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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
10 August 2023 (10.08.2023)

WIPO | PCT

(10) International Publication Number
WO 2023/148478 A1

(51) International Patent Classification:
A01H 5/12 (2018.01) *A24B 15/10* (2006.01)
A01H 6/82 (2018.01) *C07K 14/415* (2006.01)
A24B 15/00 (2006.01) *C12N 9/12* (2006.01)
C12N 15/82 (2006.01)

(21) International Application Number:
PCT/GB2023/050211

(22) International Filing Date:
31 January 2023 (31.01.2023)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
2201414.6 03 February 2022 (03.02.2022) GB

(71) Applicant: **NICOVENTURES TRADING LIMITED** [GB/GB]; Globe House, 1 Water Street, London WC2R 3LA (GB).

(72) Inventors: **BEN KHALED, Sara**; c/o Nicoventures Trading Limited, Globe House, 1 Water Street, London WC2R 3LA (GB). **ANASTACIO DE ABREU E LIMA, Francisco**; c/o Nicoventures Trading Limited, Globe House, 1 Water Street, London WC2R 3LA (GB). **GALDON-ARMERO, Javier**; c/o Nicoventures Trading Limited, Globe House, 1 Water Street, London WC2R 3LA (GB).

(74) Agent: **D YOUNG & CO LLP**; 120 Holborn, London EC1N 2DY (GB).

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

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

Published:
— *with international search report (Art. 21(3))*
— *with sequence listing part of description (Rule 5.2(a))*

(54) Title: METHOD OF MODULATING ALKALOID CONTENT IN TOBACCO PLANTS

(57) Abstract: The present invention provides a method for modulating the alkaloid content of a plant (e.g. a tobacco plant), the method comprising modifying said plant by modulating the activity or expression of a FAD synthetase. The present invention also provides for the use of a FAD synthetase for modulating the alkaloid content of a plant, as well as tobacco cells, plants, plant propagation materials, harvested leaves, processed tobaccos, or tobacco products obtainable in accordance with the invention.

WO 2023/148478 A1

METHOD OF MODULATING ALKALOID CONTENT IN TOBACCO PLANTS

FIELD OF THE INVENTION

The present invention relates to methods of modulating the alkaloid content a plant or part thereof or cell or cell culture. The invention also extends to methods of modulating the expression and/or activity of polypeptides which modulate alkaloid content within plants. Alternatively, the invention provides methods of modulating the expression and/or activity of genes which encode polypeptides which modulate alkaloid content within plants. The invention also extends to constructs, which can be used to modulate the polypeptides. The invention further relates to plant cells and plants modified to achieve a modulation in alkaloid content. The invention also relates to a processed and harvested leaf from such modulated plants and use thereof in a tobacco industry product, including combustible smoking articles.

BACKGROUND

Alkaloids are a group of naturally occurring compounds which mostly contain basic nitrogen atoms and are produced by a large variety of organisms including bacteria, fungi, plants and animals.

Alkaloids may be classified according to the similarity of the carbon skeleton e.g. indole-, isoquinoline- and pyridine-like. Pyridine derivatives are one class of monomeric alkaloids; this class includes simple derivatives of pyridine, polycyclic condensed and noncondensing pyridine derivatives and sesquiterpene pyridine derivatives. Examples are nicotine, nornicotine, pseudooxynicotine, anabasine, myosmine and anatabine.

Most of the known biological functions of alkaloids are related to protection. Neuroactive molecules, such as caffeine, cocaine, morphine, and nicotine, act as defence compounds against invading predators. The accumulation of these alkaloids is the result of signal transduction cascades that monitor gene expression, enzyme activities, and alkaloid concentrations. The fine-tuning of alkaloid content in the plant involves negative feedback loops and degradative pathways.

Nicotine occurs naturally in several varieties of plants but is found at the highest level in the tobacco plant. Cultivated tobacco produces 2-4% alkaloids of total dry weight. Nicotine is produced in wild and cultivated *Nicotiana* species and plays an important role in plant defence against herbivores and insects (Voelckel et al. (2001) Oecologia 127(2): 274-280, incorporated herein by reference). It accounts for ~90% of the total alkaloid content. The remaining 10% of the alkaloid pool is mostly constituted by the structurally related compounds nornicotine, anatabine, anabasine and pseudooxynicotine (PON).

The regulation of alkaloid content in tobacco is complex. Several factors including genotype, environment, fertilization and agronomic practices (e.g. topping) affect alkaloid levels in tobacco plants. Some key regulators of nicotine biosynthesis are well characterized, for example putrescine N-methyltransferase (PMT), which plays a pivotal role in this pathway, is activated by members of the ethylene responsive factor (ERF) superfamily, the largest transcription factor family in the tobacco genome (Rushton et al. (2008) Plant Physiol. 147(1): 280-295 incorporated herein by reference).

Tobacco pyridine alkaloids are precursors of tobacco-specific nitrosamines (TSNAs) that form during the post-harvest leaf curing. The four primary TSNAs found in cured tobacco leaves are N'-nitrosonornicotine (NNN), N'nitrosoanatabine (NAT), N'-nitrosoanabasine (NAB) and 4-(methyl nitrosamino)-1-(3-pyridyl)-1-butanone (NNK). During the post-harvest leaf curing, reactions between pyridine alkaloids and nitrosating species leads to the formation of TSNAs. PON is likely to function as the direct precursor in the synthesis of the TSNA NNK (Bush *et al*., 2001 incorporated herein by reference). Reducing the production and accumulation of TSNAs is of high importance. The CYP82E family of nicotine demethylase genes is one of the primary regulators of nicotine to nornicotine conversion, and altering their activity or accumulation may result in a decrease in NNN levels.

As described in the Examples, the inventors sought to investigate genes responsible for alkaloid and/or TSNA precursor synthesis, with the aim of modulating alkaloid content in plants, e.g. decreasing TSNA content in tobacco.

SUMMARY OF THE INVENTION

It has been surprisingly found that by modulating the activity or expression of a Flavin adenine dinucleotide (FAD) synthetase, the alkaloid content and/or TSNA content or precursor of TSNA content of plants can be modulated. The gene(s) as taught herein, for example Nitab4.5_0005997g0050.2, are regulators of alkaloid and TSNA precursor content in cultivated tobacco. In particular the gene(s) as taught herein, for example Nitab4.5_0005997g0050.2, is a regulator of alkaloid content in cultivated tobacco. Nitab4.5_0005997g0050.2 encodes a FAD synthetase according to the present invention. Exemplary homologues of Nitab4.5_0005997g0050.2 are provided in Table 1. Suitably FAD synthetases and homologues thereof according to the present invention may comprise a FAD synthase domain.

According to the present invention, tobacco industry products with modulated alkaloid content and commercially desirable traits sought after by consumers of tobacco industry products can be produced. In some instances, consumers may desire a product with low levels of alkaloid

content. In some instances, consumers may desire a product with low levels of TSNA precursors.

The present invention may be particularly useful in the field of plant molecular farming, where plants (such as tobacco and other *Nicotiana* spp.) are used for the production of proteins, peptides, and metabolites e.g. for the production of therapeutics and pharmaceuticals such as antibiotics, virus like particles, or neutraceuticals or small molecules. Tobacco has been used for the development of an HIV-neutralising antibody in an EU-funded project called PharmPlant and Medicago Inc., Canada have worked on a tobacco-based platform for the production of virus-like particles for flu vaccine manufacture.

Thus, a plant according to the present invention may be used for molecular farming to reduce or eliminate the presence of nicotinic alkaloids. The use of a low nicotine plant or rootsock is beneficial in molecular farming and would reduce downstream processing costs associated with purification.

The present inventors have surprisingly determined a method for modulating (e.g. decreasing) the alkaloid content, of a plant (e.g. a tobacco plant) by modulating (e.g. decreasing) the activity or expression of a FAD synthetase. The alkaloid content (e.g. the content of one or more of nicotine, nornicotine, PON, anabasine, anatabine or myosmine, suitably the content of one or more of nicotine, nornicotine, PON, anabasine or anatabine) of a plant (e.g. tobacco plant) may be decreased by decreasing the activity or expression of a FAD synthetase or may be increased by increasing the activity or expression of a FAD synthetase.

Prior to the present invention it had not been known that modulation of the activity or expression of a FAD synthetase as described herein could be used to modulate alkaloid content or modulate TSNA precursor content in particular, nornicotine, PON, anabasine and/or anatabine content.

In one aspect, the present invention provides a method of modulating (e.g. decreasing) the alkaloid content of a tobacco plant or a part thereof or tobacco plant cell, the method comprising modifying said plant or plant cell by modulating (e.g. decreasing) the activity or expression of a FAD synthetase.

In one aspect, the present invention provides a method of modulating (e.g. decreasing) the content of a tobacco specific nitrosamine (TSNA) or a precursor of a TSNA in a tobacco plant or plant part thereof or tobacco plant cell, the method comprising modifying said plant or plant cell by modulating (e.g. decreasing) the activity or expression of a FAD synthetase.

In another aspect, the present invention provides a method for producing a plant or part thereof, a cell or cell culture, a plant propagation material, a leaf, a cut harvested leaf, a processed leaf or a cut and processed leaf which has modulated (e.g. decreased) alkaloid content, the method

comprising modifying said plant or cell culture to modulate the activity or expression of a FAD synthetase.

In another aspect, the present invention provides the use of at least one gene encoding a FAD synthetase for modulating alkaloid content of a tobacco cell or tobacco plant or part thereof; wherein the FAD synthetase.

In one aspect, the FAD synthetase may:

a) comprise an amino acid sequence as set out in SEQ ID No. 3; or a functional variant or functional fragment or orthologue of SEQ ID No. 3; or a sequence which has at least 80% identity to SEQ ID No. 3, or a homologue or SEQ ID No. 3; or

b) be encoded by a nucleotide sequence as set out in SEQ ID No. 1 or 2; or a functional variant or functional fragment or orthologue of SEQ ID No. 1 or 2; or a nucleic acid sequence which has at least 80% identity to SEQ ID No. 1 or 2, or a homologue of SEQ ID No. 1 or 2.

Suitably, the alkaloid content may be modulated (e.g. decreased) in comparison to a plant or cell culture which has not been modified to modulate the activity or expression of a FAD synthetase.

