Subtraction Kit (Clontech). Probe for lola-3⁰ (5'-GATGCGCCATGTGAGAAGAG-3' (SEQ ID NO. 3)) and lola-3⁰ (5'-GATGCGCCATGTGAGAAGAG-3' (SEQ ID NO. 4)) was a labeled 1427 bp fragment of the lola gene was generated by PCR with primers lola-C-3⁰ (5'-GGTTGCGGTGCTTCTAATACTTGAC-3' (SEQ ID NO. 5)) and lola-C-3⁰ (5'-GATGCGCCATGTGAGAAGAG-3' (SEQ ID NO. 6)) (SEQ ID NO. 1). cDNA-based northerns were performed with complete cDNA, which was gel fractionated and Southern blotted to Hybond™N+ nylon membranes (Amersham Pharmacia Biotech). Southern blotting of DNA by alkaline transfer, as well as dot blotting onto Hybond™N+ nylon membranes (Amersham Pharmacia Biotech) and DNA hybridizations were accomplished with standard protocols (Ausubel et al., 2001). Membranes were washed with 0.1×SSC, 0.1% SDS, once for 15 min at room temperature, then for 20 min and again for 30 min at 75°C. (membranes with cDNA) or at 68°C (membranes with genomic DNA). Chemiluminescent detection of probes hybridized to DNA with anti-DIG antibodies was performed according to the protocol of the supplier (Roche-Boehringer). To visualize labeled probes hybridizing to DNA, membranes were exposed to Hyperfilm™ ECL™ Chemiluminescence film (Amersham Pharmacia Biotech).

PCR screening for lolA was performed on endophyte genomic DNA with primers lolA-3⁰ and lolA-5⁰. PCR screening for lolA employed primers lolA-C-3⁰ (5'-GGTTGCGGTGCTTCTAATACTTGAC-3' (SEQ ID NO. 7)) and lolA-C-5⁰ (5'-GGTTGCGGTGCTTCTAATACTTGAC-3' (SEQ ID NO. 8)). PCR was performed with 35 cycles of 95°C for 30 s, 62°C for 30 s, and 72°C for 1 min. As a positive control for DNA integrity in this screening, a tub2 gene fragment was amplified by PCR with primers 5'-TGCTAACAGCTCAGAC-3' (SEQ ID NO. 9) and 5'-GAGAAATGCTGAGATGTG-3' (SEQ ID NO. 10) (Byrd et al., 1990), with 35 cycles of 95°C for 30 s, 55°C for 30 s, and 72°C for 1 min.

cDNA Library Screening and Conversion of Phage to Plasmid Clones.
Screening of the cDNA library as described by Ausubel et al. (2001). Phage were plated in a lawn of E. coli XL1-Blue, and plaque lifts were on Hybond™N+ nylon membranes (Amersham Pharmacia Biotech). To convert clones in λTriplEx2 to plasmid form, plasmids were added to E. coli strain BM25.8 (which expresses Cre-recombinase) as per the supplier’s protocols (Clontech). Single, isolated colonies were selected on LB agar with ampicillin, picked and grown in LB with ampicillin, and plasmids isolated by a rapid alkaline procedure (Ahn et al., 2000, BioTechniques 29: 266–368). To verify that a plasmid carried the desired insert, 3 μl of each plasmid was spotted onto a nylon membrane for dot blotting, and the membrane hybridized to the probe initially used to identify the λ-phase clone.

Plasmid DNA Isolation, Sequencing, and Database Search of cDNAs.
Plasmid DNA was isolated from bacterial cells by the rapid alkaline miniprep procedure (Ahn et al., 2000). Plasmid inserts were sequenced with primers L-triplEx 5' (5'TCCGGAGATCTGACGACGAGC-3') (SEQ ID NO. 11) and L-triplEx 3' (5'TAATACGACTCACTATAGGG-3') (SEQ ID NO. 12), specific to vector regions flanking the cDNA inserts. DNA cloned into TOPO vector (Invitrogen) was sequenced with M13-reverse (5'CAGGAAAACCTAATGACGAC-3') (SEQ ID NO. 13) and M13-forward (5'CAGGAAAACCTAATGACGAC-3') (SEQ ID NO. 14) primers.

Sequencing of DNA was performed with the BigDye Ter-
Sequences of subtracted cDNA clones, as well as inserts in library clones that hybridized to subtracted cDNAs, were used to query databases by various BLAST algorithms. For several subtracted cDNAs (Table 1) and cDNAs from the λ-library (Table 2), matches to known protein sequences in NCBI, or sequences in the N. crassa database were identified. Five library and subtracted clones had significant similarity to the C-terminal amino acid regions of aspartate kinases (Tables 1 & 2). Two of the library clones with similarity to aspartate kinase were identical to each other in sequence (P2 and P16). However, a third library clone (P17) differed from the other two (94% identity), but had 100% sequence identity with two subtracted clones (B8 and C5). The detected variation in sequence among the clones suggested more than one form of this gene in N. uncinatum. The presence of two genomic alleles of the AspK-related gene in N. uncinatum was verified by PCR with primers with allele-specific nucleotides at their 3'-ends (data not shown). Because other results of this study (described below) strongly associate these sequences with LA production, we will hereafter designate the corresponding genes as lola alleles; the two allelic sequences have been submitted to GenBank (accessions AF439396 and AF439395).

The predicted proteins (209 and 210 aa) encoded by lola alleles were smaller than known aspartate kinases (which usually exceed 500 aa) and had similarity only to the C-terminal regions of aspartate kinases. This was indicated by protein sequence alignments with known aspartate kinases that gave the most significant matches in BLAST. These proteins, from Schizosaccharomyces pombe and Saccharomyces cerevisiae (GenBank accessions T9822 and P10869), aligned with a region starting at amino acid position 47 and ending at amino acid 204 of the predicted lola proteins and had 25–26% identity to lola, but only within a region of the known aspartate kinases starting at about amino acid 351 and ending at amino acid 495. PROSITE searches with the predicted amino acid sequences of the two lola alleles indicated that both lacked the aspartate kinase signature, defined by PROSITE as [LIVM]-x-K-[FY]-G-G-[ST]-[SC]-[LIVM], a conserved region located near the N-terminal end of aspartate kinases.

One subtracted clone, D6 (Table 1), gave highly significant matches with fungal (N. crassa, Emericella nidulans, and S. cerevisiae) genes for O-acetylhomoserine-thiol)lyase (homocysteine synthase), and the related enzymes, cystathionine γ-synthase and cystathionine β-lyase, all of which are γ-type pyridoxal phosphate-containing enzymes in sulfur-containing amino acid biosynthesis and interconversion pathways. Additionally, significant similarity was found with an enzyme in the biosynthesis of the bacterial compound riboflavin. The molecular size of the transcript (between 1.5 to 2 kb, FIG. 3) predicted a protein similar in size to known homocysteine synthases, which are 430–450 amino acids (Sienko et al., 1998). A related sequence was recently identified at a locus associated with LA production in Epicloë fuscata (Wilkinson et al., 2000; Spiering et al., 2000). Because these data and further evidence presented below associated this sequence with LA production, the corresponding gene was designated lola. (GenBank accessions AF461175, AF461176).
TABLE 1

<table>
<thead>
<tr>
<th>Clones(^1)</th>
<th>Length in bp nr matches, identify (%), and E values</th>
<th>Ns. crassa matches, identify (%), and E values</th>
</tr>
</thead>
<tbody>
<tr>
<td>K8, C37, D5</td>
<td>448</td>
<td>—</td>
</tr>
<tr>
<td>B8, C5(^2)</td>
<td>633 Sc. pomb(e) aspartate kinase gene, 24%, 5e-07</td>
<td>—</td>
</tr>
<tr>
<td>N17, C7</td>
<td>1521</td>
<td>—</td>
</tr>
<tr>
<td>C2, D1</td>
<td>724 Krüppel-like C2H2 zinc finger transcription factors, 44%, 7e-08</td>
<td>various contigs (1.246; 1.392; 1.622; 1.686; 1.151), 35-52%, 4e-20 to 1e-05</td>
</tr>
<tr>
<td>C1, C3</td>
<td>283</td>
<td>—</td>
</tr>
<tr>
<td>E21</td>
<td>388</td>
<td>—</td>
</tr>
<tr>
<td>A6</td>
<td>370</td>
<td>—</td>
</tr>
<tr>
<td>A7</td>
<td>379</td>
<td>—</td>
</tr>
<tr>
<td>A8</td>
<td>554</td>
<td>—</td>
</tr>
<tr>
<td>C8</td>
<td>430 Sc. pomb(e) hypothetical protein, 42%, 2e-05</td>
<td>Contig 1.291 (57.61-57.83 kb), 56%,</td>
</tr>
<tr>
<td>D2</td>
<td>472</td>
<td>—</td>
</tr>
<tr>
<td>D3</td>
<td>249</td>
<td>—</td>
</tr>
<tr>
<td>D4</td>
<td>694 rRNA intron-encoded homing endonuclease, 80%; 2e-11</td>
<td>various contigs (2.820; 2.798; 2.816; 2.793; 2.790; 2.943); 2.786; 2.843; 2.831; 2.957); 4e-10 to 3e-08</td>
</tr>
<tr>
<td>D6(^3)</td>
<td>374 homocysteine synthase/D-acylhomoserine synthase, 53%, 1e-22; related enzymes in methionine/cysteine biosynthesis, 1e-07; RtxA, enzyme in rhizobiotin biosynthesis, 37%, 1e-10</td>
<td>Contig 2.65, 34%, 3e-22; Contig 2.688, 34%, 3e-11</td>
</tr>
</tbody>
</table>

\(^1\) Listed in order of frequency from most common to least common.
\(^2\) = No significant match.
\(^3\) = lola clones.
\(^4\) = Saccharomyces pomb\(e\).
\(^5\) = loloA clone.

TABLE 2
cDNA library clones hybridizing to total subtracted cDNA.

<table>
<thead>
<tr>
<th>Clone(^6)</th>
<th>Length of nucleotide sequence [bp]</th>
<th>Length of putative ORF [aa]</th>
<th>Identical in sequence to subtracted clones</th>
</tr>
</thead>
<tbody>
<tr>
<td>P2, P15(^7)</td>
<td>838, 880(^7)</td>
<td>210</td>
<td>none</td>
</tr>
<tr>
<td>P3</td>
<td>446</td>
<td>35</td>
<td>K8, C37, D5</td>
</tr>
<tr>
<td>P15</td>
<td>725</td>
<td>111</td>
<td>none</td>
</tr>
<tr>
<td>P17</td>
<td>774</td>
<td>209</td>
<td>B8, C5</td>
</tr>
<tr>
<td>P18</td>
<td>449</td>
<td>30</td>
<td>none</td>
</tr>
</tbody>
</table>

\(^6\) = lola clones
\(^7\) = Difference in length of 3’ untranslated region.
\(^\ldots\) = No significant match.

EXAMPLE 4

Association of the lola and loloC Genes with Loline Production.

Genomic sequences of loloC and one allele of lola (clone P1.7) were obtained by using primers based on the cDNA of the lola gene and genomic sequence of the loloC gene from E. festucae (data not shown). This information was used to design primers for specific probes and detection of lola and loloC sequences in cDNA-based northern analysis of complete cDNAs from LA-producing and suppressed cultures (FIG. 3). Both sequences were expressed in the LA-producing cultures. Strong hybridizing bands were detected from the complete cDNA from LA-producing cultures, whereas faint bands were obtained from the complete cDNA from the suppressed cultures.

LA production is a trait specific to endophyte species (Christensen et al., 1993; Siegel et al., 1990;TelPaske and Powell, 1993) or even isolates within species (Wilkinson et al., 2000). Consequently, we reasoned that genes associated with LA production would be present in all LA-producing endophytes, but might be absent from endophytes that do not produce LA. For many endophyte species and isolates available from our collection the LA phenotypes were known from the literature (Table 3), and these were confirmed by GC analyses of plants symbiotic with these endophytes. Additional species or isolates included in this survey were similarly assessed for LA production (Table 4). In Southern-blot analysis of genomic DNAs from two LA producers and two nonproducers, lola and loloC sequences hybridized only with DNA from the endophytes that produce LA (FIG. 4). The probes used to detect lola and loloC did not have sites for the restriction enzyme used in the genomic digests, so for each putative allele one hybridizing band was expected. In N. uncinatum, two bands were observed from the genomic DNA probed with lola, indicating at least two alleles of this gene; hybridization with loloC gave only one band, suggesting only one allele of this gene was present, but the possibility that this single band represented multiple alleles of loloC could not be excluded. In E. festucae, hybridization with the two probes gave one strong hybridizing band for each, suggesting one allele of each gene. The additional, fainter hybridizing bands present on the blots corresponded to some bands on the ethidium bromide-stained gel (not shown) and were, therefore, likely due to non-specific binding of the probes to mitochondrial or repetitive genomic DNA.

Diagnostic PCR was used with primers specific to the lola and loloC genes for detection of these sequences in all
species and isolates listed in Table 3. Detection of the lolA and lolC genes in endophytes was strictly associated with the LA-producing phenotype (FIG. 5). In addition, the two genes were detected in N. chissoa (ATCC 64037).

The high expression of lolA and lolC in LA-producing 5 cultures of N. uncinitatum, and the strict correlation of LA production with presence of the two genes in the different endophytes, lent strong support to involvement of the lolA and lolC genes in LA production.

### TABLE 3
LA phenotype of endophyte species and isolates used in this study. Indicated are the respective grass hosts which were used in the determination of the LA, and from which the endophytes in this study were originally isolated.

<table>
<thead>
<tr>
<th>Species/isolate</th>
<th>Grass host</th>
<th>Loline phenotype</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) E. festucae CBS 102477</td>
<td>Festuca rubra</td>
<td>–</td>
<td>1</td>
</tr>
<tr>
<td>2) E. festucae CBS 102475</td>
<td>N/A</td>
<td>+</td>
<td>2</td>
</tr>
<tr>
<td>3) E. typhina</td>
<td>Lolium perenne</td>
<td>–</td>
<td>3</td>
</tr>
<tr>
<td>4) Neotyphodium avenae CBS 109345</td>
<td>Echinochogon ovatus</td>
<td>–</td>
<td>4</td>
</tr>
<tr>
<td>5) N. avenae ATCC MYA-1231</td>
<td>E. ovatus</td>
<td>+</td>
<td>5</td>
</tr>
<tr>
<td>6) N. australiensis CBS 109346</td>
<td>E. ovatus</td>
<td>–</td>
<td>6</td>
</tr>
<tr>
<td>7) M. cernuum ATCC 90604</td>
<td>Lolium arundinaceum</td>
<td>–</td>
<td>7</td>
</tr>
<tr>
<td>8) N. hungaricum ATCC 604040</td>
<td>Festuca arizonica</td>
<td>–</td>
<td>8</td>
</tr>
<tr>
<td>9) N. inornata CBS 109347</td>
<td>Achthatherium inornata</td>
<td>–</td>
<td>9</td>
</tr>
<tr>
<td>10) C. lolii 138</td>
<td>L. perenne</td>
<td>–</td>
<td>10</td>
</tr>
<tr>
<td>11) N. meliciola CBS 109342</td>
<td>Melica decumbens</td>
<td>–</td>
<td>11</td>
</tr>
<tr>
<td>12) N. occultans 999</td>
<td>Lolium multiplanum</td>
<td>–</td>
<td>12</td>
</tr>
<tr>
<td>13) N. steigelti ATCC 74483</td>
<td>Lolium perenne</td>
<td>–</td>
<td>13</td>
</tr>
<tr>
<td>14) Neotyphodium sp. 55</td>
<td>Poa annua</td>
<td>+</td>
<td>14</td>
</tr>
<tr>
<td>15) Neotyphodium sp. 87</td>
<td>Festuca paradoxa</td>
<td>–</td>
<td>15</td>
</tr>
<tr>
<td>16) Neotyphodium sp. LpTG-2 Lp1</td>
<td>L. perenne</td>
<td>–</td>
<td>16</td>
</tr>
<tr>
<td>17) Neotyphodium sp. 269</td>
<td>Hordeum bogdanii</td>
<td>–</td>
<td>17</td>
</tr>
<tr>
<td>18) Neotyphodium sp. 270</td>
<td>Hordeum brevisubulatum</td>
<td>–</td>
<td>18</td>
</tr>
<tr>
<td>19) Neotyphodium sp. 361</td>
<td>Hordeum arundinaceum</td>
<td>–</td>
<td>19</td>
</tr>
<tr>
<td>20) Neotyphodium sp. FcTG-3 Tg18</td>
<td>L. arundinaceum</td>
<td>–</td>
<td>20</td>
</tr>
<tr>
<td>21) Neotyphodium sp. FcTG-2 Tg14</td>
<td>L. arundinaceum</td>
<td>–</td>
<td>21</td>
</tr>
<tr>
<td>22) Neotyphodium sp. 4096</td>
<td>Achthatherium robustum</td>
<td>–</td>
<td>22</td>
</tr>
<tr>
<td>23) N. uncinitatum CBS 102646</td>
<td>L. pratense</td>
<td>+</td>
<td>23</td>
</tr>
</tbody>
</table>

*CBS = Centraalbureau voor Schimmelcultures, Fungal Biodiversity Center, Utrecht, The Netherlands (http://www.cbs.knaw.nl), ATCC = American Type Culture Collection, Manassas, Virginia, USA (http://www.atcc.org). Other designations are from the referenced papers or are laboratory isolate numbers.

* a = loline-producing, ‘—’ = loline non-producing.

*References are (1) Leuchtenmann and Schardl (1998), (2) Wilkinson et al. (2000), (3) Siegel et al. (1999), (4) this study (see Table 6), (5) Miles et al. (1996), (6) TePaske and Powell (1993), (7) Craven et al. (2001), (8) Chitester et al. (2001), (9) Leuchtenmann et al. (2000).

*Numbers before each isolate correspond to the numbers indicated above gel lanes in FIG. 6.

*NA = not applicable. Isolate derived by mating E. festucae CBS102477 with an E. festucae isolate CBS 102474 from Lolium giganteum.

### TABLE 4
LA in plants with endophyte species for which the LA phenotype was previously unknown.

<table>
<thead>
<tr>
<th>Species/isolate</th>
<th>Host grass</th>
<th>Loline</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neotyphodium avenae CBS 109345</td>
<td>Echinochogon ovatus</td>
<td>1780</td>
</tr>
<tr>
<td>N. avenae ATCC MYA-1231</td>
<td>E. ovatus</td>
<td>2120</td>
</tr>
<tr>
<td>N. australiensis CBS 109346</td>
<td>E. ovatus</td>
<td>nd</td>
</tr>
<tr>
<td>N. meliciola CBS 109342</td>
<td>Melica decumbens</td>
<td>nd</td>
</tr>
<tr>
<td>Neotyphodium sp. FcTG-3 Tg18</td>
<td>Lolium arundinaceum</td>
<td>670</td>
</tr>
<tr>
<td>Neotyphodium sp. FcTG-2 Tg14</td>
<td>L. arundinaceum</td>
<td>nd</td>
</tr>
<tr>
<td>Neotyphodium sp. 269</td>
<td>Hordeum bogdanii</td>
<td>nd</td>
</tr>
</tbody>
</table>

*CBS = Centraalbureau voor Schimmelcultures, Fungal Biodiversity Center, Utrecht, The Netherlands (http://www.cbs.knaw.nl), ATCC = American Type Culture Collection, Manassas, Virginia, USA (http://www.atcc.org). Other designations are from the referenced papers or are laboratory isolate numbers.

*Reported is the sum of N-formyl and N-acetyl loline in μg g⁻¹ dry weight plant tissue.

*nd = not detected (limit of detection = 10 μg g⁻¹).
EXAMPLE 5

Additional Subtracted cDNAs Matching Known Genes and Genomic Sequences.

As shown in Tables 1 & 2, several other cDNAs isolated by the subtraction method also gave highly significant matches in BLAST searches of the nr and *Ns. crassa* databases. Matches included a zinc-finger transcription factor, a hypothetical protein in *S. cerevisiae*, and a homing endonuclease. Additionally, matches with *Ns. crassa* sequences were identified for putative ORFs of one hybridizing library clone, P15, and one subtracted clone, A8.

A survey of the distribution of putative zinc-finger transcription factor among eight endophytes differing in LA production (four LA producers and four non-producers) was performed by diagnostic PCR. There was no association of this putative zinc-finger transcription factor gene with LA production; its presence was detected only in two isolates, one LA-producer (*N. uncinatum*) and one non-producer (*N. huerfanum*) (data not shown).

As further indicated in Tables 1 & 2, a number of sequences from library and subtracted clones gave no significant matches with known genes in the nr database and sequences in the *Ns. crassa* genome. For one subtracted clone, N17, the full-length cDNA sequence was obtained by PCR, using an aliquot of the amplified cDNA library and gene and vector-specific primers. A predicted ORF of 363 amino acids was found within the N17 cDNA (data not shown), but this amino acid sequence did not give significant matches with any genes or sequences in the nr and *Ns. crassa* databases or known protein patterns in the PROSITE database.

EXAMPLE 6

Identification of the LOL1 and LOL2 Gene Clusters

Central to the present invention is the identification of the loline alkaloid gene clusters LOL1 (SEQ ID NO. 15), and LOL2 (SEQ ID NO. 16) which apparently may also include lolF2 and lolM (SEQ ID NO. 17). The association in *Neotyphodium uncinatum* of lolA and lolC was tested by long-distance-PCR. The 8.2 kb product contained the expected sequences of both, plus two additional open reading frames between them. We then walked outward from this fragment by vector-mer mediated PCR; and in the process identified two gene clusters (LOL1 and LOL2 in FIG. 6).

In addition to lolC and lolA, at least 8 (LOL1) or 7 (LOL2) ORFs were inferred within LOL1 and LOL2 by using a program with an algorithm for prediction of fungal genes. PCR analyses/Southern hybridization on a cDNA library/total cDNA from *N. uncinatum* showed expression of the ORFs lolM, lolF, lolC, lolO, lolA, and lolE, indicating that these contain active genes. The details of the gene predictions and coordinates, i.e., location of the exons in the ORFs of LOL1 and LOL2 are given below. LOL1 and LOL2 differ in sequence (LOL1 has ~95% nucleotide sequence identity to LOL2), thus represent two truly distinct genomic regions. Altogether, ten genes were inferred in the gene clusters, with most of the genes shared between the clusters. PCR and Southern-blot analyses indicated that all ten genes were unique to the loline alkaloid producers among the isolates surveyed in Table 3 (Fung. Genet. Biol. Spiertz et al.). Nine of the genes, lolE, lolT, lolP, lolU, lolA, lolO, lolD, lolC, and lolF, were found in two allelic forms.

The amino acid sequences deduced from the LOL gene ORFs gave highly significant matches (E<10^-5, except lolU and lolM for which E=0.01) with known enzyme sequences in the protein databases curated by the National Center of Biotechnology Information (http://www.ncbi.nlm.nih.gov; NCBI). The gene functions predicted by GenBank searches of the databases at NCBI, and gene orientations within the clusters, thereby indicate that LOL2 contains eight genes (i.e., lolE, lolT, lolP, lolU, lolA, lolO, lolD, and lolC) representing alleles of genes present in LOL1. LOL1 contains an additional gene, namely lolF, hitherto not found in LOL2. The genomic location of the additional ORFs, lolF2 and lolM, relative to the two LOL clusters is presently unknown, but we postulate that lolF2 and lolM are located close to LOL2. lolF2 and sequence adjacent to it has ~93% identity to lolF (and sequence adjacent to it) in cluster LOL1.