In one aspect, the present invention provides a tobacco plant or part thereof or a tobacco cell or cell culture which has been modified to modulate (e.g. decrease) the activity or expression of a FAD synthetase and the tobacco plant or part thereof or tobacco cell or cell culture has decreased alkaloid and or TSNA precursor content in comparison to an unmodified plant or unmodified cell or cell culture.

In another aspect, the present invention provides a plant propagation material obtainable (e.g. obtained) from a plant according to the present invention or from a plant or cell or cell culture produced by a method or use according to the present invention.

Suitably, the alkaloid content of the plant may be decreased in comparison to a plant or cell culture which has not been modified to modulate the activity or expression of a FAD synthetase.

Suitably, the content of one or more alkaloids selected from nicotine, nornicotine, PON, anabasine, myosmine and anatabine is modulated (e.g. decreased), preferably the content of nicotine, nornicotine and/or PON may be modulated (e.g. decreased).

Suitably the nicotine content may be decreased.

In another aspect, the present invention provides the use of a tobacco plant or part thereof or tobacco cell or cell culture according to the present invention, or of a plant produced by a method according to the present invention to breed a plant.

In one aspect, the present invention provides the use of a tobacco plant or part thereof or a tobacco cell or cell culture according to the present invention, or of a plant produced by a method according to the present invention for production of a product.

In another aspect, the present invention provides the use of a tobacco plant or part thereof according to the present invention, or of a plant produced by a method according to the present invention to grow a crop.

In one aspect, the present invention provides the use of a tobacco plant or part thereof according to the present invention, or of a plant produced by a method according to the present invention to produce a leaf.

In one aspect, the present invention provides a harvested leaf of a plant according to the present invention, or obtainable from a plant propagated from a propagation material according to the present invention, or obtainable from a plant obtained by a use according to the present invention, or obtainable from a plant produced by a method according to the present invention.

Suitably, the harvested leaf of a plant may be a cut harvested leaf.

In one aspect, the present invention provides a processed leaf, preferably a processed tobacco leaf, preferably a non-viable processed tobacco leaf:

obtainable (e.g. obtained) from a plant obtainable from a use according to the present invention;

obtainable (e.g. obtained) by processing a plant according to the present invention;

obtainable (e.g. obtained) from a plant propagated from a plant propagation material according to the present invention; or

obtainable (e.g. obtained) by processing a harvested leaf of a plant according to the present invention; or

obtainable (e.g. obtained) from a plant produced by a method according to the present invention .

Suitably, the leaf may be processed by curing, fermenting, pasteurising or a combination thereof.

Suitably, the processed leaf may be a cut processed leaf.

In one aspect, the present invention provides cured tobacco material made from a plant or a part thereof according to the present invention, or a harvested leaf according to the present invention, or a processed leaf according to the present invention.

In one aspect, the present invention provides a tobacco blend comprising cured tobacco material according to the present invention.

In one aspect, the present invention provides a tobacco industry product prepared from:

a tobacco plant or part thereof or tobacco cell or cell culture according to the present invention;

a tobacco plant or part thereof propagated from a tobacco plant propagation material according to the present invention;

a harvested leaf of a plant according to the present invention;

a processed leaf according to the present invention.

Suitably the tobacco product may be a combustible smoking article.

Suitably the tobacco product may be a smokeless tobacco product.

Suitably the tobacco product may be a non-combustible aerosol provision system such as a tobacco heating device or an aerosol-generating device.

In another aspect, the present invention provides a combustible smoking article, non-combustible aerosol provisioning system, smokeless tobacco product or tobacco heating device comprising a plant or a part thereof according to the present invention or an extract (e.g. a tobacco extract) thereof or a tobacco cell culture according to the present invention; or a cured tobacco material according to the present invention; or a tobacco blend according to the present invention.

In another aspect, the present invention provides the use of a polynucleotide sequence encoding a FAD synthetase protein which:

a) encodes an amino acid sequence as set out in SEQ ID No. 3; or a functional variant or functional fragment or orthologue of SEQ ID No. 3; or a sequence which has at least 80% identity to SEQ ID No. 3, or a homologue of SEQ ID No. 3; or

b) comprises a sequence as set out in SEQ ID No. 1 or 2; or a functional variant or functional fragment or orthologue of SEQ ID No. 1 or 2; or a nucleic acid sequence which has at least 80% identity to SEQ ID No. 1 or 2, or a homologue of SEQ ID No. 1 or 2;

to select a plant having modulated (e.g. reduced) alkaloid content and/or modulated (e.g. reduced) content of TSNA or a precursor of a TSNA.

In one aspect, the present invention provides a mutant of a plant carrying a heritable mutation in a nucleotide sequence which:

a) encodes an amino acid sequence as set out in SEQ ID No. 3; or a functional variant or functional fragment or orthologue of SEQ ID No. 3; or a sequence which has at least 80% identity to SEQ ID No. 3 or a homologue of SEQ ID No. 3; or

b) comprises a sequence as set out in SEQ ID No. 1 or 2; or a functional variant or functional fragment or orthologue of SEQ ID No. 1 or 2; or a nucleic acid sequence which has at least 80% identity to SEQ ID No. 1 or 2, or a homologue of SEQ ID No. 1 or 2;

wherein said heritable mutation modulates (e.g. decreases) the activity or expression of FAD synthetase and wherein the mutant plant has modulated (e.g. decreased) alkaloid content and/or modulated content of a TSNA or a precursor of a TSNA relative to a comparable plant which does not carry said heritable mutation.

In one aspect, the present invention provides progeny or seed of a mutant plant which carries the heritable mutation according to the present invention.

In one aspect, the present invention provides a harvested leaf, a processed leaf or cured tobacco material produced from a plant comprising a modification in a nucleotide sequence which:

a) encodes an amino acid sequence as set out in SEQ ID No. 3; or a functional variant or functional fragment or orthologue of SEQ ID No. 3; or a sequence which has at least 80% identity to SEQ ID No. 3, or a homologue of SEQ ID No. 3; or

b) comprises a sequence as set out in SEQ ID No. 1 or 2; or a functional variant or functional fragment or orthologue of SEQ ID No. 1 or 2; or a nucleic acid sequence which has at least 80% identity to SEQ ID No. 1 or 2, or a homologue of SEQ ID No. 1 or 2;

wherein said modification modulates (e.g. decreases) the activity or expression of FAD synthetase and wherein said plant has modulated (e.g. decreased) alkaloid content and/or modulated content of a TSNA or a precursor of a TSNA relative to a comparable plant which does not carry said modification in said FAD synthetase.

In one aspect, the activity or expression of at least one Nic1 ERF gene (such as any one or more of those in Figure 7) and/or at least one Nic2 ERF gene (such as any one or more of those in Figure 7) is modulated in addition to the at least one FAD synthetase. Suitably, said at least one Nic1 ERF and/or Nic2 ERF gene may comprise a mutation which decreases its expression and/or activity. In one aspect, the activity of ERF199 is modulated (e.g. decreased) in addition to the modulation (e.g. decrease in activity and or expression) of the at least one FAD synthetase. In one aspect, the activity of ERF189 is modulated (e.g. decreased) in addition to the modulation (e.g. decrease in activity and or expression) of the at least one FAD synthetase. In one aspect, the activity of ERF199 and ERF189 is modulated (e.g. decreased) in addition to the modulation (e.g. decrease in activity and or expression) of the at least one FAD synthetase. Exemplary sequences of ERF 199 and ERF199 are provided in Figures 8-15 (SEQ ID Nos 35-42). See also WO2018237107, which is incorporated herein in by reference in its entirety.

BRIEF DESCRIPTION OF THE DRAWINGS

Embodiments of the invention will now be described, by way of example only, with reference to the accompanying drawings, in which:

Figure 1 shows the alkaloid content of 5-week-old TN90 leaves silenced for Nitab4.5_0005997g0050.2. Content is represented relative to control and comprises three biological replicates analysed by one-way ANOVA and Tukey's multiple-comparison post-test. Values are shown as means ± SEM. Asterisks indicate statistical significance of P value ≤ 0.001. Pyridine alkaloids: nicotine, nornicotine, anabasine (ANAB), anatabine (ANAT), pseudooxynicotine (PON).

Figure 2 shows measured endpoint absorbance values (570 nm) for the production of FAD over a 60 minute reaction by 0.5 µg of Nitab4.5_0005997g0050.2 enzyme in the presence of 10 µM adenosine triphosphate) ATP and 10 µM FMN. The results for the control reactions ((maltose binding protein (MBP) and the no enzyme control (-ve)) are also shown on the same graph and indicate that the reaction is FMN-dependant and that the reaction has reduced activity in the absence of added ATP. Errorbars are standard deviation from four replicates.

Figure 3 shows the genomic sequence of Nitab4.5_0005997g0050.2 (SEQ ID No. 1).

Figure 4 shows the coding sequence of Nitab4.5_0005997g0050.2 (SEQ ID No. 2).

Figure 5 shows the amino acid sequence of Nitab4.5_0005997g0050.2 (SEQ ID No. 3).

Figure 6 shows the sequence of TRV1 used in Example 1.

Figure 7 provides a table of Nic1 ERFs and Nic2 ERFs

Figures 8-15 show the sequences of ERF199 (SEQ ID Nos 35-38) and ERF189 (SEQ ID Nos 39-42).

Some sequences disclosed herein contain "X" or "N" in nucleotide sequences. "X" or "N" can be any nucleotide or a deletion or insertion of one or more nucleotides. For example, in some cases a string of "X"s or "N"s are shown. The number of "X"s or "N"s does not necessarily correlate with the actual number of nucleotides at that position. There may be more or fewer nucleotides than shown as "X" or "N" in the sequence.

DETAILED DESCRIPTION

For the first time the present inventors have shown that by modulating (e.g. decreasing) the activity or expression of at least one FAD synthetase in a plant (e.g. a tobacco plant) or a cell (e.g. tobacco cell), the alkaloid and/or TSNA precursor content of the plant (or processed plant) or cell can be modulated (e.g. decreased).

FAD synthetase catalyses the adenylation of flavin mononucleotide (FMN) to form flavin adenine dinucleotide (FAD) coenzyme. FAD can be present in different forms (FAD, FADH,

FADH2), being converted between these states by accepting or donating electrons. These forms changes are possible because of its redox-active isoalloxazine ring in its structure.

This cofactor is involved in multiple biological oxidative-reduction reactions in metabolism, being essential for all forms of life. In these reactions, using organic compounds as electron donors, organisms can generate metabolic energy that can be driven using different energy carriers, such as ATP. However, these reactions are also important for biosynthesis, where this energy can be used to make complex compounds.

The present invention provides a method of modulating (e.g. decreasing) the alkaloid content of a plant or a part thereof, the method comprising modifying said plant by modulating (e.g. decreasing) the activity or expression of at least one FAD synthetase.

Also provided is a method of modulating (e.g. decreasing) the content of a TSNA precursor in a tobacco plant or plant part thereof, the method comprising modifying said plant by modulating (e.g. decreasing) the activity or expression of at least one FAD synthetase.

The at least one FAD synthetase may be selected from an amino acid sequence as set out in SEQ ID No. 3, or a functional variant or functional fragment or orthologue thereof, or a sequence which has at least 80% identity to SEQ ID No. 3, or a homologue of SEQ ID No. 3; or said at least one FAD synthetase may be encoded by a which comprises a sequence as set out in SEQ ID No. 1 or 2, or a functional variant or functional fragment or orthologue of SEQ ID No. 1 or 2, or a nucleic acid sequence which has at least 80% identity to SEQ ID No. 1 or 2, or a homologue of SEQ ID No. 1 or 2.

Suitably, the protein may comprise a FAD synthase domain. In one embodiment at least two genes encoding FAD synthetases are modified selected from the group of: genes which encode polypeptides comprising an amino acid sequence as set out in SEQ ID No. 3 or a functional variant or functional fragment or orthologue thereof, or a sequence which has at least 80% identity to SEQ ID No. 3, or a homologue of SEQ ID No. 3; or genes encoding a FAD synthetase comprising a nucleotide sequence as set out in SEQ ID No. 1 or 2, or a functional variant or functional fragment or orthologue of SEQ ID No. 1 or 2, or a nucleic acid sequence which has at least 80% identity to SEQ ID No. 1 or 2, or a homologue of SEQ ID No. 1 or 2.

In one embodiment, at least three, such as at least four, such as at least five, such as at least six, such as at least seven, such as at least eight, such as at least nine, such as ten FAD synthetases are modulated, wherein the FAD syntehtases comprise an amino acid sequence as set out in SEQ ID No. 3 or a functional variant or functional fragment or orthologue thereof, or a sequence which has at least 80% identity to SEQ ID No. 3, or a homologue of SEQ ID No. 3; or wherein the at least one FAD synthetase comprises a nucleotide sequence as set out in SEQ ID No. 1 or 2, or a functional variant or functional fragment or orthologue of SEQ ID No. 1 or 2, or

a nucleic acid sequence which has at least 80% identity to SEQ ID No. 1 or 2, or a homologue of SEQ ID No. 1 or 2. Suitably, the protein may comprise a FAD synthase domain. In one aspect, the at least one FAD synthetase comprises or consists of an amino acid sequence as set out in: SEQ ID No. 3 or a functional variant or functional fragment or orthologue thereof, or a sequence which has at least 80% identity to SEQ ID No. 3; or wherein the FAD synthetase comprises a nucleotide sequence as set out in SEQ ID No. 1 or 2 or a functional variant or functional fragment or orthologue of SEQ ID No. 1 or 2; or a nucleic acid sequence which has at least 80% identity to SEQ ID No. 1 or 2. Suitably, the FAD synthetase may comprise a FAD synthase domain.

In one aspect, the activity or expression of at least one further gene is modulated. Suitably, at least two (or at least three or at least four or at least five or at least six or at least seven or at least eight or at least nine) additional genes selected from Table 1 or a sequence having at least 80% sequence identity thereto may also be modulated.

The “expression” of a FAD synthetase may refer to the level of transcription, translation i.e. protein expression.

Measurement of the level or amount of a gene product may be carried out by any suitable method, for example comparison of mRNA transcript levels, protein or peptide levels, and/or phenotype of a plant, between a modified plant and comparable plant which has not been modified according to the present invention.

The term “a comparable product” as defined herein would be one derived from a plant (e.g. a tobacco plant) which had not been modified according to the present invention, but in which all other relevant features were the same (e.g. plant species, growing conditions, method of processing the plant, e.g. tobacco, etc.). The comparable product according to the present invention may mean a plant (e.g. a tobacco plant) or a part thereof, such as a leaf (e.g. a tobacco leaf), a harvested leaf (e.g. a harvested tobacco leaf), a cut harvested leaf (e.g. a cut harvested tobacco leaf), a processed leaf (e.g. a processed tobacco leaf) or plant propagation material (e.g. tobacco plant propagation material), or a product comprising said plant or part therefore, e.g. a tobacco industry product or combinations thereof obtainable or obtained from a plant which has not been modified in accordance with the present invention, e.g. to modulate the activity or expression of gene encoding a FAD synthetase. In one embodiment a comparable product is one which does not comprise gene encoding a FAD synthetase whose activity or expression has been modulated.

The term “modifying” or “modified” as used herein means a plant (e.g. a tobacco plant) or nucleic acid sequence that has been altered or changed. The present invention comprises the modification of plants using techniques for genetic modification of plants or non-genetic

modification of plants. Such methods are well known in the art and examples of genetic modification techniques include transformation, transgenics, cisgenics, and gene editing methods. Examples of non-genetic modification techniques include fast-neutron mutagenesis, chemical mutagenesis e.g. ethyl methanesulfonate (EMS) mutagenesis and modern population analysis approaches.

In one embodiment a natural variant which has a modified gene encoding a FAD synthetase is selected and that trait or gene is bred into a second plant which may have commercially desirable traits.

In one embodiment the plant according to the present invention is a transgenic plant. In one embodiment the plant according to the invention is a non-transgenic plant.

The term “unmodified plant” as defined herein would be a plant (e.g. a tobacco plant) which had not been modified according to the present invention, e.g. to modulate the activity or expression of a FAD synthetase or to modify the nucleic acid sequence of at least one gene encoding an FAD synthetase; and in which all other relevant features were the same (e.g. plant species, growing conditions, method of processing tobacco, etc.). In one embodiment an unmodified plant is one which does not comprise a gene encoding a FAD synthetase whose activity or expression has been modulated. In one embodiment, an unmodified plant is one which does not comprise a modified nucleic acid sequence which encodes at least one gene encoding a FAD synthetase protein.

FAD synthetase

A “FAD synthetase” as used herein has its usual meaning in the art and refers to an enzyme which catalyses the adenylation of flavin mononucleotide (FMN) to form flavin adenine dinucleotide (FAD) coenzyme (ATP + FMN → diphosphate + FAD).

Methods for measuring FAD synthetase activity are known in the art and include, for example commercial assay kits sold by Abcam (Flavin Adenine Dinucleotide (FAD) Assay Kit - ab204710). Typically colorimetric or fluorometric methods are used to assay the activity of FAD which functions as the cofactor of an oxidase which catalyzes the formation of a product that reacts with a probe generating colour and fluorescence.

An illustrative sequence of a FAD synthetase protein from tobacco is shown in SEQ ID No. 3.

A “FAD synthase domain” as used herein has its usual meaning in the art and refers to a domain required for FAD synthase activity.

An illustrative sequence of a FAD synthase domain is set forth at amino acids 94-279 of SEQ ID No. 3.

A FAD synthase binding domain may be identified by comparing the protein in question to amino acids 84-279 of SEQ ID No. 3.

Suitably, a FAD synthase domain as used herein may refer to a sequence set forth in amino acids 94-279 of SEQ ID No. 3 or a sequence which has at least 80% identity thereto. Suitably, a FAD synthase domain as used herein may refer to a sequence which corresponds to amino acids 94-279 of SEQ ID No. 3 when aligned with SEQ ID No. 3.

Domains within the amino acid sequence of a protein may be identified using domain prediction software known in the art. Domains are also described in protein databases such as UniprotKB.

Without wishing to be bound by theory, it is hypothesized that modulating content of a FAD synthetase in a plant cell or modulating activity, such as FAD synthtase activity, of a FAD synthetase in a plant would alter the metabolic pathways producing alkaloids and TSNA precursors, resulting in modulated alkaloid and/or TSNA precursor content.

In one embodiment, a FAD synthetase comprises an amino acid sequence shown as SEQ ID No. 3, or a sequence which has at least 80% identity thereto, or a homologue thereof. Suitably, the FAD synthetase comprises a FAD synthase domain. Suitably, a homologue of SEQ ID No. 3 may be selected from the group comprising the amino acid sequences provided in Table 1 or a sequence which has at least 80% identity thereto. Suitably, a homologue of SEQ ID No. 3 may be selected from a sequence having at least 80% sequence identity to a sequence provided in Table 1 and wherein said sequence comprises a FAD synthase domain.

In one embodiment, a FAD synthetase comprises an amino acid sequence shown as SEQ ID No. 3, or a sequence which has at least 80% identity thereto (preferably at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity thereto). In one embodiment a FAD synthetase comprises an amino acid sequence shown in Table 1, or a sequence which has at least 80% identity thereto (preferably at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity thereto). Suitably, said FAD synthetase comprises a FAD synthase domain.

In one embodiment, the FAD synthetase according to the present invention comprises or consists of an amino acid shown as SEQ ID No. 3. In one embodiment, the FAD synthetase according to the present invention comprises or consists of an amino acid shown in Table 1.

Suitably, the protein may be from *Nicotiana tabacum*.

In one embodiment, a FAD synthetase is encoded by a polynucleotide sequence wherein the gene (prior to mutation) comprises a polynucleotide sequence shown as SEQ ID No. 1, or a sequence which has at least 80% identity thereto, or a homologue thereof. Suitably, the FAD synthetase comprises a FAD synthase domain. Suitably, a homologue of SEQ ID No. 1 may be selected from the group comprising the polynucleotide sequences provided in Table 1 or a

sequence which has at least 80% identity thereto. Suitably, a homologue of SEQ ID No. 1 may be selected from a sequence having at least 80% sequence identity to a sequence provided in Table 1 and wherein said sequence comprises a FAD synthase domain.