The LOL1 gene cluster spans about a 25.3 kb region and consists of 9 ORFs. Open reading frames of LOL1 are indicated relative to nucleotide numbers annotated to SEQ ID NO: 15; mRNA sequences of each gene are given by joined exons determined by cDNA sequencing or predicted by the fgene server (Neurospora) gene prediction program at “Softberry”, http://www.softberry.com/beryl.php?topic=fgnd; gene orientations are indicated by “+” (forward strand) and “-” (reverse strand).

ORF1 of LOL1 is lolE: +strand, join 23457–24195, 24275–24306 (predicted by fgenesh at Softberry).


ORF3 of LOL1 is lolP: +strand, join 19245–19554, 19639–20225, 20287–20694, 20818–20846, 20919–21045 (predicted by fgenesh at Softberry).


ORF5 of LOL1 is lolA: +strand, join 14951–15476, 15545–15648 (determined by sequencing of lolA cDNA).

ORF6 of LOL1 is lolO: –strand, join 13961–13770, 13781–13677 (predicted by fgenesh at Softberry).


ORF8 of LOL1 is lolC: +strand, join 6903–7000, 7063–7114, 7199–7282, 7364–7723, 7810–8364, 8435–8709 (determined by sequencing of lolC cDNA).

ORF9 of LOL1 is lolF: –strand, join 5905–5028, 4960–3509, 3448–3346 (predicted by fgenesh at Softberry).

The LOL2 gene cluster spans about a 16.4 kb region and consists of at least 8 ORFs. It appears that LOL2 may include lolF2 and lolM (SEQ ID NO: 17) linked to the 5’ end of SEQ ID NO: 16, in which case the LOL2 gene cluster would span about a 24 kb region, consisting of 10 ORFs (i.e., ORF1’ through ORF10’). ORFs of LOL2 are indicated relative to nucleotide numbers annotated to SEQ ID NO: 16; mRNA sequences of each gene are given by joined exons determined by cDNA sequencing or predicted by the fgenesh (Neurospora) gene prediction program at “Softberry”, http://www.softberry.com/beryl.php?topic=fgnd; gene orientations are indicated by “+” (forward strand) and “-“ (reverse strand).

ORF1’ of LOL2 is lolE: +strand, join 15210–15946, 16026–16057 (predicted by fgenesh at Softberry).


ORF4’ of LOL2 is lolU: 3-strand, join 10438–9597, 9531–9816 (predicted by fgenesh at Softberry).

ORF5’ of LOL2 is lolA: 5-strand, join 8006–8534, 8605–8706 (predicted by sequencing of lolA cDNA).

ORF6’ of LOL2 is lolA: 3-strand, join 7190–6999, 6907–6011 (predicted by fgenesh at Softberry).


It also appears that LOL2 may include lolF2, an allele of lolF, and lolM, probably linked to the 5’ end of LOL2 (SEQ ID NO: 16). The ORF’s of lolF2 and lolM are indicated relative to nucleotide numbers annotated to sequence of SEQ ID NO: 17; mRNA sequences of each gene are given by joined exons predicted by fgenesh (Neurospora) gene prediction program at “Softberry”, http://www.softberry.com/berry.phtml?topic=gflnd; gene orientations are indicated by “+” (forward strand) and “-” (reverse strand).

ORF9’ is lolF2: 2-strand, join 5804–4342, 4281–2027, 3905–3821 (predicted by fgenesh at Softberry).

ORF10’ is lolM: 3-strand, join 1689–1525, 1430–1332, 1231–1174, 1085–1021 (predicted by fgenesh at Softberry).

EXAMPLE 7

Functional Assignment of the Loline Alkaloid Gene Clusters

Most of the predicted gene products show highly significant BLAST matches (E ≤ 10^-7) with known biosynthetic enzymes and motifs. The closest BLAST matches and/or motifs of the ten genes follow in the order that they occur in the clusters: lolE gave a match to epoxidases; lolT and lolT2 matched the diagnostic domain of ɑ-type pyridoxal phosphate (PLP)-associated enzymes, including class-ɑ amine oxidases; lolP matched cytochromes P450, with closest relationship to pisin demethylase from Nectria haematococca; lolU gave no significant match or diagnostic motif; lolA closely matched the Asp kinase amino acid binding domain; lolO matched nonheme-Fe oxidoreductases, especially isopenicillin N synthase; lolD matched ornithine decarboxylase (ɑ-type PLP enzyme); lolC appeared to be a ɑ-type PLP enzyme; lolF and lolF2 appeared to encode an FAD-containing monoxygenase with closest match to cyclohexanone oxidase; lolM had no significant BLAST match or motif.

EXAMPLE 8

Hybridizability Variants

The nucleic acids of the present invention comprise at least a nucleotide sequence of all or part of SEQ ID NO: 15 or SEQ ID NO: 16 or variants thereof. It also appears that SEQ ID NO: 17 or variants thereof may be part of the LOL2 gene cluster linked to the 5’ end of SEQ ID NO: 16, and therefore, nucleic acid sequences that hybridize to all or part of SEQ ID NO: 17 are also encompassed by the present invention. Variants of the present invention encode isolated nucleic acids that at least hybridize to all or part of SEQ ID NO: 15 or SEQ ID NO: 16 or the complements thereof under hybridization conditions of, at, or between, low and high stringency conditions, and have insecticidal activity. Low stringency conditions are generally about 3xSSC at about 45° C. to about 65° C., and high stringency conditions are generally about 0.1xSSC, 0.1% SDS at about 65° to 68° C.

Preferably, the hybridization conditions are highly stringent at 0.1xSSC, 0.1% SDS at 65° C. Variants are made by methods known to one of ordinary skill in the art and as set forth in Maniatis et al: Molecular Cloning: A Laboratory Manual (Current Edition). Preferably, the hybridized nucleic acids code for a polypeptide that has one or more or all of the physical and/or biological properties of loline alkaloids, such as insecticidal activity and feeding deterrent properties.

EXAMPLE 9

Host-Vector System

Identification and cloning of the loline alkaloid gene clusters is useful for the development of a host-vector system for the efficient recombination production of both novel and known alkaloids. The coding sequences which collectively encode a loline-type alkaloid gene cluster, including variants, hybrids, mutants, analogs or derivatives of the loline alkaloid gene cluster, can be inserted into one or more expression vectors, using methods known to those of skill in the art. The replacement gene cluster need not correspond to the complete native loline alkaloid gene cluster, but need only encode a functional gene cluster to catalyze production of an alkaloid.

The recombinant vector(s) of the present invention includes replacement gene clusters derived from a single gene cluster, or may comprise hybrid replacement gene clusters with, e.g., a gene of one cluster replaced by the corresponding gene from another gene cluster. For example, the oxidoreductase of LOL1 may be replaced with the oxidoreductase of LOL2 without an effect on the product structure. Accordingly, these genes may be freely interchangeable in the constructs described herein. Thus, the replacement clusters of the present invention can be derived from any combination of alkaloid gene sets, which ultimately function to produce an identifiable alkaloid.

Expression vectors also include control sequences operably linked to the desired loline coding sequence. Suitable expression systems for use with the present invention include systems, which function in eucaryotic and proaryotic host cells. However, proaryotic systems are preferred, and in particular, systems compatible with Neotyphodium, Epichloë, Adenocarpus and Argyreia mollis species are of particular interest. Control elements for use in such systems include promoters, optionally containing operator sequences, and ribosome binding sites. Particularly useful promoters include control sequences derived from alkaloid gene clusters. However, other bacterial promoters, such as those derived from sugar metabolizing enzymes, such as galactose, lactose (lac) and maltose, will also find use in the present constructs. Additional examples include promoter sequences derived from biosynthetic enzymes such as tryptophan (trp), the beta-lactamase (bla) promoter system, bacteriophage lambda P1, and T5. In addition, synthetic promoters, such as the tac promoter, which do not occur in nature also function in bacterial host cells.

Other regulatory sequences may also be desirable which allow for regulation of expression of the replacement gene cluster relative to the growth of the host cell. Regulatory sequences are known to those of skill in the art, and examples include those which cause the expression of a gene to be turned on or off in response to a chemical or physical stimulus, including the presence of a regulatory compound. Other types of regulatory elements may also be present in the vector, for example, enhancer sequences.

Selectable markers can also be included in the recombinant expression vectors. A variety of markers are known
which are useful in selecting for transformed cell lines and generally comprise a gene whose expression confers a selectable phenotype on transformed cells when the cells are grown in an appropriate selective medium. Such markers include, for example, genes which confer antibiotic resistance or sensitivity to the plasmid.

The various subunits of gene clusters of interest can be cloned into one or more recombinant vectors as individual cassettes, with separate control elements, or under the control of, e.g., a single promoter. These subunits can include flanking restriction sites to allow for the easy deletion and insertion of other subunits so that hybrid gene clusters can be generated. The design of such unique restriction sites is known to those of skill in the art and can be accomplished using the techniques described above, such as site-directed mutagenesis and PCR.

Further, the vectors, which collectively encode a replacement gene cluster can be inserted in to one or more host cell, using methods known to those of skill in the art. As such, the present invention also provides host cells which have their naturally occurring gene substantially deleted, transformed with vectors encoding a replacement gene cluster or parts thereof, for the production of active alkaloids. The invention provides for the production of significant quantity of product at an appropriate stage of the growth cycle. The alkaloids so produced can be used as insecticidal and feeding-deterrent agents to protect plants. The ability to recombinantly produce alkaloids also provides a powerful tool for characterizing biosynthetic enzymes and the mechanism of their actions.

More particularly, host cells for the recombinant production of the subject alkaloids can be derived from any organism with the capability of harboring a recombinant gene cluster. Thus, the genetically engineered host cells of the present invention can be derived from either procaryotic or eucaryotic organisms. Preferably, the host may be E. coli. However, more preferred host cells are those constructed from the Neotyphodium species, among others, will provide convenient host cells for the subject invention.

The above-described host cells are genetically engineered by deleting the naturally occurring loline alkaloid genes or genes encoding tailoring enzymes therefrom, using standard techniques, such as by homologous or heterologous recombination. One or more recombinant vector, collectively encoding a replacement gene cluster of the present invention, is then introduced into a host cell. The vector(s) can include native or hybrid combinations of loline alkaloid gene cluster subunits, or mutants, analogs, or derivatives thereof. Methods for introducing the recombinant vectors of the present invention into suitable host cells are known to those of skill in the art and typically include the use of CaCl₂ or other agents, such as divalent cations and DMSO. DNA can also be introduced into bacterial cells by electroporation. Once the genes or gene clusters are expressed, the alkaloid producing colonies can be identified and isolated using known techniques. The produced alkaloids can then be further characterized, e.g. by NMR and mass spectroscopy.

Although illustrative embodiments of the present invention have been described in detail, it is to be understood that the present invention is not limited to those precise embodiments, and that various changes and modifications can be effected therein by one skilled in the art without departing from the scope and spirit of the invention as defined by the appended claims.