In one embodiment, a FAD synthetase is encoded by a polynucleotide sequence shown as SEQ ID No. 1, or a sequence which has at least 80% identity thereto (preferably at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity thereto). In one embodiment a FAD synthetase is encoded by a polynucleotide sequence shown in Table 1, or a sequence which has at least 80% identity thereto (preferably at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity thereto). Suitably, said FAD synthetase comprises a FAD synthase domain.

In one embodiment, the FAD synthetase is encoded by a polynucleotide sequence which comprises or consists of a polynucleotide sequence shown as SEQ ID No. 1. In one embodiment, the FAD synthetase is encoded by a polynucleotide sequence which comprises or consists of a polynucleotide sequence shown in Table 1.

In one embodiment, a FAD synthetase is encoded by a polynucleotide sequence wherein the gene (prior to mutation) comprises a polynucleotide sequence shown as SEQ ID No. 2, or a sequence which has at least 80% identity thereto, or a homologue thereof. Suitably, the FAD synthetase comprises a FAD synthase domain. Suitably, a homologue of SEQ ID No. 2 may be selected from the group comprising the polynucleotide sequences provided in Table 1 or a sequence which has at least 80% identity thereto. Suitably, a homologue of SEQ ID No. 1 may be selected from a sequence having at least 80% sequence identity to a sequence provided in Table 1 and wherein said sequence comprises a FAD synthase domain.

In one embodiment, a FAD synthetase is encoded by a polynucleotide sequence shown as SEQ ID No. 2, or a sequence which has at least 80% identity thereto (preferably at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity thereto). In one embodiment a FAD synthetase is encoded by a polynucleotide sequence shown in Table 1, or a sequence which has at least 80% identity thereto (preferably at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity thereto). Suitably, said FAD synthetase comprises a FAD synthase domain.

In one embodiment, the FAD synthetase is encoded by a polynucleotide sequence which comprises or consists of a polynucleotide sequence shown as SEQ ID No. 2. In one embodiment, the FAD synthetase is encoded by a polynucleotide sequence which comprises or consists of a polynucleotide sequence shown in Table 1.

Table 1. Homologues of SEQ ID No. 1, 2 and 3

SEQ ID No:	Sequence
4	**>Nitab4.5_0004979g0070.1 Nitab4.5_0004979:152018-157676** ATGTTGGGTGGGTCTCGCTGTATATCCCAGCACTTAAGAGACACTAATTTTCACCAATTTTGCACTTCCACAACTCCCATTTTTACTGCTGCTTCTCCCAC CAATCTCAACTTCTTCATTCACCTACTTCAGATTCAGATTAAGGAATAAAGTTAGTGATTTTGGGGTCAGTTGTGTTCAGCCCAATTCTTCAAATACCCA AAGCTCTTTGCTTTCTCAAACGTTGAGGTAAATCTTTTACTCTTTCAACATTAATTGTAATGATTTTTGTGCGTAAATTTGAGCTGTAAA TTTAGTCTTTTTGGGTACTAGCCAAATCTCCAGAATATTCAAAGCTCTTTACTTTCTGAAAGTTTGAGGTAGATTATATAACACATCTTTTACTCTTTCAGC AGCAATTTTAGTGATTTTTGGGTTTAAATTTGAGCTGTTAAATTTAGTGATCTTTTGGGGGTGTACATTTGAGCTGTAAAATTTAGTTTTTTGGGTACTTTCT TTGAATTGTGTACTAGCCAAATCTTCAGAATATTCAAAGCTCTTTGCTTTCTGAATGTTTTGAGATAAATTTATAACATCCTTTACTCTTTCAGCAGTAAA TTTAGTGATTTTTTGGGAGTAAATTTGAGCTGTACATTTAGTTTTTTGGGGGGGGTACTTTTTATGAAGAGTGTAAAACCCAAGTCTTCAGAATATCTAAA GCTCTTCACTTTCTGAAAGTTTGAGGTACTTATAATAGCTGTTCTTGTTTGAGCAATAAATCGAGTGAGCTTTGATCTAGTGGTAAGAGTGCAGCACATG ATTCAAACCCTGCAGTGGATTTTGGTATCTAAGTGAGGAGGGTAGAGGAGGCTCGTTATTATTCAGGAATTTCTCGATATTAAAGAATAAAAATT GAGTAAGTTTGGGGTCTGTTTTATACATTATGTGCAACCACAATCCATAGCTCCTTGCTTTCTAATAGTTTGAGATAAATTTTCTTTTCACTTGTATTTC TAAGCCTTTTTGTCATTTAAAAGTAAGTTTAAAATAAAATATTTGTAAGCTCATTATTCTTGCAAATAGAGTTAGTGAAA GAAGGACTGGAAGGAGACTTTCTTTTATAACTGCAATAAGATTGAAAATTTATTGCAACATTAAGTTGGGATTACATTTTTGACGAAAATAATA TGATAAGTTTGTTGGAATATTATATAAGTAGGCGTTCTCTCAGACAGTTTAAAGTTTTAGATGAGCAGATTCACACGCTTCAACCGGATAACAGAGCAGA CACATGCCTGTGTTTGAATCTCGTTGCGTGTCATCTATAAAAAATATTCCGGCGTCTAAAAAAGATTCAGGCTGTGCACATGAGAGCATGTTGTAGACCT AAAACAAATAAATAAGACTGTCTTCCTTCTAATTGTTATAAGTTTTTTAAATGAGAGGGTTATACATTTAAATAGTTTCATCTGGACATGGATTTTGCCGG TAGTTTCCATTTTATGTTTAAATGAAGCCTACCGAGGTCGGGTTTCCATTAATTCTATGTCAGAGTTCTTTATTCCATTGCATTGTAATGTCGT TTTGTATTCTAAAATATGACTTACTTTAGAATCCCAGTGCTTATATGTGTTTTCTTCACAATTTTCATGGAAAAGATGAGGAGCCTC CATCAGAAAGATTGTCCATGCTTGCAGGTGTGCTCTAACCCATGTACTACTTGTAATAATATTGTTGAATGGTATTACATGTATGCTGAATATATATCAA TTCCATTTTTTATCATTTATTGTAAATCATTAAGTTTCTTCAGTGGTTAAGTGCTACACAGATACCTTGTTCTACAACTGTTATCAGCCATTCTTATTTC ATAATTCTGCTCGTTCTTTCCTTTGTAATTTTTTCTTTTCTCTTGGTTCTCTAAAGAAAATGGGGAAAGAGGATAACAAGATTTCCTGCAATAA TGGAAGGCTCTTCTACTTCAGTTCTAACATAAGCTGGAGATCAACTATCTTATCAAAACATGAAATTTCACACAATGAAATCAACCATCAAATTTCCT GCAATCACTTCAGGAAATTACAAAATATACTGCAAGGATATCTTGTCTTTATGTTCTTGGATAACTTCCATTTCAAGGTTAGGGAAAATTTCTCTCACCA AAGCACACCTAGAGATATTGGTAACATGAAGGGTATGTATGACCCTGTGAAAGCTTCAGTTTAATGCCCTCCTCTGACGGAAATTTGACCACAAGATCAT AGAAAAGAAAAAGTGAAAGCCATTGATACATGCTCCAACACGGGTCATTATATAAAGGGAAATATATTGTCCATTCTTTCCTCAATTTGTTTTGCTAACACTTT AGACCTGTTCTTTCCATACCGTGCACTTATTTTTGTATAAAAGAAATAGGGATAATCTAGCTGTGAACGGTTGAATGGTCAATCTGCAAATTTTATTCT AAATTTACTTTGGACCTACAATTACATTTTTAGATTAGTGAGTTTCATTTCTACTTGTTGCTTATCATCAATATATAAAAAAAAGAGTTTCAATTCCACTTG CTAATTTGTGTTGTAAATGAGAAGAAATAGAGGATTCTTGGATATCATTTATGTGTAAATTTCTCGATCATTGGTTTCATTTCAATAGGTGGGATAGTAGC TCTGGGGAAGTTTGATGCACTCCATATTGGTCATCGTGAGCTTGCAATCCAAGCAGCTAAGAGAGGAATTCCATTTCTTCTTCATTGTTGGAATGGCG