---

SEQUENCE LISTING

| SEQ ID NO | 1 |
|-----------|
| LENGTH    | 23 |
| TYPE:     | DNA |
| ORGANISM: | Unknown |
| FEATURE:  |
| OTHER INFORMATION: Artificial primer sequence |
| SEQUENCE: |
| gttgtctcc agatccgtg agg |

| SEQ ID NO | 2 |
|-----------|
| LENGTH    | 20 |
| TYPE:     | DNA |
| ORGANISM: | Unknown |
| FEATURE:  |
| OTHER INFORMATION: Artificial primer sequence |
| SEQUENCE: |
| gttgtccg agtttcgac |

| SEQ ID NO | 3 |
|-----------|
| LENGTH    | 23 |
| TYPE:     | DNA |
| ORGANISM: | Unknown |
| FEATURE:  |
| OTHER INFORMATION: Artificial primer sequence |
| SEQUENCE: |
| gttgtctcc agatccgtg agg |
gtctggcgaatcttcagacag

SEQ ID NO 4
LENGTH: 21
TYPE: DNA
ORGANISM: Unknown
FEATURE: OTHER INFORMATION: Artificial primer sequence
SEQUENCE: 4
gatggcatagtgagaaagag

SEQ ID NO 5
LENGTH: 24
TYPE: DNA
ORGANISM: Unknown
FEATURE: OTHER INFORMATION: Artificial primer sequence
SEQUENCE: 5
cgggtgctgctttcctaaacttgac

SEQ ID NO 6
LENGTH: 25
TYPE: DNA
ORGANISM: Unknown
FEATURE: OTHER INFORMATION: Artificial primer sequence
SEQUENCE: 6
gaatctttccgtgcaagccttcag

SEQ ID NO 7
LENGTH: 23
TYPE: DNA
ORGANISM: Unknown
FEATURE: OTHER INFORMATION: Artificial primer sequence
SEQUENCE: 7
ggtcttagctactgtggccaggg

SEQ ID NO 8
LENGTH: 22
TYPE: DNA
ORGANISM: Unknown
FEATURE: OTHER INFORMATION: Artificial primer sequence
SEQUENCE: 8
gttgctcaacgttgctgcagttccc

SEQ ID NO 9
LENGTH: 20
TYPE: DNA
ORGANISM: Unknown
FEATURE: OTHER INFORMATION: Artificial primer sequence
SEQUENCE: 9
tggtgtaacccgttacggccacc

SEQ ID NO 10
LENGTH: 20
TYPE: DNA
ORGANISM: Unknown
FEATURE:
OTHER INFORMATION: Artificial primer sequence

SEQ ID NO 11
LENGTH: 10
TYPE: DNA
ORGANISM: Unknown
FEATURE:
OTHER INFORMATION: Artificial primer sequence

SEQ ID NO 12
LENGTH: 20
TYPE: DNA
ORGANISM: Unknown
FEATURE:
OTHER INFORMATION: Artificial primer sequence

SEQ ID NO 13
LENGTH: 17
TYPE: DNA
ORGANISM: Unknown
FEATURE:
OTHER INFORMATION: Artificial primer sequence

SEQ ID NO 14
LENGTH: 16
TYPE: DNA
ORGANISM: Neotyphodium uncinatum
FEATURE:
OTHER INFORMATION: Artificial primer sequence

SEQ ID NO 15
LENGTH: 25346
TYPE: DNA
ORGANISM: Neotyphodium uncinatum
FEATURE:
OTHER INFORMATION: Artificial primer sequence

FEATURE:
OTHER INFORMATION: Artificial primer sequence

SEQ ID NO 16
LENGTH: 16
TYPE: DNA
ORGANISM: Neotyphodium uncinatum
FEATURE:
OTHER INFORMATION: Artificial primer sequence
-continued

atactagact tttctttaaa gtaaacccctt tttatatttat aaagcttaaag attaastact 540
tagggaagtc tttttaaatat taatttacgc aggttatagt gttttttctat cttattatagt 600
tagccttaga tsaataagct taactaccc ttagttagta gttttttccg ttaataatttt 660
ttagtttat ataataggtt ttttaagttt atacctatgc aactcttatt tttatcttaat 720
tagtaaatct ataatagact ttttacactct atatttattaat ttaataaatag 780
ttagtaaattt actaatctat cttcctttaa gtatcttact atctctacact 840
aggttacactt actatctccttt ataatctcctt tataaatataa ataatagtt 900
taatattactc cttaataaact tttataaatat tttttttttta aatttattttta aacctgttagta 960
ggtaggacca atctatcctaa ttaataagtaa ttaataaat tttatatcactt acacctgtag 1020
tataataaat actaatatatc tatactttaact tttataaat ttaataaactc aactcatgac 1080
aagtttttaag ttaaatggtt tggggttc taaaagttta acttaaatgtaa 1140
rttagtagctt ttagtagcct aatagtttagt ttttattttta taatagcct 1200
cgtttagataa atcatataact attataacta ttttatatatag taatagcct tataataaat 1260
actataattct cttaataagga cttaatatattg taagcccctc aacctgttagta 1320
aggttagcctt tcctattaaact tcctattttta cttatttagt aatttttaaccctt 1380
aactagtttaa atactattttta ataatctaggt ttttattttta ttttatttttt 1440
aagtttttaag ttaaatagcct aatagtttagt ttttattttta ttttatttttt 1500
atatattctt aatagtttagt ttttattttta ttttatttttt 1560
ttacttaaatct ttagctatgct actactataact ttttattttta ttttatttttt 1620
aggttagcctt tcctattaaact tcctattttta cttatttagt aatttttaaccctt 1680
aggttagcctt tcctattaaact tcctattttta cttatttagt aatttttaaccctt 1740
atatattctt aatagtttagt ttttattttta ttttatttttt 1800
aggttagcctt tcctattaaact tcctattttta cttatttagt aatttttaaccctt 1860
aggttagcctt tcctattaaact tcctattttta cttatttagt aatttttaaccctt 1920
aggttagcctt tcctattaaact tcctattttta cttatttagt aatttttaaccctt 1980
aggttagcctt tcctattaaact tcctattttta cttatttagt aatttttaaccctt 2040
aggttagcctt tcctattaaact tcctattttta cttatttagt aatttttaaccctt 2100
aggttagcctt tcctattaaact tcctattttta cttatttagt aatttttaaccctt 2160
aggttagcctt tcctattaaact tcctattttta cttatttagt aatttttaaccctt 2220
aggttagcctt tcctattaaact tcctattttta cttatttagt aatttttaaccctt 2280
aggttagcctt tcctattaaact tcctattttta cttatttagt aatttttaaccctt 2340
aggttagcctt tcctattaaact tcctattttta cttatttagt aatttttaaccctt 2400
aggttagcctt tcctattaaact tcctattttta cttatttagt aatttttaaccctt 2460
aggttagcctt tcctattaaact tcctattttta cttatttagt aatttttaaccctt 2520
aggttagcctt tcctattaaact tcctattttta cttatttagt aatttttaaccctt 2580
aggttagcctt tcctattaaact tcctattttta cttatttagt aatttttaaccctt 2640
aggttagcctt tcctattaaact tcctattttta cttatttagt aatttttaaccctt 2700
aggttagcctt tcctattaaact tcctattttta cttatttagt aatttttaaccctt 2760
aggttagcctt tcctattaaact tcctattttta cttatttagt aatttttaaccctt 2820
aggttagcctt tcctattaaact tcctattttta cttatttagt aatttttaaccctt 2880
agtcgctaco atcagctacg cgtgttttgt ggccgcccgt atcatacagc acagtacctc 2940

ctctgatttt ggttaaacgt ggttagcacc ttttagggcc ctgagccaga tggcggcaga 3000
cgtttgag ctagctgctg ccatcaaccc ttttaggggcc tggcgtgtgg ctggtacggta 3060
cgggggctag ctagctgtga ggtatgcgca atagctgctg ccatcaaccc cgggggctag 3120
gagaaggtga ctacgctgcg ccaagggaga aacgacatcc catcccttcg tggcatacga 3180
tctgagacgt atgggtggtt ttgcaataac acagctcagt caggtggaac cgggggatt 3240
ggggagaact actgctctag cgctctcata aaagaaaaaa acocactgct acctgtgctc 3300
tagctctgaa ggtggtatcg actcgataaa ccagagaaag ttgcttcaac cggagacgct 3360
tctctcata tgcgcctgct caaacccctc cagatgttgg gcagacagct ttttagggcc 3420
cgagccagcc acagcactgc ccacccagcc acgtgtgtgg ttagagcagta ccagagaaag 3480
aacagcagc aaggctcgctg ccaacccactc aagttttctt acctctgtag ctctgctctg 3540
ctctgctttg ttgctgtggc ggtgttcggt ttggactagc aacccaggct cgggaccatt 3600
cctctccccct tgcagcgtgg tgtttgagaa cctctctgct gcacatgtgc cgcctgctg 3660
ggaccctgct caggtctcag caaacccactc cgggaccatt ctctgctttg 3720
cccgctctct atagagacgc tcctgaccc cgtggtctcg tggcatacga cggagacgct 3780
agctctgct cagctgctag cggagacgct ggtggtatcg actcgataaa aaagaaaaaa 3840
cagctgctag cggagacgct ggtggtatcg actcgataaa aaagaaaaaa 3900
aatagctgtcc cgcagctgct cgcagctgct cgcagctgct cgcagctgct cgcagctgct 3960
actggttgct gccagctgct cgcagctgct cgcagctgct cgcagctgct cgcagctgct 4020
agattttcta cccagagata ctcatctttg tttctggag cggagacgct ggtggtatcg 4080
cagagatac cggagacgct ggtggtatcg actcgataaa aaagaaaaaa aacaactgct 4140
ataagctctg caagctgctag cggagacgct ggtggtatcg actcgataaa aaagaaaaaa 4200
actggttgct gccagctgct cgcagctgct cgcagctgct cgcagctgct cgcagctgct 4260
actgaagctg ccctctgctg agctggttgt ctatagaaag ttgcttcaac cggagacgct 4320
tgcagctgct cgcagctgct cgcagctgct cgcagctgct cgcagctgct cgcagctgct 4380
gctggttgt ccctctgctg agctggttgt ccctctgctg agctggttgt ccctctgctg 4440
tctctctctag cagatgttgg gcagacagct ttttagggcc tggcgtggtt ttttagggcc 4500
cctctctctag cagatgttgg gcagacagct ttttagggcc tggcgtggtt ttttagggcc 4560
ggggagacgc aacagtacgt gcctctctctag cagatgttgg gcagacagct ttttagggcc 4620
tgcagctgct cgcagctgct cgcagctgct cgcagctgct cgcagctgct cgcagctgct 4680
gctgtgtttct ccctctgctg agctggttgt ccctctgctg agctggttgt ccctctgctg 4740
tctctctctag cagatgttgg gcagacagct ttttagggcc tggcgtggtt ttttagggcc 4800
tctctctctag cagatgttgg gcagacagct ttttagggcc tggcgtggtt ttttagggcc 4860
gctggtgtcc ggtggtctgc gcagctgct ggtggtctgc gcagctgct ggtggtctgc 4920
gctgtgctgc ggtggtctgc gcagctgct ggtggtctgc gcagctgct ggtggtctgc 4980
tctctctctag cagatgttgg gcagacagct ttttagggcc tggcgtggtt ttttagggcc 5040
gctggtgtcc ggtggtctgc gcagctgct ggtggtctgc gcagctgct ggtggtctgc 5100
tgcagctgct cgcagctgct cgcagctgct cgcagctgct cgcagctgct cgcagctgct 5160
gctgtgtttct ccctctgctg agctggttgt ccctctgctg agctggttgt ccctctgctg 5220
<table>
<thead>
<tr>
<th>Sequence</th>
<th>Counts</th>
</tr>
</thead>
<tbody>
<tr>
<td>gaccaagtca gctctggttg agaaccctac caaaccctgg caagcgttac tagaccttgtga</td>
<td>7680</td>
</tr>
<tr>
<td>ggccagtctcg cttgctgc acacaaccgg ccctctttt atgtgata tccsactagsa</td>
<td>7740</td>
</tr>
<tr>
<td>aaccgctgot gccatagggct gttggttgggg taaaatcctg gggagatgcttt tccoaatccca</td>
<td>7800</td>
</tr>
<tr>
<td>agatgattagt gctgataaac cttgcggctgt gcaggggactt ttgctgctcc cttacactaac</td>
<td>7860</td>
</tr>
<tr>
<td>ggccagacat atcctgctga cccctcttga atgctgctcc cttctccagag ccacacacttc</td>
<td>7920</td>
</tr>
<tr>
<td>gcggtagcgaa cagctgagcc cggcagcccgg gcacacgctga ctagaggtct</td>
<td>7980</td>
</tr>
<tr>
<td>catctcaagct ccacacacgtg gcaagacttg gacccgtggcc cctgctgctcg</td>
<td>8040</td>
</tr>
<tr>
<td>cggtgtccct tcctggtcct gcctggtac ccgccctggcc tcggcctcgc ccgctggccag</td>
<td>8100</td>
</tr>
<tr>
<td>caagggcttg gctctgctcg atcccttggcc gctgctgccg acggcgaagcc gcggctctcc</td>
<td>8160</td>
</tr>
<tr>
<td>gccaagacat cggactgcttg gcgtgacctc ccccctggtgc cttgggctcag ctatgtggtg</td>
<td>8220</td>
</tr>
<tr>
<td>ccaacccacgcc ccaagcctgctgc cgccggtctc gccacagggc tctggctctcc tggcgggt</td>
<td>8280</td>
</tr>
<tr>
<td>tgaatctgctc tgtgctacag gcggggttct cgaaccaggt gcgctgcttc gctgctgtct</td>
<td>8340</td>
</tr>
<tr>
<td>gatagctggc atcccaacct ccaagactgg cttgagccttc gcggctgctgc gcggctgctgc</td>
<td>8400</td>
</tr>
<tr>
<td>ttcctcagc cttccactct gcagctgctac ctgctgctgc gcctgctgctgc gcggctgctgg</td>
<td>8460</td>
</tr>
<tr>
<td>tgcagactgca ctaatccttc ggtctgctgc gcggccaggt gcgctgctgg cggctgctgc</td>
<td>8520</td>
</tr>
<tr>
<td>ccaactgaaact gcagcctgctgc cctcctgtgg gcggctgctgc gcggctgctgc gcggctgctgc</td>
<td>8580</td>
</tr>
<tr>
<td>actggccct gcggctgctgc gcggctgctgc gcggctgctgc gcggctgctgc gcggctgctgc</td>
<td>8640</td>
</tr>
<tr>
<td>tgcggtctgc gctgctgctgc gcggctgctgc gcggctgctgc gcggctgctgc gcggctgctgc</td>
<td>8700</td>
</tr>
<tr>
<td>ggtctgctgc gctgctgctgc gcggctgctgc gcggctgctgc gcggctgctgc gcggctgctgc</td>
<td>8760</td>
</tr>
<tr>
<td>ttttttttgtg acgctatgtt ttcggtcctt ttcgctgtt gctctcctc gctctcctc gctctcctc</td>
<td>8820</td>
</tr>
<tr>
<td>ttttttttgtg acgctatgtt ttcggtcctt ttcgctgtt gctctcctc gctctcctc gctctcctc</td>
<td>8880</td>
</tr>
<tr>
<td>cagctgaact ttttgggtgc gactgtgctc cttgacact gttgactacgc gggccttcttc</td>
<td>8940</td>
</tr>
<tr>
<td>ggtctgctgc gctgctgctgc gcggctgctgc gcggctgctgc gcggctgctgc gcggctgctgc</td>
<td>9000</td>
</tr>
<tr>
<td>actgctgctgc cccctggtgc cttgagccttc gcggctgctgc gcggctgctgc gcggctgctgc</td>
<td>9060</td>
</tr>
<tr>
<td>gttgagacac ccggagatgt gttgagccttc gcggctgctgc gcggctgctgc gcggctgctgc</td>
<td>9120</td>
</tr>
<tr>
<td>cggtgctgctgc gctgctgctgc gcggctgctgc gcggctgctgc gcggctgctgc gcggctgctgc</td>
<td>9180</td>
</tr>
<tr>
<td>ccgggctgctgc gctgctgctgc gcggctgctgc gcggctgctgc gcggctgctgc gcggctgctgc</td>
<td>9240</td>
</tr>
<tr>
<td>tggagcctgc gcggctgctgc gcggctgctgc gcggctgctgc gcggctgctgc gcggctgctgc</td>
<td>9300</td>
</tr>
<tr>
<td>gacatcatact cttctgtgt tttctgctgc gcggctgctgc gcggctgctgc gcggctgctgc</td>
<td>9360</td>
</tr>
<tr>
<td>actgctgctgc gctgctgctgc gcggctgctgc gcggctgctgc gcggctgctgc gcggctgctgc</td>
<td>9420</td>
</tr>
<tr>
<td>tttctgtgt tttctgctgc gcggctgctgc gcggctgctgc gcggctgctgc gcggctgctgc</td>
<td>9480</td>
</tr>
<tr>
<td>ggtctgctgc gctgctgctgc gcggctgctgc gcggctgctgc gcggctgctgc gcggctgctgc</td>
<td>9540</td>
</tr>
<tr>
<td>ccgggctgctgc gctgctgctgc gcggctgctgc gcggctgctgc gcggctgctgc gcggctgctgc</td>
<td>9600</td>
</tr>
<tr>
<td>ccgggctgctgc gctgctgctgc gcggctgctgc gcggctgctgc gcggctgctgc gcggctgctgc</td>
<td>9660</td>
</tr>
<tr>
<td>ccgggctgctgc gctgctgctgc gcggctgctgc gcggctgctgc gcggctgctgc gcggctgctgc</td>
<td>9720</td>
</tr>
<tr>
<td>ccgggctgctgc gctgctgctgc gcggctgctgc gcggctgctgc gcggctgctgc gcggctgctgc</td>
<td>9780</td>
</tr>
<tr>
<td>ccgggctgctgc gctgctgctgc gcggctgctgc gcggctgctgc gcggctgctgc gcggctgctgc</td>
<td>9840</td>
</tr>
<tr>
<td>ccgggctgctgc gctgctgctgc gcggctgctgc gcggctgctgc gcggctgctgc gcggctgctgc</td>
<td>9900</td>
</tr>
<tr>
<td>ccgggctgctgc gctgctgctgc gcggctgctgc gcggctgctgc gcggctgctgc gcggctgctgc</td>
<td>9960</td>
</tr>
</tbody>
</table>
atattgcatct ttcctgacct taacctcaacgg ttctcaaeqgt gtatgatctc attagacctt 14760
cgggatatct ttcctgccag acgtacacaa tccttcttgag ctacgttttcgc cactgctattg 14820
agaaagtggtct gcacatattc tgcataatata agcgcagtcgc gtctgcggac gactcagc 14880
acgagcaagct gctgcacacac aacccatcaac gcgcacccgc ccaatcaacaa tctcagctgta 14940
tttcttatacgc ttgctcagatg aacacccaaagg gaagacgggt ggttctgtaa gaacacgaca 15000
acacccaggg gcctggagatg ctctgcgttttc cactctctcag acacccacaa tccacctgc 15060
tgtctccacc gcgcacagct gcctacaaact gacgctgaca ttgtgcgtgtcct 15120
gagactaaggc cccctgagaa gttgccagat aatagctccttt ttctggagac gactcctcct 15180
actagagacagtggacgtgcagttgctcctgactgcccaggtcacccgactcaccctct 15240
ggagcgtgagg gcctggccagctgcacagctgcctctgctgcagatccagacgc 15300
cagacatgacg ggcgcagctgcacagctgcctctgctgcagatccagacgc 15360
gttctctcct ccacatgaaagtctctcctctgtgcacagctgcctctgctgcagatccagacgc 15420
ccacatgacgc cagacatgacg ggcgcagctgcctctgctgcagatccagacgc 15480
gagacgcagga ggcgcagctgcctctgctgcagatccagacgc 15540
tcctgctcctg ccctgcctcctgctgcagatccagacgc 15600
tttgtggtct cagacgcagga ggcgcagctgcctctgctgcagatccagacgc 15660
gtctctcct cccctgctgcatgctgcagatccagacgc 15720
gttctgtggct atgtgcttttc tgcgtgtata gacgctaaatgctctctcctctgtgcagatccagacgc 15780
acacccaggtct ttcctgctgcagatccagacgc 15840
taggtacacaac aacgcttttcctgtgcagatccagacgc 15900
agtctgctctc gccatgctctc gcctgcctctgctgcagatccagacgc 15960
tccctgtactaagagcctgcctctgctgcagatccagacgc 16020
tagctgctctc gcctgcctctgctgcagatccagacgc 16080
acacccaggtct ttcctgctgcagatccagacgc 16140
gcgcctgttcttg cctctgctgctctgctgcagatccagacgc 16200
agacagcagctgctgcctctgctgcagatccagacgc 16260
gctgcctgctctgctgcagatccagacgc 16320
gctgcctgctaagagcctgcctctgctgcagatccagacgc 16380
 tgctgtgctgcxgcctgcctctgctgcagatccagacgc 16440
attgagcactgctgcctctgctgcagatccagacgc 16500
aatctcgcctgcctgcctctgctgcagatccagacgc 16560
 ggctgcctgcctgcctctgctgcagatccagacgc 16620
 ggctgcctgcctgcctctgctgcagatccagacgc 16680
 cactgcctgcctgcctctgctgcagatccagacgc 16740
 ccctgcctctgctgcagatccagacgc 16800
 acctgcctctgctgcagatccagacgc 16860
 ttgctgctctgctgcagatccagacgc 16920
 gcgcctgcctctgctgcagatccagacgc 16980
 gcgcctgcctctgctgcagatccagacgc 17040
 gcgcctgcctctgctgcagatccagacgc 17100
aactgcaggg aagtagacgc ccgctttggtgc gacgacgcc caactacgctt ctcacagatg 19500
atgcgaacct ttcagacagc gtcctacccgc ctcagacaca taacccccgta atggggttaa 19560
gtcgtctccat atcacagtcg tttttaaatag tagggagact cagagacccca ctcacacntg 19620
gggctccagc gattgaacgc ttcagagcgct aatgccacgg ccactatggg toccatactg 19680
acegctgcc ttcagacgccc atcacaagga ccagagatgg ggggtacacg ggcgcaagag 19740
ggtagacagt ccgaggggagc atcgccagccc atcagacagc ccttggcagct gaggatcccg 19800
gccgtccagct tgggacccct gtgcagttct cttgtctgtg gtcgctagatg ggcgctgcaag 19860
cgccacccg ccgtacccccg ttgggttcttc gcctgccccg gacgacgacgg 19920
tgtctgctga ctcagcgcgc cgtgacaaaa tgcgacactt cttcagacccc ggcgacgcac 19980
tgccgctgggt gcgcgcttgcc gcgctgctcg cgcgctgggt gcgcgcttgcc gcgctgctgggt 20040
tgtctataac gcgtacagca gcctacgcag cccgctgagcg tgggagcgcgc 20100
agccgccgca gcacgggagc aagggacgca tcgagccctc attatcagcag aagggacgca 20160
cgcagatgc gcgctgctgc ggacagccac gcagccagatgc gcgctgctgc ggacagccac 20220
tacgctgctgc gcgctgctgc gcgctgctgc gcgctgctgc gcgctgctgc gcgctgctgc 20280
ccttacgagc ccgcttacgt gcgggcttgt gcgggcttgt gcgggcttgt gcgggcttgt 20340
gccagccgc gcagctgcgc gcgtggtgct gcgtggtgct gcgtggtgct gcgtggtgct 20400
cgcgctgct gcgtggtgct gcgtggtgct gcgtggtgct gcgtggtgct gcgtggtgct 20460
gtcggtacgcc gcgtggtgct gcgtggtgct gcgtggtgct gcgtggtgct gcgtggtgct 20520
gcgctgctgc gcgctgccgc gcgctgccgc gcgctgccgc gcgctgccgc gcgctgccgc 20580
cgcgctgccgc gcgctgccgc gcgctgccgc gcgctgccgc gcgctgccgc gcgctgccgc 20640
cagtcggctgc gcgctgccgc gcgctgccgc gcgctgccgc gcgctgccgc gcgctgccgc 20700
gctgctgctgc gcgctgccgc gcgctgccgc gcgctgccgc gcgctgccgc gcgctgccgc 20760
tgtctataac gcgctgccgc gcgctgccgc gcgctgccgc gcgctgccgc gcgctgccgc 20820
acctgcaact ctcgagcctt atgcagcttg ctcgggcaag ctcgggcaag ctcgggcaag 20880
ggtcctgtca acaccccccg atactacttc gcagctgagc gcgctacggt gcgtgcagct 20940
cgtcgtacgg atatctgagc gcgaacccag tttcgcttta cctcgtgcag ggggtctgta 21000
ttccgagagc atacgagcgc gcgttcgagt gcgttcgagt gcgttcgagt gcgttcgagt 21060
agtatagcgt ctcggtcaag ctcggtcaag ctcggtcaag ctcggtcaag ctcggtcaag 21120
aggggttagc gtagggtagc aatctctctc gtacgctgcc gcgctgccgc gcgctgccgc 21180
gggctagga attatcgcaaca tcagtaatag tttcgtggtt gcggctgctgc gcgttcggtct 21240
cgacgctactg acatcacaag gtcgctcgcg gcggctggttc gtcgctggttc 21300
cattgtctg gcggcttata cttttagtaagg ccatcctctc ctcctctctc ctcctctctc 21360
atctaattatatcggtgc gtcctggttc gtcctggttc gtcctggttc gtcctggttc gtcctggttc 21420
ggctgataac gcacccgtcct gtcgatcggc gcgtgtggtc gcgtgtggtc gcgtgtggtc 21480
gcatattatatcagctggtgacctttc gcacgctgc gcgtgtggtc gcgtgtggtc gcgtgtggtc 21540
cggccccttc cactacgcc ccagctgcgc gttcctgcgc gcgtgtggtc gcgtgtggtc gcgtgtggtc 21600
catgctctgc gacagctggt gcgtgtggtc gcgtgtggtc gcgtgtggtc gcgtgtggtc gcgtgtggtc 21660
agatcgactgc gcgctgctgc gcgctgctgc gcgctgctgc gcgctgctgc gcgctgctgc gcgctgctgc 21720
gggtggtagc gcgctgctgc gcgctgctgc gcgctgctgc gcgctgctgc gcgctgctgc gcgctgctgc 21780
cttctctact gcggctctgc gcgcggcctc atccacacgc ctcctctctc gcgtggtgct gcgtggtgct 21840
agastatcgg cgatctgctt accaccctct ttcgcgagcc acctttatgt caatgtctaatt 21900
gcgcctgcg caccacacag gtcctgcgcc tgcgggaggg cgccgggaggt gcaacagatgg 21960
ggaattgat cgctgagcg gttagaggcc caaagatgtct caaaggaaga gtcgctggc 22020
cgccgtaacga ccacgaagga acagccgagct ggtccattct gaggataaat 22080
cgaacagag aggacagact gcgggatcag agttgaggct tcgctggcgg aacggtatat 22140
acagacacaag tgggaggagca aataaaaccct ctgatgtagct cattggctag aagacacacoa 22200
tggacagccc gtatattaaggg ctggtgctgac gttgagatta catattctgtg gcacatcgcag 22260
acagagaagt cggggtcag ctcgctgaga tgtacccctga actgagcccaac gctatcgccc 22320
cgctcagcca acctactgtgg gctcctcttc tggctacatgc gcaacagatgc ttcgacagccc 22380
actctatagcg cggccgaggt tcgctagcct tacattttct ctggccggccct tcgtagacgct 22440
tccgctatctg ttcgctgact gcgcgtcctca accctgggac caactctttg gcatctgcgct 22500
gggagctcaaa actcacaactt aagttccctg tcggagtcag agacagtaccct 22560
ccctgtccga tggccgaatgc gttcaggttg cgggacctggt ctcgacagccct ttcctcacaat 22620
gcttttggtct aeeagagttg gagaattccag gttcgccgtg ctcgagccgac agcgtaccttc 22680
agasaacggc ctaggaatagtg ctgacagttc ggcgccaggg aacagatccg 22740
tccggaaac ccagattcgctgt ggctccagcg agggacaccgc aatctctgat 22800
tgctgcgccc accaccctgg gccgacccac acaagatctg acaacggctc cttgcttgatg 22860
ccagatacgc caaacacgga attgaaacgg gcacatcgtyg tcgagggcata taagccagga 22920
ttcgattttg tggattctcg accatcctcag aacgtcttccca gcatcatgcttg ggcgacagctg 22980
aatcgtttag cttgctgatt catctggcag accgattgac accgacacag aagtgattgg 23040
tcagattagg tagagatgcc accagcactt gtgtaaaggc gccacgctca gtaaagccac 23100
acgctgagcc ttcctattgtatatcctgaccgc gactcatatt acaccccttc accacccctg 23160
gttgattttag gttgtgagcgaccacatgcgtgttcgggctggcgttaaattctgcgatg 23220
caagcttaatt gcagctctcttg ttttattatatatgtagttc tttgtgattt ggaagctgtcgt 23280
ggcgagcggac gctattgatt gctgctgctg gaaggctttg aaggtgcgctctg 23340
agatatataa taccctctct tcttcgacgcc ctcacccccc cttgctgagc ctccttataa 23400
aacaatccca cttcgagac aacaccctgtg aacccctcttt ctttacttaa tccatcatgtg 23460
ccgctgtttc ttccttccttc ctaaggctctt ctcaggggag cagcaamgtc taccacgcc 23520
acggtatatg tcgctgccgc caaagaggct acggcgtggt gcaagctgcttg gcaaaatggc 23580
agcaacaggt tggcagactta cttcgctgcc ctgctgctgg gggtgagctatg gcaaaagctc 23640
acggagacgc gaaagtcgctct gtttctctct ggagggccga ggaagctctgt gtaaaccacgc 23700
cggccagagcc atagtgtctag ggtgggcggg ccagccgttact cccttttgtc cccgagtttc atccgagata 23760
tcgcgctttg gatttcttg tgtgctgccg ccgctgtgtc gtagctttgg cagcgctctg ctcgctgctg 23820
ccgcttctca ctcgacgcacg cccgctctct ccaatttttc cccgagtttc atccgagata 23880
tgattggcatt cgcagcagcct ggaagctttg aagttccctg acggtcttgc acggtcttgc 23940
acagcgccgcttc gttgccccgc ggtgacctgg ctgctggccac ggcgacgcgc 24000
agagttgggt ntccgcrtcct gccgagcctgc gcagcagcttc gtagcttcct gcagctgcgtg 24060
ctcattgattt ggtgtcattc aagagcgtttcc aagatggcgtt gccgacttcct gcagctgcgtg 24120
actttgaccg ttcagccggc aacaaattctca cccgagctgg ctccacaccac 24180
<210> SEQ ID NO 16
<211> LENGTH: 16397
<212> TYPE: DNA
<213> ORGANISM: Neotyphodium uncinatum