	GAAGTACTTGGATGGGAACCGAGGTATAACTTCTGTTCTCTCCTTTACTCTACTCTAATAGAATGTTAATAGAGAAATATTCTCTCCCTGGAAGA
	CCAGAAAAGTAGAGAATGGTGAACTTCGACTCAACTCTCGCTGCTACACCCAAACCTGTAATAGTACCTAGTTCAACCAGGGAGTTCAATCTCGCATTCT
	CTAAATGTGCACTGTCTCTCTTTGCAACCTCTAAGTTGTAGCAATACTGCACCTTCATGCCTCTATTTTTAGCGTAATTCCGTGGACCTACGGATGCTCTTT
	CTCCGAAGCAGTAGATTTGTTGATTTGCACTCTCTTTAGTTTCTGTTAATAGGATGATCAGTTCAGAGTATGTGCAAAGAGCACGCAAATATAA
	CCATATGTTTTTTAATTCAAGTTTTAACAAAAAGGATAAATTAAAGAAATACACTGGCTGAATACATAGATGGCCCTTTAAACTTGCCCTATTTCCC
	ATTAAGACACCTGAACTAATGCTTGTACCTGTTGAGCACTTCACCTATATTCGAGCTGTATCAATTGAACACATATTTTACAATCTAAACTAAGTTCAAC
	ACATGAGTACCACTCTGTATGTAGCAAATATGACCCATAAAAAAGAGCCATGTTAGATAAGGAAAAAGTCAATAGCTGGAGGAAAAAAGGAAACAGTAA
	CCGGAAACAGAGGGACAATACCAACCGGAATTACGTTTCTTGTTATCAAATTATCATCTCCCCGACACTGGCAAAAACAAGTTCATCCCAGATTGAC
	CCTGAATCCCCGTCAACTGCCCGGGCGACTCCATTTTCACCGGATAAAGTTGCCTTGTCGCCAGTATCAATGAATCATCATTTTTGTCATCTCCAAGCTC
	GATTGAGTCATCTTTGGAGAAAGGTTATTTTGGGGGGTAAGTCACATTTTGAGTTCACAAATTGCTGAAACATCCAGATCCCCCTTTATCCCTTTCTTC
	TTCTTCTTCCATTCACAAATACACACCATCCTGGTTTGCCATTGTTCCACCACCGAGGAATCTCACCCCCATGGCATCAACAGATAGGCACCAGATCACA
	CCCTTCCTCCCACAAATTACCATTACATACACAAAGATTCCATCAACTAACCTCTCTCCCCTAAGTGATGGATTTATAGATACGTGAATGTATATATT
	ACTCCTGTATTGTTCATTTAGAATTTTGTATAGTAGCTTTCTTGTTTTAAAAAATAAGATTTTGAAAATTTGCTAAAAAGGTACGACTATTATTATTCTTG
	AGGGAGAATGGGTCAGTTGACGTATTTTTCCTTGGCTATATCATGGGCCGTTTTCTGATGAAGGCGGCGCGTATTATTTAGACGGGTCAGAAAAGAGAAT
	GGTGTCTGAATAGAATGATGGAGAGGAGTGGTGTTTGATAGGCTTTTTATAGGATAAATAAGAAGAAGATGAAAATACTTTATTGGATTTGAATCTGGTAT
	GGAGTCTAAAAAAGAAGAGGAAAATGAGTCGGAGAAGATGACGGAAAATAGCAGTTGTCAGAAATATCAATTTACCGATGATAACTTCATA
	TCTTTTTCAGTAAAATGCGCGTAGTTTCGGGTGAAAGACACGCTACGTGCCATGTCGAGCGAGTTGTGTCTCATAGGCACAGTTTGGCTCAAGTTCAAGTA
	TTCAATAGGCTCAATAGTTAATTTAGGTGTCTAGATGATAAAATTGGACAAGTATAAGGGGCTGCATATGTATTTGGCTTAAATCAATTAATTCCTGCAA
	CAGTGTATATTCCCACAGGAAAACTTAAGTGGATGTCCTTTTTTTTCTTTTTTCTGTTTTTGGTGACTAGTAAAAGTTCCTCCATCTTAAATTCAAAGGA
	TATCCTTAACCTCATGTATTTGCTCCTTGAAAATAAAATATTTATTCTGCTCAAGTAGTATATAGGAAAAGTCATGGTTTATAAAATTTG
	GTCTATGAAGCTGGAAGGTGTTTATAATCTGCTGAACATATCCTCCTAATCTTACTCAAGCGTGTTGAGTAGTTCCTTCTTTAGTTAGTGGTCTTTATGA
	AATGCTACATCATGAAATGAGAAAGAATCATTGCATAATTTTGTGTATATTTTTCTACACAGGGCTCCCATTGTTGCTGACTGTGATCGCAAAAGGATTCT
	TTCATCTTGGGCTCCTTATTGCGGTAATGTAATCCCAAGGGAATTCCAGATAAAATTTCCAAAGTTCGATCTCTTACGCCTCGTCAGTTTGTGGAGAAG
	CTGTCCAAAGAGCTGGGAGTGCGAGGAGTTGTTGCTGGTATGTTGTTACATCCTAACTTGACATAGAGATTCCAATATAAACTAAGAAAAGAAAAGAAATA
	TATGAGGATGTCTTGATGATCTCAGGCGAAAACTATCGTTTTTGGATGATGCATCGGACCTTGTGAAGCTCTGTAAAGAGTATGGAT
	TAGAGGCTTATATAATCAATTCTGTCATGGACAAGAATCAAATTTCTGGAGACTTAAACTCTCGCGAAAAAAGGAGAGAGGGCAAGTATCATCTACTCG
	TGTTAGGTATGCCCTTGACAAGGGAGACATGAAATATGTGTCAGAGCTGTTAGGTCGCAACCATCGTCTTGTTTACTGATGGAGGACCAAGAAAGATTT
	ATTAGTGAGAAAAATAGGCTGTCAGCTCCAAGTCTTGTTGTTGAATCTTGCACCCAAGGAAGGTCTTTATGAGAATTGTTCAGTTTCGATTGACGAGA
	ATGTTGTACCCTGCAGAGTCACAGTTAATACAACTGATATTCATTTGGAATGCTATGAG
5	>57f5c3e6-41cd-48df-b95b-8207e8a6d89e (sequence:mRNA) 1026 residues [Nitab4.5_0004979:152018-157676 strand]
	+
	[cds]ATGTTGGGTGGGTCTCGCTGTATATCCCAGCACTTAAGAGACACTAATTTTCACCAATTTTCACTTCCACAACTCCCATTTTTACTGCTGCTTC
	TCCACCAATCTCAACTTCTTCACCTACTTCAGATTCAGATTAAGGAATAAAGTTAGTGATTTTGGGTCAGTTGTGTTCAGCCCAATTCTTCAAAT
	ACCCAAAGCTCTTTGCTTTCTCAAACGTTGAGCCAATCTGGAAAAGATGAGGAGCCTCCATCAGAAAGATTGTCCATGCTTGCAGGTGGGATAGTAGCTC

	TGGGGAAGTTTGATGCACTCCATATTGGTCATCGTGAGCTTGCAATCCAAGCAGCTAAGAGAGGAATTCCATTTCTTCTTTCATTTGTTGGAATGGCGGA AGTACTTGGATGGGAACCGAGGGCTCCCATTGTTGCTGACTGTGATCGCAAAAGGATTCTTTCATCTTGGGCTCCTTATTGCGGTAATGTAATCCCAAGG GAATTCCAGATAAAATTTTCCAAAGTTCGATCTCTTACGCCTCGTCAGTTTGTGGAGAAGCTGTCCAAAGAGCTGGGAGTGCGAGGAGTTGTTGCTGGCG AAACTATCGTTTTGGATATAGAGCTTCTGGTGATGCATCGGACCTTGTGAAGCTCTGTAAAGAGTATGGATTAGAGGCTTATATAATCAATTCTGTCAT GGACAAGAATCAAATTTCTGGAGACTTAAACTCTCGCGAAAAAAAGGAGAGGGCAAGTATCATCTACTCGTGTTAGGTATGCCCTTGACAAGGGAGAC ATGAAATATGTGTCAGAGCTGTTAGGTCGCAACCATCGTCTTGTTTACTGATGGAGGACCAAGAAAGATTTATTAGTGAGAAAAATAGGCTGTCAGCTC CAAAGTCTTGTTGTTGAATCTTGCACCCAAGGAAGGTCTTTATGAGAATTGTTCAGTTTCGATTGACGAGAATGTTGTACCCTGCAGAGTCACAGTTAA TACAACTGATATTCATTTGGAATGCTATGAG
6	[Nitab4.5_0004979:152018-157676 + strand] [peptide]MLGGSRCISQHLRDTNFHQFCTSTTPIFTAASPPISTSSFTYFRFRLRNKVSDFGVSCVQPNSSNTQSSLLSQTLSQSGKDEEPPSERLSM LAGGIVALGKFDALHIGHRELAIQAAKRGIPFLLSFVGMAEVLGWEPRAPIVADCDRKRILSSWAPYCGNVIPREFQIKFSKVRSLTPRQFVEKLSKELG VRGVVAGENYRFGYRASGDASDLVKLCKEYGLEAYIINSVMDKNQISGDLNSREKKERGQVSSTRVRYALDKGDMKYVSELLGRNHRLVLLMEDQERFIS EKNRLSAPKSCLLNLAPKEGLYENCSVSIDENVVPCRVTVNTTDIHLECYE
7	**>Nitab4.5_0002045g0010.1 Nitab4.5_0002045:184352-203199** TATGACAGCCTATTGTCCATGCAAGTTTGATATGGGTCGGAAAGTTGCTGGTGGGCCAATCCATTTTCAAGATTGTCAGCTAAGATATGTAAAAATTTCA GGGTTTGGTATTGGGTTTTCATATGGGTATAGGCGTGGGAGGTTGAGATGGGTTCAAAGGAGAAGACAGAAGAAGCTTGTTGTCAGTGCTGAGCTCTCTA ATGCTTTCTCCTCTAACGTTGGATTTGATTCTCAGGTTGCTTTAAATTTTTTCATTTTTCCATTTACCTACTAATGTCATTCTCTCAGTTAAGCTTTGAT TGAATACTTGCCCTCTTATCTTGTTAATTTGTTGGTTGAAAGGGATGTTATCTGAAGCTTATGTTTGATCTTGTTTGAAATGTACTACTGATGCCAAAAT TTGGGTCTTTCTTTATATCCAACCACCTAGCTTAGAAATTATGGGTTCAACTGAACCTAATAGCTTTGACTCAAACCCAGTCCATGTGTTAAAACGTCTA CTAAATAAATACAAATAACAGATTTTGAGCCTAATAAGTCAAATGGGTTGTGATAAAAATTTCGAACTCGCACCCATAAAGTGCAAAATCTTGGCTCCGCCT CTGTGCACTAATGATGACTTTATAGGCCAGTGAGTCTTCAGATAGAGCAGACTTGAGATCGTGTTTCGGAGCATTTGATTCATGTAAGATATGCCAAGGT TGAACCTTGATATGCCCTGGTATAGGAAGTAAAAGTCAGATAGATGAATTCTAACCAGAAGTCCAAGCGATCCAGATCCGATTCTAGCTATTTTGTTTTTG CCCCAATGTGGTGGCATAGAGTATCTGATAGCCTGTTTCACAGTTTTGGTGTCTCTTGTATAGACTGATACGAAAAACAGAGTGGTTCTTGCTTGAAACT CGACGCGAGAAAACATCTTAGTTGATTAATTTAAATTTTAAGCAAACAACATTTGCACACAATTCTTTCAACTATGTTATGTTCTCATTACCTTTTTTA GTTGTTTCTAATTAGTTCTTAATTTGAAAGTTTTGCGTTTTGTTTCTTATGGTTTGTTTACTTGATCCGGGATGGTTTCTAGTCAAAATTTATGTTTTT ATTTTCTGATATAAGTAACTCAATTAACAAAAAAATGGGACTACCCAGCATACTAGAGGTTATTATGAAGTTCTAGATAAATAGTTTCCAACTATTAATCTC TGCCTGAAAGTTGTTGGTTTGAATTTTTTTGTTATTAGCTGACAGCTTTCCTGAAATTTTGAATTCTCCAGCCCCGTGATATACTGAAGTTACCCTGGATT GGCCCTCTACCTGGCGATATTGCTGAGATAGAGGCTTATTGTAGAATCCTTCGAGCTGCCGAACAACTTCATAACACATTAATGGAGACACTATGCAATC CAGTGACTGGAGAATGTAGTATCTCATATGATGTACCCTCAGAAGATAAACATTTACTGGAGGATAGAATAGTTTCTGTTCTAGGATGTATGGTATGTCT TCTAAACAAAGGGAGGGAGGATGTTCTATCTGGGAGATCCTTTATCATCAACTCATTCAGTGATTTTGATGTGCATGTAATGGATGATAAGCTTCCTCCA CTAGCTTTTTTTCAGAGGTGAAATGAAGAGGTATTGTGAGAGCTTGCATGTTGCTCTTGAAAACTTTATAACACCTGATGATCCTACGAGCATTAATGTAT GGAGAAAGCTGCAAAGACTGAAGAATGTATGTTACGATTCTGGATTTCCCCGTGGAGATGACCATCCCTACCATACATTGTTTGCTAATTGGAATCCTGT TTATTTTTCATCAGAAGAAGAGACAGAATCTGCAAGTTCTGAGGTTGCTTTTTGGACAGGTGGACAGGTAACTGAAGAAGGTCTTAGATGGCTGTTGGAG AGAGGATTCAAACTATTATAGATCTCAGAGCTGAGACTATAAAGGACAACTTCTATGAAAAATTGCTAGATGAAGCCATTTCCTCTGGAGATATTGAAG