<400> SEQUENCE: 16

tataagcttt tacataatt tttataaagta aactgtcag tagctattc 60
taaagcta caggaggtta cttatcttat aattttatat aaaaatagct ctcttcttg 120
tagggtctg ttaattctt ataataagaa actaactact gcatgttttc gttggtctg 180
acaagacotg tggcctcgg tccaaaccc ccctctaagtg ccctggtttgc cctagactg 240
taaacaggca tctctccctt cttgatcctt actattctt actattctt actattctt 300
tatatcgtt cagctactatt ttaattttt ttagatgctt tttgatgctt 360
tatatcgt ctggttattc cgctcactc tttcaaagta aacactaacgt ttaaagcagtt 420
cgggactg tttgttattg tggcctcgg tctcact ctcttcctctg ccctggtttgc 480
agaagactg tttggttttc ggctcctcag ggtactacac gctcactcg ctcttcctc 540
gtgggttttc ggctcctcag ggtactacac gctcactcg ctcttcctc 600
cttctctt gcctcactc tttgcttattc gcctcactc tttgcttattc 660
ccctcctatt taaaagctt ccctcctatt taaaagctt ccctcctatt 720
cycaatcaca ttaacagcctga acagctcagc cggcttttcgtc 780
acctggttac ccacctccac ccctccctctc ctctactacac cgctcactcg 840
tttctcttc ggtactacac gctcactcg ctcttcctc 900
tttcctcctt gggcagctc ctctcactacac cgctcactcg ctcttcctc 960
-continued

ggtttccaac tggttacaaa cactaagct aggcggcgct gcaattgagca attacgaga 3360
cgtgacgct atagacgcca gatatctcag atacagatcc atacagatgc cgcggcaaat 3420
gccctcgatt atgcaaaaag ccagaaagcgc cagaaagcgc tccaaagcttc gcctgtaaatc 3480
cgggtctcct cttggaacac tttgaagcaco tctgctacaa tctgcagtcg tctttcatac 3540
agtggctaca gaggttctga ggacacagaa cagacaagct tgaacaatttg gcagaggaagt 3600
agctgccaca ggtgttccga ctgttctgac gctgtacacac caataattgga gcgcttgccag 3660
tctggetto atttatatgt ccgattactt cttcgyttttc gtaaaactoa actgyccagc 3720
cttcatcctg tcggcccaccc gatggacttt cttccgact tttcccatct cccgagaagt 3780
atctgcaaca atttccagtc gcgcacagca agcgacggct acagaaagca tctggaaca 3840
gttcgacggt gcggactgat ccgcaagccg ccgcaagccg gttctgcttc tctgagcct 3900
cctccttttg cttccgactcc aagcgagctc tttgaagctg gcggtgctgc aagcgagctc 3960
tccatgagat cactctttgctt ttggcaagctt ctccaatcctt cccacttttct atgttctctct 4020
aagcgtgagaa gcggcgccgag gggggaagcc ttcagactgt cgygggagcc ctgggaaga 4080
agacagatgac gggctgctgat ccgactccttg gcagcactgtg gcagcgagctg tctggtctgc 4140
tggaggagy aggctggagt gatcagctcc tgggcatttg cggccagcaag aatctctctc 4200
aactcaagt gcaagcggctgc cttccagcttg gcggcagctc gcagcgagctg tctggtctgc 4260
atcttgcttg ccgactctgc cttccagcttg gcggcagctc gcagcgagctg tctggtctgc 4320
atctctgttg caacagagatt tggctttgtgg aagcgagctc ttcagactgt ccgactccttg 4380
agctctgtgct gcgcacagca cgcagagctc cgcagagctc cgcagagctc cgcagagctc 4440
gggccgggag gcgggcaatgg ttaaggagtt ggaactagtcc gctggggagc aatctctctc 4500
cgtgctgctgc cggcttctcc ttcagagtgtgc cttccagctct gcggcagctc gcagcgagctg 4560
gactgtcatacg ctgctgctgc ctacagagct ttcagactgt ccgactccttg gcagcgagctg 4620
agctctgtgct gcgcacagca cgcagagctc cgcagagctc cgcagagctc cgcagagctc 4680
gggccgggag gcgggcaatgg ttaaggagtt ggaactagtcc gctggggagc aatctctctc 4740
ccatgagagt gtcggagctgc cggccagcaag aatctctctc atacagagct ttcagactgt 4800
tggaggagy aggctggagt gatcagctcc tgggcatttg cggccagcaag aatctctctc 4860
agctcttgtt ccagagagatt tggctttgtgg aagcgagctc ttcagactgt ccgactccttg 4920
ccatgagagt gtcggagctgc cggccagcaag aatctctctc atacagagct ttcagactgt 4980
ttcagactgt ccagagagatt tggctttgtgg aagcgagctc ttcagactgt ccgactccttg 5040
ccatgagagt gtcggagctgc cggccagcaag aatctctctc atacagagct ttcagactgt 5100
ccatgagagt gtcggagctgc cggccagcaag aatctctctc atacagagct ttcagactgt 5160
ccatgagagt gtcggagctgc cggccagcaag aatctctctc atacagagct ttcagactgt 5220
ccatgagagt gtcggagctgc cggccagcaag aatctctctc atacagagct ttcagactgt 5280
ccatgagagt gtcggagctgc cggccagcaag aatctctctc atacagagct ttcagactgt 5340
ccatgagagt gtcggagctgc cggccagcaag aatctctctc atacagagct ttcagactgt 5400
ccatgagagt gtcggagctgc cggccagcaag aatctctctc atacagagct ttcagactgt 5460
ccatgagagt gtcggagctgc cggccagcaag aatctctctc atacagagct ttcagactgt 5520
ccatgagagt gtcggagctgc cggccagcaag aatctctctc atacagagct ttcagactgt 5580
ccatgagagt gtcggagctgc cggccagcaag aatctctctc atacagagct ttcagactgt 5640
ccatgagagt gtcggagctgc cggccagcaag aatctctctc atacagagct ttcagactgt 5700
-continued