	TACTGAAACTCCCTGTTGAAGTTGGCACAACACCTTCAGTGCAGCAGGTTGAGAAGTTTGCAGCACTGGTCTCTGATGTATACAAAAGACCCATATTTCT
	CCACAGTAAGGAAGGAGTTTGGAGGACATCAGCTATGGTCTCCAGATGGAGACAATACATGACTCGCTACACATCACACTTTGTGCCGAATGCTAGTAAG
	GATGTGACCTCCAGTGTGAACTCATTTTGTGGTAGTAGAGGAACACAGGAGGCAGGCACGCCAGTTAACTCAGAAGAAAATAAAACTTCTACTTGTGAAG
	GTGTGTCTGCATCTGATCACAAAAATGGGACACTACCTACAAGATCAAACAGCATAAATTCTGCTGGGAAACTTTTCAAACAAATTCCTGAAGCTATAGA
	AAACAAAGACCTAAGTAAGAATGAAGCTGATGATACTGTTGAATTCACTTGGAAAGGCACATTGCTTACTGCTGATGGTGGGGTAGTATCCTATAATAAA
	ACAAACCCTCTTAAGTCGCAACTGCCTCCCCCTAAATTTTTCTCCCGGACTGAGATGTCTACATATTTCAGGAGTAGAAAGGTTTCACCAGAGACATATT
	TCACTCATGGAAAGAAAAGATTGGAGGGGCTCCATGCTTCGAGGTATTACTACAAGAGAGTACCTAAAGGGAATGAAATCATAGACAGTTACACTGAGGA
	CAGAGCTATGGATTCTAGAAATCCAAATGGGCCATCTAGTAATATGCGCTTATCAACGAAACCTTCGAATTCCTCTGCGAATATGGAAAAGTATGAGGGT
	CATAATGGCTCTGCAGTGCCAATTTTGAACAGATTCAACAATGGTGAAGTACATACCTCTGTTAAAAGCAGTGGTCTTATTGATGCAAGTAACGAGTTGG
	ATGCTAATGCTGTGTCCTCAGCCACTGCCATTGAGAGGAGGAACGTTGAGGCCCCCAGGCCTTCAGTTGATGATAACATGGAGCTAATTGAAGGAAATAT
	GTGCGCTTCTGCAACTGGTGTGGTAAGATTGCAGTCCAGAAGGAAGGCAGAGATGTTCTTGGTTCGCACAGACGGGTTTTTGTGCAACAGAGAAAAGGTA
	ACAGAAACTTCCTTGGCCTTCACTCATCCTAACACCCAGCAGCAGATGCTTTTATGGAAATCCACGCCAAAGACCGTACTACTTTTAAAGAAGCTGGGAC
	AAGAACTCATGGAAGAAGCCAAAGAGGTGCTTACACAGATCACTTATCGATCTTGGACTCTGATATATGACTACTCTGTTAATGCTGGACTGTCCCTGTT
	GTTTTTCATGAAGATGCATACAGTTACCAATACTTTTAATAGGGCTGTTAATTTACTATAGCAACGCTCACATTTGTGTTTTTTACAATTGAAAGGAAGA
	CCCGCAAAGGACCTTCCTTGAGAGCAAGAGATTTATTTTGGCTCATTTGTAATTTCTGTGCTACACGATAATCATCTTATCTGCTCTGTGAAATCAATGT
	ATACACTTTGAAACTCCAGATTGCGTGATCAAGTAGGCCTGCTTGCTCCCATTTTTTCTTTATCAACAAAAAGAAAAACAGAATATGCTCGACATATGTT
	CCTGCATCATACTAGCGCGACCTCTCAGCATATTCAAACAAATTACAAAAGCTGCACTTCGTTTGTGAGTGGACATTTAGTTATTGTCAATCATCTGTAA
	TATGAATTCCATCTACTTTCTGTCCAATTTTGTATCCACCAACCAACTTGAGAAAGGAATATATGACCATTAGTTTTCTCTACTCATTTCTTCTTTGCCA
	TTGACTTCCATCCCTCTTACATCATGAATTTCAGTGTCATGCAAGTCTCTACCTAATTTTGCTGAAGTGAATCTTGGTTAGTGTTGACATACAAGTTGTC
	TAGCACATGTATACACATGTACTACATCATAGCTAGCACATGTATATGAGTTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN

	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNGTTTCTAACAGTTAGCT TAGGCCTTGAGATTTTTTCCCATTTGTTTCATTTTGAGAACCTTTAGTTTCAGCTTGTTGTGCATAGCTCTATGATCCATTTATGAAACTTATCTTTGAA GGTTCCATATTTAAGAAAATATGTAGTTCTGGAAAATTCGCTATTTTCTTTTGCAACATGCTTATGCAGTCGAAAACTATGTCCTTTTCGTTCTTTAGAT AAGTATCTCTAGCAAATAGTTTACTATCTCATAGTTTTAACTCTTTTGCTTTTTTCTTTCTTTCTAATTCTTCTTTGAAGGTTGCTTCTTTCTTGTATTA CCAAGAGAAAATGAAAGTTCTTGTTGAACCTGAGGTGCATGATATATTTGCACGAACTCCAGGCTTCGGATTTGTTCAGACCTTTTATTGTCAAGATACC AGGTGAAATGACCCACAAACTTTACACTCAGATAGCACTCTTCATATATTTGTTTGTTTGTCTCGGTTCACTTCTTTTACTTCACTAAGATATGCTGACC TTGCCTGTTGTACCTAATGCAGTGATCTTCATGAAAGTGTTGATTTTGTTGCCTGCCTAGGCGGAGATGGAGTTATACTCCATGCTTCTAAATTATTTCG AGGTGCTATCCCGCCCGTTGTGTCATTTAATTTAGGATCCCTTGGATTTCTCACTTCCCATACAGTATGCTTCTTGACTGAAAACATTATGATTTGGTTC TTAATTCACTTAGGAAGTCTCTCCTCTGACTTTTTCTTATTAATCCGTAAAGTTTGAAGATTATAAGAAGGATCTAAGACAAGTCATTCACGGGAATAGC ACTCTGGATGGTGTTTATATAACTCTGAGAATGCGTCTTCGATGTGAGATATTCAGAAATGGAAAGGCAATGCCTGGAAAGGTGTTCGATGTCCTAAATG AAGTTGTAGTTGATCGTGGTTCTAATCCATATCTCTCCAAAATTGAGTGTTATGAACATGATCGCCTCATTACCAAGGTGCACTTTTTGTAAACCTAACT GCCAACCACCCCGCTGCCCCTGCAGAAGAACCAAAAATTATCCCATTTTTCATATTTCAGACTTCGTAGACCTTCCTAGCCAACTTTTTTATTTCAACAA GCATCCACTTGATTAAGGAGCTATAGGACATGGAGATACCTGTGACACCATTTTATGTTATTTGGCAGGTGCAAGCCGATGGGATCATAGTAGCCACACC AACTGGAAGTACTGCTTATTCTACGGCTGCTGGAGGTTCCATGGTACCTATTATCCCTAAGAGATATGAACCTTTTAACTTACATGAGTAATTGAACATA TGTGTCCTGAAATACCACAGGCAATGAATTGTATCCAAAGCAAATATTCCCTCCGTTTCAATTTATGTGACATAGTTTGACTTTGGCACAGAGATACGAA AAAAAAGAGAAAAAGAAAAAGAAGACTTTTTAAATAGTGGTCTTAAAAGCTTAAAGGGTAAAAGCTTTGTGGGGCCATAACATTTGTGTCGTTATAAAAG ATTCTCATTAAGGCTAAACTGGGTAAAATGAAAAGTTTAAAGTTAAATTGTTTCCAAATATATAAACATGCCATTCTTTTTGAAACGGACTAATAAGGAA AGTGTGTCACATAAATTGAAACCGAGGTAGTAATAAGTTAAAGCTGAATTACAGAAATAGTGCAAGATGGTGAGGTTTATAGCTTTAACAGTTTGTTGTG GCATTGGTGAACAGCTAAAAAACGAGCCACAGAGAAGTGGCTGACTACCCCCTCCCCTCTCATTAAGAAATAAGAGAAGAAGAGACAGGAAAAGAAAGTA GACTGCATTGTGCACTTAATAGTAAAGATATAGACGTTTCCTGACATATTTTACGAGTTATTGGTAGAAGCTAGTTGATTCCTCCTGTCATATCGTTATT ATTGATGTGGAAATGGGGTGATGCAACAGACACTTGAATTGAAATAGATGGAAAGGTTACTTTCGAAGATATTGGGGAAAACTGCTAAGCTAGTCAAATT TGGACTCCTGATTGAAAACCTACTATCGGGGTTTTTCTTTAAAGGTTGAAAGCCTAATATTTAGCATTAAACAAAACTGATAGCAAGACTTCTTTGCATG GTTCTGTATTCAGTGGTGATTATGTTTATGATCCTTAGCTGTGTTACATAATCTGATACTGAACAATTGATTTATTTTTTTCCACTTCCTATTTGTTATT

TACTGAAAGTAATAAATGCATAATCTTGTTTCCATTTCTATTTATTGTATGACTGAGAATAGTGGAGATCGTAAAATTGCTTCATAAAATGTCACTCTA
AGTGTAGATGTGTGCAACAAAGCAATAACTTGCAGTAATCAATCAGGTTATCAGCAAGTATGAACCAATTTTTGGCATCGAAAGCAAATTGACGGAACTC
AATATTAAACAAATTAGTTTCAGTGTACTACGAAGCAGTTTTAGCTAGTTGTGTGAAGATTTCTTTCTGATTTCTTTCGTTCTACTTAGTTGCAATCTTG
AAGTTACATCTTGCAATAAGATGGTTTTGTTTATTTGATTTCATACACAGTTCTTGATCTATATTTTTACTTTTGAAAGTTATGTTGGTTGTTATTGTTC
ATTTAATTGATGCTGAAACTTTAGCTGTGCTCCTTCTGGTCAATAATTGTTGATTCGCCTTTACCAAGTGGATGCTAGAAATAGTATATCACAAGGA
AATTTATGTCATGCAGGTGCACCCAAATGTTCCTTGCATGCTCTCACACCAATCTCGCTCCCACACTCGCTTCAGACCTGTCATTCTTCCGGACTCT
GCAAAGCTGGAGCTAAAGGTGACTTTCTCCATGAATATTTAAATAGTAATGGTATATATTTACTTTTAACTACTAGTGCTAGTACCTTGGTACAGTGTGT
TAATACTCTTGGGGGCTTCTTTAGTAAATTTTGAAAGATCATCTTAGACAATGGAATAAGTTATTGTCCGGAAGGGTCGCACCCAAGGGTGTGGCCTAAC
GGTCAGTTAAGATATTGTCTGGAAGGATAGCACCCAAAGGTGTGGCCTAGTGGTCAATGACGTAGGTTGAGAACCATGAGGTCTCAGTTTCAAATCCTAG
CAGAGGCAAAAACACTAGGTGATTTCTTCTCATGTGTTGTCCATGGACAAAGTTGCCCGATACATGTGGCTGGTGGGAGGTAGCAGGTACGC
TGTGAAACTAGTCAAGGTGCACGCAAGCTCGGCTGGATACCTCGATTCTCAAGAAAAAGTTAGTGTCTGGAAGGGTAGTTAATTGATTTACTTTCAAGGT
TTTGCACCTCTCAAGCACAAGAAGAGAGAAAAACAAAAGCACCCCGCTAATGTTTTTCACACAAACTAGCGAAAGCGTTCTTCAAGCTCACGTTAGAATG
TAAGATGAAAAAAAAACCTCACCCACACGTTCTTCAATCCCGCAAGATTGCAAGAAGCGTTGAATAAAGGAAGAACAATGAAATAAAAGTAAATCACGTT
TTTGAAGCAAGAACTTAGAAAAACTAATCACTAATGTTCAGCATATTAAGTCCTTTAAATAGGCCTTACACAATGAGCAAAATTGATAAAGTTAAGTAAAAT
TAATCTTAATTAAACTAGAATAAAAATAAGGCAAAGAGAACCTATCCTAACTAAAATAGGAAAAAGAATTCTAGTTGGAATGGACTACCAACTAACAATC
CTAGTTTGACTAGAACTATATAAATATATATAATCCTACTATTGTAATCTTCAAGGAAAAGAAATCCTAATCTTGAATGGATTCTTGGACTTCTTCAATTTC
TTGAAATTCTTGGGCCTCATTGGATTAAGTCTAATTAGTGTAAGCATTTTAAATTCGTGGCCTAATTCTCCAAATTCATTGGCACCTTCATGAGCCCAAT
AAGCCTCCAAATGTTGTAGAAAACATGGCCTCCATAGGTTCATCATTCTCCCTTGGTTGAAAAGGACTCGCCCTCGAGTCATCGAAATAGTAGGTGATGCT
TCAAAAAGTCATTGGGAAGCTCCAAGCGATAAGCATTATCACCAATCTTTTGCAGTACCTTAAAGGTCAATTAGCTCGCGTTTGTAATGTAGAGAAGTA
GTGGGGATGCAAACACTACGTGGTCGACGACGGCGGACTGTATAAGGAGGGCGGCGAGAGAGGTGTTAGGGATATCTTTGGGCCGCACCGGTGGCCACAT
AGGATACTGGTGGTGGAATGCAGTTGTCCAAGGTAAAGTGGAAGCAAAGAAGGCGGCTTACCTGCGGTTAGTAGGGAGCACTGGCGAGGAGGAAGAGA
GCGAACATGGAGAGGTATAAGGTAGCTAGGAAGAGGCGAAGATGGCAGTAACGGAGGCTAAGACGGCATCTTTTGCTCGTCTGTATGAGGAATTAGGGA
ACAAAAGCGGGGAGAAGAAGTTATTTCGACTCGCTAAGGTGAGAGAGAAGACGACTCGAGATCTGGACCAAGTGAGGTGCATAAAAAATGATGGCAGCAA
AGTTTTGATGGGGGATGACCAGATTAAGAGGAGGTGGCAGACCTACTTTGCTATTGTACTAGGTGAATTGAGGAATGCCGATAGTCCCTATGAATTAAGT
TATTGTAGGGACATTGAGGTCATGGAGGCAATGCGTAAGATGAGAAGGCAGAGCTAGACGAAATTTCGGTTGAACTTTGGAGGTGTGTAGGTAGAGCA
GGTGTGGAATGGCTTACTGGGTTGTTTAATATTCAAGATTAATAGGTTGCCTGAAGAGTGGAGGTGGAGTACAATGGTTCCGTTGTATAAGAACA
AAGGCGATATCCAGAGCTGTAACAACTATAGGGTTATCAAATTACTGAGTCATACCATGAAAGTCTGGGAGAGAGTGGTAGAAATGAGAGTGCGAAGGAC
GGTGTCTATTTCAGACAAGCAGTTCGGGTTCATGCCGGGGCGATCTACCACAGAAGCTATCCACCTTATTAGGAGGATGGTGGAACAATATAGGGATAAG
AAGAAGGATCTCCACATGGTGTTTATTGATCTAGAGAAAGCGTACGACAGGGTTCTTAGGGAGGTCTTATGGAGTCCCTTAGAGGTTAAAGGGGTCCCGG
TTGCCTACATTAGGGTGATTAAAGACATGTATGATGGAGCTAAGACTCAGGTTAGGATAGTAGGAGGCAACTTCGAACATTTTCCTGTTGTTACGGGGTT
GCACCAAGGGTCTGCGTTCAGCCCTTTCCTATTTGCCCCGGTGATGGATGCACTAACTCATCATATTCAAGGGGAGGTGCCATGGTGCATGCTATTTGCT
GATAACATAGTTCTAATTGACGAGACACAAGGCGGCGTCAACGAGAGGCTAGAGGTTTGGAGACATGCCCTTGAGTCTAAAGGTTTCAGGTTGAGCACGA
CGAAGACGGAATACCTCGAGTGCAAATTTGGGGCCGAGCCGACGGAAGCGGGAGTGGAAGTGAGGCTTGACTCTCAAGTCATTCCTAAGAGGGGTAGTTT
CAAGTACCTTGGGTCAGTTATTCAGGGTATCGGAGAGATCGACGAGGATGTCACACACCGTATAGGGGTGGGTGGATGAAGTGGAGGTTTAGCGTCGGGA

	GTCCTGTGTGACAAGAAAGTGCCACTGTTACTAAAGGTACTCCCTCTGTTCCAATTTATGTGAACCTGTTTGACTGGGCACGGAGTTTATGAAAAAATGA AAACTTTTAGAATTTGTGGTCCTAAACAAATCAAAAAGGGGTCCAGGGTATTTGTGTGGTTATAAAAGCTTCTCATTAAGGGTAGAATTGTAAGTTTAAG CTAAATTGCTTCCAAATTTAGAAAGGGGTCATTCTGTTTGGAACGGACCAAAAAGAAATAGGTTCACATAAATTGGAACGGAGGGAGTAAGTTTTATAGA GCAGTGGTTAGGTCGGCCATGTTGTATGGGATCGAGTGTTGGCCGATTAAGAACTCACACATCCAGAAGATGAAAGTAGCAGAAATGAGGATGTTGAGGT GGATGTGCGGGCATACAAGTATGGATAAGATTAGGAATGAAGATATTCGAGAGAAGGTGTTCGGGCACATTCAGAGGAAGAGCACTGATGCACCGGTGAG GAGGTGTGGGCGACTGGCTGTGGTGGGCACGAGGAGAGATAGAGGGAGACCTAAGAAGTATTGGGGAGAGANNNNNNNNNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNGNTGTTGTTGTTGTTGTTGTAATGTAAAAAACTTGCGGTTTGGAAATCTCTCCT TACTTAAATGAACCCAAACCAGGTCGCCTTCCTTAAACACCACATGTTTCCTACGTATATCAGCCTTTTCTTTATATCTAGCATTTTGCTTCTTGATATG TGACCGAACTTTCTCATGTAGTTTCTTGAGTTGCTTCACTCGATCAGCAACATCTCGGCTAAAAAAATGAGTGGCTAGAAGAGGAGAAAGATCAAGTGGT TCAAGAGGATTCTTACCATAGACAACTTCGAATGGACTCTTTTCCGTAGTGTGGTTGGTTGACTTGTTGAAGGCAAACTCATCAGTAGGAAGTAACGATT CCCACTCTTTGAGATTCTTCTTTATTAACCCACGAAGCAAAGCACCTAAACTTCGATTGACAACTTCGGTTTGCCCATGATGAGAAGTGCTAGGTGTTAG CGTAGTACCAAGTTTAGCCCATAAAGTCCTCCAAATATGCCCTAGGAACCTTGAATCCCTATCTGAAACAATAGTTCTTGGAATTCCATGCAATTTCACA ATATGATTAAAAAAAAAAGATCAGCAATTCTTATCGCAAGGAATAAAGTGAGCCATCTTTGAAATTCGATCGACAACCACGAAGATGGAATCCTTGCCTC TTTGATTCCTTGGAAGGCCCAACAAAAAGTCCATGCTAATGTCTTCCTAAAGTTGAGTAGGAATTGGCAAAGTAGTGTACAGACCCGAATTGCCTTTGTG CACTTTTGCCTTTATACAAGTTTCACATCTTGCAACCAGCTTATTGACATCCCGAACCATCTTAGGCCAATAGAAGTTTTCATGAACTAGGCGTAGAGTG TTTGTTTGCCCAAAGTAGCCAGCTAACCCTCCCTCATGGGACTCACGGATTATAAATTCACGCTAACTTCCTTGTGGAACACAAAGCTGCTAACCATGAT AAAGTTAGCCATCCCTAAGCACAAACATACGCTGAGGACCCTCCCAGCATTGACCCCAAATAGCTCCAAAGTCTGGATCTGTTGCATATAATTCCTTTGT GTTCAAAGCCCACAATAGTGTTTTGCATGACAGAAACAAATCAAGGCTTGCGACTTAATGCATTAGCTACTTGATTGACTATCCTAGCCTTGTGTTTAAT AGCAAAGTGAAAGGCTTGCAAAAATTCTACCCACTTGGCATGCCTAATGTTGAGTTTATGTTGGCTATTCACATACTTCAAGACCTCATGATCAGAATAA AGCACAAATTCTCGAGGTAAAAGATAATGACTCCAATGAGAAAAAGCTCTGATAATGGTGTAGAACTCTTTGTCATAAGTAGAATACTTGGAACTTTGCA