tgtgaacaac gaccaaggcca acccaagagac caatgtcctt gtttcacattc accagcaacg 8100
tcagatcata cacttgctct coccaaggtc agtctggccc accagcatag gctcaagcgg 8160
tgccatgctg tgctgagact tacggccgct ctaaaagttg cccaactaact gccctttgtg 8220
cogagttctt gagaagccag aaaaatgcca gttgcaagct gacatatgt ctcaagcctt 8280
gggggagaa atcttggccg tgtgggggccg agccotgcag gccgagaccgg tagccttgg 8340
cagctcagc aatgctcaag tggcccaggta ttctcagctac gggagccgca agctgcctcc 8400
ggagacgctc atattagcctc ctttcttcc caagccgac atotcggctg tgggacaccc 8460
gagccagcgg atctcgggccc acatcctccgg caacattggag gccaatgata ttctcagcgt 8520
catgatctgct cagctgttcc gcccttactct ttctccttcct ctaaaagtgt catcctgccca 8580
cctctctgctgc cagctggtcc aagcagccgc aggctggtta tagcctgtgt gcctctcgag 8640
cggcctaaag ctaaaaagtt gggtttccgc aggcttttcc tttttgaggcc gataagcacc ctaaaacctaa 8700
catggaac agtctctcgtgt ttgctctcct ctcctcctac ccacccctccc cctctttggat 8760
tgcaccgcc gcccttctggct gcctgatgtcc tttttgccc gcctgtcacc ctaaaacctaa 8820
attctcattag tagctcttacag cagcttttact ccotcctcgc atctgctgctt tttttgccc 8880
gcgtcagcc gcctgcagttg acagcacaac gcgctttccgc cctctctccggt atcttgcctt 8940
tctgagacct cctctccgctg gaccccaaca tctctgggag gaccccctccg ccacccctccc 9000
aagcactcgaacc aaaaatgcca gaaaggttgcctt tcagacaatt caattaac 9060
atagtgccgc gggagtcgaa ttcctggacag tcagctggtgc gctcagtcgt tggccatcag 9120
eggcagctc tccatgttctc ccaacacggag gttggtgcctt ctctgttggcc gtaaacagtct 9180
gotagttact ccctccacaaa cggcggcgcc gccgctggtg gctcagtcgg gccgctggtg 9240
cggcgagcc gcggcggagcc agggagccgc gcctctggcc gcgtctcctag cggcgcacgg 9300
tcgcggggc agccttcggc tgtgcacgctg tgtggggccg gaccccctccc 9360
tacacaacgc ccagctttccgct gtgtgtctgc ctcattgctc gcgctttccg ctaccggtttc 9420
aceaagatcc agaggtcgag tgcgatgtc gccacccccc atggggtgc cagaattgta 9480
cgcggctcg ccaggaggcgc atcattgcgt cgggggctgt gcgtctgggg gtggctgggg 9540
agagccgcgc ccagacacccgc gcagcactttgc gcgtctgctc gcgtctgctg gcgtctgctg 9600
agctcctgca cccacccgctg atttcctcatg cgcggctggt gcgtctgctg gcgtctgctg 9660
gttgggttct gcgggagccgc gcgcctcctgc gagcctcagg gcgcttttccg gcgtctgctg 9720
acaacaagct atggggtggc cgcataagcg cgcgatggcgc gcgcttttccg gcgcttttccg 9780
cgggtcttcct ctcgagccgc ccacccctccc ttagagctgc gcgggctggg ggcgggtggg 9840
rtaagcctcg gcggggggtg ccaactgtag ctggggtgcct ggcgggagcc gcgtctgctg 9900
caaacggcgc cagagctttgc tttgtgctgc tgtgtcttgc ctcacaagttc aactgcggt 9960
gagtcggcc aatcaccgag cggacaggcc gcaccccccac gctgcgggtcg atttcctcatg 10020
gagtgcgaaa gatcataagcg gcgttttggc gctctttttgc cgcgctttttgc cgcgctttttgc 10080
tacgtggaa cagcttttgc gcggttgcct gtcgcttttc cgcgctttttgc cgcgctttttgc 10140
agaagctttgc ctcacaacgc gcggttgtgc gcgtctgctg gcgtctgctg gcgtctgctg 10200
ctgggtcttct ccaggttcag cgcggtgtcg gcgtctgctg gcgtctgctg gcgtctgctg 10260
gcagcttgtgc tgtgggttcag ccggtgtcgg gcgtctgctg gcgtctgctg gcgtctgctg 10320
gcgcgggtttg gcgtctgctg gcgtctgctg gcgtctgctg gcgtctgctg gcgtctgctg 10380
gcgttggcgt gcgcgggtttg gcgtctgctg gcgtctgctg gcgtctgctg gcgtctgctg 10440
ctgccctgcc tgtttaagac ggccgggtga gtcctgggga gtaaagcagt gaggycgatt 10500
ggatatgcac gcagaaaaacg aaggasgattt ccaactegaa ccaatagaac gcaaactattt 10560
tgaggagcct tacattaatg ccagtaattg tccgattcgct tgaatttctcg 10620
aawgctatgg tegctgtaactg catagatgtaa cccgaccaacag gtcgcaaatg acgttgctaa 10680
taaatcgtt atcagcaacta gacagcctaatgattcgact agtcgcatg ggtcaattaaa 10740
tgagaattcg aagcgcttggg tctaatcctcg ggttaastgct acacacactaatgcagata 10800
tagttatatcctataatag tcagaatagtt aacatttgat aacttttagtgagcttattaa 10860
ataactgattt aataaatgata attaatcctt aattattagtt gttttaattaa tgattactaa 10920
ataatatttt ataataatcct ttaaatcactct actaatatcct aaataatatttg aattatatattt 10980
cttgagtttat attataagcgt aatctattaaa gattttataaa acaacgaggtg ttcacaaaag 11040
gctttaaat aatataatttcg aatgttcctt gtttatttgtygttactota ggcotctgcgt 11100
agcactataa atataagggtt taatttagtgo gaatctcctgta aacccaggtgaa atctgtacac 11160
acgatgcaac gtttaagacg aacatattcg aatctgacgcc tggcaccggagc 11220
cgacagcgcc gggcgccgg ggtgtctcttc ttttggtgtc ttcctttctgt cgaaagcgcc 11280
gacatcgata atcataacttc aacaacattc gccacaaacct ggggtcaata cataagtcag 11340
ccccttttggtt cccatagggg ccctttatcttg tcccatagcgc gcctctttgccgct cctctgtcgg 11400
gacacacgac cttcaccacag ttgacacgctt gcacacttaa cagaacttaa tggatctgcg 11460
coaatctcaac aacagcgtgaa tggagtctgtct gcacagtggtc gacattgcac gocatgcctttatacttatt 11520
cctctctctct ctggtcttctct gggtggtct gctagccac ccctttcctttgtgctc 11580
cctgctttca ctgcttttggg aatattttaca atcattgact ggcctttaagc agcctcttctc 11640
cgacagctgc ccaggttatg aacgctgcctg ggtgtaacacc ggccaaact cagtcctcaca 11700
agaacagcgg gaattttggc ttcctttgtcgc gtccttttgcct acgcctaccc cccctttttactg 11760
gtctgctct gttcactata cacccttttt gaatttgagg gagaactcagaga caactaaattt 11820
tacggatctc ctgcttttcttg gaggcttgca attgcaactag aatggggcgc 11880
atatcagac gtcggtcagcc ggcactcaac aacagacggtg agggttgggg tacaagcagc 11940
agattcttgc agacccctgcag ggcccgactcg gtcgctttgctt gcagacatttct gagcga 12000
tccgggttcg aacggagctac gacgctgcttg ggcttgctgcg aagggagtctg 12060
cgagaatcg cacagtgcagc gcccgctagcg gtcgcttttc gcctttgtcgtgtc 12120
ggcttcttcc agtgccagcg cgcctgctgt gaggcaactg aagctcgctc gcggcgcg 12180
cgataaagtctc gggctgctgct aagctgctcgc gggcctctcg gttccttactac 12240
agcatgtcgc gggctgctgct aagctgctcgc gtgctgctgct acacgtgcgc 12300
tttctttctcg gcacgtgctgct ggcctgctgct gtggctgctgct ctcttattctttct 12360
cgggaggctgc gtcggtctcg cgttttttcga ttccttttttgccttttctctgc 12420
tttatttcgc ggctgctgctgct ggctgctgctgct ctcttattctttttgccttttctctgc 12480
tggtgcgcgg ttccttttttgccttttctctgc gggtgtcgcct ctctttttttgccttttctctgc 12540
attcatgtg cagcaagctgt gccacctgct ggcctgctgct gcacgtgcgc 12600
ggcatagcct gcacgtgcgc gtcggtctcg cgttttttcga ttccttttttgccttttctctgc 12660
gtgggttcg cagcgtgctgct tggattttgcagcagtcctcgc gccacgtgcgc 12720
agggggctgg ttaatttcgc gcaataacttc ggctttgcgc tggcttgtacgtg 12780
gaagccctct tgtttcccgtgt atcaatactgc agcgtgcttt ctccccctcaca cccagggtc 15240
tctgcagaggt cacactaatc ctatccagcac acgcgaatacg tttgcttcacc ccaagggcct 15300
cagcggtcgg tacgctctgt gccagagctc cagatcagctgg tggcagaaact cttccagttg 15360
gggtggaaaa cggggaatct gcagacattatat tccagagccac caaaagtcgcag gcattctctc 15420
tggyggcagg aagaatcacta aagctactaca gcgcctagctg gaaagttgtatt tttctggcag 15480
gcaccctctgc cgctgctcaag gcagctatgaa caggctgttc caaggctaca 15540
atagggctga aatcctccag cgggaagggg ggcggttccttc acctccagcc gcggctgtac 15600
tcgggttggc cccgtgcaggtt ctcgctagtc atgatgcgcg tagatgcgca taccycgcag 15660
aatgggttggt gctgcgtttgt gccagctctc caaagcgaag cgaccccgat cttccgcgcag 15720
ggcgcctatgt gctccgcggtgc gctagagaggg aggctgttcaata tccocccagg tctgagccg 15780
gggtgctctgcgtatccgaa cggagagcgtg ggcgctcggt ggcagcttga caagacagcgc 15840
cggagaagtc caagctgtttt cggcacctac cactttgaoca cttccacagcc gcggctgagg 15900
gataatcttc acggccacag cgtctctcac gggccctcgac acacaccgaa ggcgttcccag 15960
tggaaaagtc agttgtgcat gtttggagc caaagctatac atggtacgat gcaccatatca 16020
cgtgatgcttc ggtaggaaaag tggcagcagc ggcagcttgcg gggagttgagc gacctccag 16080
acagaagac gcggttgcgg agacaggtgac cgggctgaaat ggaacattgcg tagtaggtgcc tagatatcag 16140
tttgctttccttttgctgtga cgcgtaataa atcaactgcgcc ttgtcttctaa taactacagtc 16200
aacggcactaac aacccggtat cgcctgcttc tggcagcatt ccaacacagc agcgocacata 16260
gaggcgcaca tttacggtct cactttagcgc acatccgctgc gacgactcag tcattacacg 16320
ttccataccttg tagctgtggt atattgccttt ttactataagc tatactgcgg atataatgac 16380
ttattacagc cgggcgc 16397

<210> SEQ ID NO: 17
<211> LENGTH: 7478
<212> TYPE: DNA
<213> ORGANISM: Neotyphodium uncinatum
<400> SEQUENCE: 17

gacgyaaggygctgctaticcagca gccgagaagcg ccgggggtgt ttaacaggga tttggcagat 60
aagggctgtc aagagaagctc caagataatc taataaaccg cgggtttgtat gctoccaaac 120
gctcacaag agctttttttg accttctctta caaggcattaa tcggcagctc ctctgcttctc 180
tgctcaatgta gcgggctggt ccagtgatatgctt ccagccttaca cgggccccgc atcaaaatgc 240
gaacctgtaa aaggagcccc taacatacgg tctttgctgc cccgacgagg acggattttgg 300
cctcctaaacct gacgatcttg ggccatgctga caagctgctca aagacttgctgc acatttacctg 360
gagocctctc ttagctgcgag ctaaatcctttttat cgggctggaa ttagcagttgc 420
acgaagatacg aggctgtgagcc aagatatttc aacocgtgatt acgctcttca aaccgaatct 480
acagcggcaca cgacagcatt ccgggagctc atcaacaccca gaggacgacag gggggcaggg 540
ggggcaagcc tagtatcgcg aagtttgaagt cggcaattct cttgctcagc tctgctcagat 600
atgcgcaaga gggccctcct ctcttacagc atggaggagc gagatcactta cgggtaagg 660
gtgctctctc cagcttacgta caacagtttc ccggcattcc gcgctctgcg caccggagac 720
gtaaatatcc tttacagggc caaattttc aagggagcc gagatcctcttt ttaaatattgc 780
aagggccagc gccagcttctt catggagggag cagcaaggct cctcaccagg accagttcaca 840
-continued

caccgcatac agagactcat toccagtgcc cttaagtaca agggctgaga ggcgttaaaa 900
aaaatgctgc ctcagcgcct ttgttgtgtg acacataaat aggattttgg ggattttgct gaaagagtgt 960
gtcacgtcag acgcctcgat gaaggtattg ggagtggctct ttgagtctttg ctaaagctat 1020
ttagagggga atcagctcaaa taacttaatta ggtttcgtgaa gttgttgtggt ttctgtgtatt 1080
gccatgtgta aacctggttgaa ctttttcacc ctcgacacaga aagaattggc 1140
ccttcattgac cattctggct cttcctgctt ctaaactctc gcacagagctc 1200
cctgggttatt gcaactgtcag ggtgtactata gtttgtaaag aacgccctcct gtcgcaatcc 1260
cgattattc taactaggtt tgccgtttgg tcgggtcaagt aaacaagcct cagagtacgt 1320
aatggcggcag cacgggcctac ctttccccctt cttaagcttct gatgtctgct gcgtgctgaa 1380
cagcctctct cttgagcctc atggcggccc acacccgata acgtgtctct tttttgatt 1440
attatatgcc aagactcaag cagacaccac aocacccac ccagcccttg cccctctgctc 1500
tttagtctaca tctaaagctc ttgacccatt ctttgtgtctgt cttcgttata gttggagat 1560
gtcaacatcc gctgacagcct ttggagccag gttggtccag ttctaaaac ggctgagcct 1620
cagaggaac aactcggatt gccagcgggata cctgggtcct ctcgagccct 1680
ggcctgtcct tgggcttcct tttgctgtgta tgggctgtgta ggcgggtgtg cagttggca 1740
tgagggtcct agtgatggaa acccccctcat cagcctggtt cttggtttctgt gctgggtctg 1800
ttgccgcttt gcggccggag gttggtccag ggtggtcttt gcacccctac 1860
tcattacat ctacctctctc gctgacagcct cggccctgtt cttgtgatat cggccctgc 1920
tggccgatt cagccacacc ctcctctctc aaccctctct ctcgacacag aatgaggtttt 1980
gggtgagct ttgtgtgagg ctgggttatt ccgattctgct ctgctggctct ccgctggagt 2040
cgccacattatt cgcctcggct ctctctctgc gcgtgctgcc ctggccgtta 2100
cacggtaac cgggttctct gcgggttatt gcgggttatt ccgctgatac cgggtcgcgtt 2160
aacatagac cactcattata ctaacactgtta cttttttgcc ctttgtgttt 2220
tctgaggtgct tctacactgc gctaagcctg ctttgtttttgt ggcggtcctct 2280
gctgccgctct gcggccgcttt ggcggtggtt gggcgttggtt cagccgcttct 2340
agctggagac atggccctgc acagacagct ggcggtggttt gctgtgctttc aacgtggtcg 2400
atattttgtg cttctctgtg cttctcctct ctttctccct cttctctcct 2460
caggtgctgg gacagcagct gccaacacc ctcgagctggct ctgctgtcgtt cagcagctgg 2520
cgttcagcgg gcggcccgaac cttctctgga ctgctgtcgtt cagcagctgg 2580
aatggcgttt cgcttcctca ctcgcccccc cccccccttc gattttttttt ctgcaaccctc 2640
agaacttgct attttttttc ctttctcact gctgacactg ctttcggtct 2700
 acctcctaat ctcgaggtat cccctcctac gcacccctac gggcgtgttgct ctgcttggtg 2760
 tgtgtctcttg tcaacagctgtt aggccctcgct ctttctagttt gctaaggtctg 2820
cgcataattac ctcgaggtat cccctcctac gcacccctac gggcgtgttgct ctgcttggtg 2880
 tacaactcct ttcgaggtat cccctcctac gcacccctac gggcgtgttgct ctgcttggtg 2940
 tataggctct ctcgaggtat cccctcctac gcacccctac gggcgtgttgct ctgcttggtg 3000
 tggagaggct gcacccctac gggcgtgttgct ctgcttggtgct gcacccctac gggcgtgttgct 3060
gccagctcgg cggctcgggc cccctcctac gcacccctac gggcgtgttgct ctgcttggtgct 3120
 acctcctac gcggcgtgttc acggcgtgttc acggcgtgttc acggcgtgttc acggcgtgttc 3180
ggcggttctct gcggcgtgttc acggcgtgttc acggcgtgttc acggcgtgttc acggcgtgttc 3240
cattgacctg tgcatacag acctataattt ctaggcctag gcagaagctca ggcatacaca  3300
ccacccacct atttttaaagt gctgaattgt tcaattattatt tataaatata atcaataata  3360
ttttttttag tataataaaa tagatatag ctgaggtata ataataaata ggtttataa  3420
aataagaacct atataatatata tttttaaaggt ctttattgata ataaaataata  3480
gygacatac tgtaagataat tattgtttta ttttttttttttgc ttkatgcatta  3540
tgacatagtgt atcaatatattta taatattata gcaaatccgt atgcaatctc agttagtta  3600
atatttaag ctataaatattc caacagatata gtaaatagta aacctctgtaa ttctctaggac  3660
tatgattttaa gtcacacacat acacagggcct gcgaagtcag cgcaagatgtt cgcataaac  3720
ttggggactaa agccgtatac accataacta agtcttcacct tattttgggtc caaggacttcc  3780
ttttagattt actatgaactgt gtttaagctgt ttttatttttttct gtaggcgtta  3840
ttttctgtagc tccccctgctc agctctgcgt atactagtat gcattctgacac acagcttacc  3900
ggggcctgaco caagctgtag atttggaacac gatgctataca aacctgctatt gttcgcgccg  3960
agacagtcgt tggagatagtga caccgcgagc accagttcgc accttctttcc gccattagaaa  4020
cctgagaetzt tggatggtttc ttatatacatg atgtcatcata cttggaacatg gagagctgt  4080
gggacaaat cttctaatgg gtcatacataaa aaaaaaccat cctctatgtgc gttgcacact  4140
gacaagctga tagttggatt taatagcagaca aatgggcttc gcaagaagagc cgtagccttcg  4200
aatctccgatt cttggcacag cttgagtattt tggcaacgac ctttcaatcct ctttctgcccctcct  4260
tacagactca aaccccaaac cttgccaggt tgttagagcg agttaaaacat aaaaaacacag  4320
agagcccttt gttctccccaa cacacgtttt ttttcttctg tagctctcagc tttttttctttcttg  4380
acagttctta cagttcaattta cttcatacataa gttggcataat ctttctttctcctcctcc  4440
cacggaacagt ggccctgagg agtggccctg cttgagccaat tgggttccttccctccctc  4500
agagggattt ccacacagttct ttctctcagtg aaacagctgc agggcgccgaga gttgcacao  4560
agttttaggtg acctacagcct aataagggag aacacattag ggggtacaatt tggagccatc  4620
ccagagagag aggttgcacg tagctgcag aacgtgtctttt ctctgtgtgat cttctggcctc  4680
ctgtagatgc gtaatcagcg tgcctcagcc ctgctccaaag acaccggcgc gggggagcgt  4740
gcagtagcct cagctgtcctt cccaacgagc ggaatttctt ttctcgtatc ccaacgtttt  4800
tgggtgtctg atggatcattta aatcctcataa cttgggttagc cttgatccctt cggagcagcc  4860	catcaccacct gggggtgagat gtggcgcgcg gcacagacgag aataactgcag  4920
tcgcgttctg tgcgcgtcctg ctctcgtgctg ccgcgtacag ggctgatcct gcgggtgccga  4980
tcgggtcttg gttcgtgatcgc gtcggtcagtc ggctggtgac gcgaagggca aataactacg  5040
tccctccggca gccggaggct tgtcataata cttctgtctt gcgttgatgc gcgggttcgg  5100
gttctgtgc cccggcagct gcgttcttgg gtttattggtc cttgcgtttc gcgtgtttgcgt  5160
gcgttttatttg gtcggtggc gtttattggtc cttgcgtttc gcgtgtttgcgt  5220
tactgctgt cttgtcgcagc tcagcgtcagc tccaattcct ctttgctcgc gccacatctca  5280
tgcagcgcct ggtcgcctgct gcgtttccttc cggcgcctgg cggcgcctgg cggcgcctgg  5340
cgcctgcatcgctgcgtgcgc cttgatggaag agttgcttgc ggtgagagac agtggatggtc  5400
cgcgcttcagcttgcgctaatgcgttgcgtcgc ggggagagac cggcggcggc ggggagagac  5460
cgcgcttcagcttgcgctaatgcgttgcgtcgc ggggagagac cggcggcggc ggggagagac  5520
tcgcgcggtac gcgggcgcgc gcgggcgcgc gcgggcgcgc gcgggcgcgc gcgggcgcgc  5580
What is claimed:

1. An isolated nucleic acid molecule comprising: a nucleotide sequence consisting of or complementary to ORF8 of SEQ ID NO: 15 or a nucleotide sequence encoding a variant which hybridizes under high stringency conditions of 0.1x SSC, 0.1% SDS at 65° C. to ORF8 of SEQ ID NO: 15 and wherein the encoded polypeptide of said variant has γ-type pyridoxal phosphate (PLP) enzyme activity.

2. The isolated nucleic acid of claim 1, wherein nucleotide sequence is ORF8 which encodes lloC in SEQ ID NO: 15.

3. An expression vector comprising the nucleic acid compound of claim 1.

4. An isolated host cell transformed with an expression vector of claim 3.

5. The host cell according to claim 4 wherein the host cell is a bacterium.

6. The host cell according to claim 5 wherein the host cell is E. coli.

7. The host cell according to claim 4 wherein the host cell is Neotyphodium.