	TCATTCAGCTTCTCACTAAAATAAGCAATGGGTCTACCTTCTTGACTAAGAACAACTCTCATCCCAACTACGGAAGCATCACATTCAACTTCAACGATTA
	CATCAAAGTTGGGCAAAGCAAGCGCATGTGCTTGAGTCATCTTTGTCTTAATCAATTCAAAACTCTTTGGGCAGCATCCGTCCATTTAAATGAGTTTCC
	CTTCAAACACTCGGTAATATGAGCAACAAGAGAGCTAAAATTCTTTATAAAACGCCGATAAAAGACACCATACCATGGAAGGTGCGAATCTCTATTAAGC
	TTTTTGGAATAGGCCAGCTAGCAATAGCTTCAATCTTTTCCGAGTCGACCTGAATTTCCTCAATAGAGACAATATAACCGAGAAAGAGACTTGCTCAGAA
	AAGAACTTGCACTTCTTCTTCAAGTTGGCATAAAGGGTTTGCTCTCTTAAAACTTCAAAAATCTTTCTCAAGTGCTTCAAATGTTGTTGATCATCCTTGCTAT
	AAATAAGTACATCATCAAAATAAACACAACGACAAAATTTCCTATATATGGCTGACAGATTTGGGTCATGAGTCTCATAAATGTGCTAGGAGCGTTAGAGAG
	CCCAAATGACATAACAAGCCATTCGTACAATCCCTTGCCTGTCTTCGAATGCAGTCTTCCATCATCTCCCTTCTTCATTCGAATTTGGGGATAGCCACTT
	TTCAAATCAATTTTAGAAAATATCTTGCCCTCATGTAATTGGTCAAAAGATCATCAAACCTAGGAATAAGAAATCTATACTTGATGGTGATCTTGTTAA
	TAGCACGACTATCAACACACATGCGCCAAAATCCATATTTCTTAGGAACGAGCAAGGCAGGTACAACACAAGGACTTAAACTTTCCCTTATTACTCCACT
	CCTTTATAAGTAATTCTTCCTCTCTTGTTATGCATTTTTATGGACTTTAGGATTAACGCATTATGCTGACTTATTTGGAATGATAGAGTCAACAATCAAG
	TCAATATCGTGTTGAGCTTCTATGAAAGGAGGAAGGTCAATGACACACATTGTTGGATACTCCAATGGATCATCATAAACTTTGGTGGTAAATTCTCAT
	CTTGTAGATCCCAATGCATGTTCCTCGTCTCTTTGGTAAATCCTCATCTTGTGGATCCAATATCATTAATGTTGGATGTTAATGAACATGTTGGGAGATAT
	AGAGACGAGAGAGGGAGAGGTGAGCCAGATGAGGGTGAAAAAATGAGGAAGATGAAAGTGAACGAGAATTAGCTATGATGGCTATGATACAAAATCTAG
	TTGAAGAAGAAATCTAGGGTATTCGGTTGAGAGAGGGGAATTGTTATTCAAAATTCTTCCATTGACAAGCTATAGTGAATAAGAAGAGAAACTTATAGAG
	GCAATAGAAGTTAAGGAATAATAATGTACAAGGAGAAGTAGTGCATGAGGCGATTGTGGATTTATACATACAATTATACAGGGAAGGGAGTGCTTTAGGG
	GTCGTTTGGTATCATGGATAAAGAAAATATAATCCCATGATAAAATGTGGGATTATTTAGTCTTGTTGTTTTTTTAAAAATCTTCATCAAAAAGAAAACGAAA
	ATGAAAGAGAAAAAAAGAGGACAATACATGGCATAGTCCATGTTCTCTTTGTTGTTATCGTCTTGAGTCTATGACACCCCTTTATTTAGTACCTTTCTTTT
	CTTTATGATAACAACAATACCACCTCAATTCCAAACAAGTTGGGGTCGCTATATGAATCCTCACTGACCATGTTCTCCATTTAAATTCATCTCAGGCCAA
	TATTGTACAAAATAAAAATAAAAATGAGAAGTACTAGAAGTTCTCGATATTTCCTACTAGCATAAGAATCTGTAGTAAGACTAGAAAACTCCTAGAAAAC
	ATAGGCCTAAAGTAAACGTAATAACGTGATATTCCAGTGAATGGTAAGTGAAGGTGATAAAAGGAGACGAGATAGGCCTAAAACACATGAAAGAAGTTGT
	CTCAAAAAACTTACAATCTCTTGGAATCAATGCAACTTACTTCCATGTGATTTTAGGTCTACCTCGTCTCTTTTTAACACCTTCACTCTCCATGGTTTAA
	CACCCACAGTGTACATACGAGTGCATGTGGAGGTCAACACAAGACATGACCAAACCATGTAAAGTGACATTCTCTGATTTTATCCTCATGCTACTTGTTT
	CATCCGGCGAATGTAGTTATTTTTTAATCTTGTCTAATCTTATTTGAATGCACATCCATCTTTAGGTTTCGGATCTTTGAGATCTTTGTGGATGTGT
	TGGACTTAGCAGCCCAATATTCACATAAATATAACATTGCTGGTCTAATAGAACTTACTTTCACTTTGATATGTATCCTCCTATACATA
	ACACCTCAGTAGCACAAAATGTTCGTCGACAACAGGATTGGAACCTGCACAGGAAAATCCCAACAGATTTCAAGTCTTCTCCTTAACTGCTCGAACATAT
	AGGCAGCTTCTTTATAAGTTCTTATCTTTCCTCATTAAACAACGATGTTTTCTATTTGTAGAATGTGAGTCCTTTTATTCTGCTACTTCTGATTTTGGAC
	TTCTCTTCTACTCGTCATGAAGTTACTAAAACAAATTAAATTTGAAGTTGTGCCTGATTTTGATAGGTCATCTTAGTATATGTTCTTGTCAGTATTCTTAA
	AAGTAGACATGATTCCTCCAATGAATTTAGGTTTCTTAGGATGCTTTAAGTTTTTCCCTAAAAGAGTAATGTCGGTGCAAAAGCCTTAATTTCAACTATCG
	TGTTTTATGATTGGCACCTTGATGACTGCACGTATCCAATTTCTTGATATTACACCATTATTGTTCGACTATTGGCACCTTGATGAATGCATATCCATTT
	CTTGTTGTCAAATCATGAAATCTCTCTCATGATTTCATGGCATGTGTCTAGCAGTCAGCCAAGGCACAACTGCGTCTTTCTAAACGTTTTGCTCTTTTGTGG
	ATTCATTTCGTACCATTCAGTTAGAAAGTTAAGTTATTTGCAAACAGATTCCAGAGGATGCACGCAACAATGCTTGGGTCTCCTTTGATGGGAAGAGACG
	GCAACAACTGTCTAGAGGAGATTCTATTCGGATATGTATGAGTCAGTCAATAAATGTGACCAAACAGGTGATTGGTTTGGTAGC
	TTGATCCGTTGCCTGAATTGGAATGAAAGACTAGACCAGAAAGCACTC
8	>9c3ab348-2705-42b3-b4f8-31f1a0ec59f9 (sequence:mRNA) 3057 residues [Nitab4.5_0002045:184352-203199

	+ strand] [cds]ATGACAGCCTATTGTCCATGCAAGTTTGATATGGGTCGGAAAGTTGCTGGTGGGCCAATCCATTTTCAAGATTGTCAGCTAAGATATGTAAAAAT TTCAGGGTTTGGTATTGGGTATTTTCATATGGGTATAGGCGTGGGAGGTTGAGATGGGTTCAAAGGAGAAGACAGAAGAAGCTTGTTGTCAGTGCTGAGCTC TCTAATGCTTTCTCCTCTAACGTTGGATTTGATTCTCAGCCCCGTGATATACTGAAGTTACCCTGGATTACCTGGCGATATTGCTGAGATAG AGGCTTATTGTAGAATCCTTCGAGCTGCCGAACACTTCATAACACATTAATGGAGACACTATGCAATCCAGTGACTGGAGAATGTAGTATCTCATATGA TGTACCCTCAGAAGATAAACATTTACTGGAGGATAGAATAGTTTCTGTTCTAGGATGTATGGTATGTCTTCTAAACAAAGGGAGGGAGGATGTTCTATCT GGGAGATCCTTTATCATCAACTCATTCAGTGATTTTGATGTGCATGTAATGGATGATAAGCTTCCTCCACTAGCTTTTTTTCAGAGGTGAAATGAAGAGGT ATTGTGAGAGCTTGCATGTTGCTCTTGAAAACTTTATAACACCTGATGATCCTACGAGCATTAATGTATGGAGAAAGCTGCAAAGACTGAAGAATGTATG TTACGATTCTGGATTTCCCCGTGGAGATGACCATCCCTACCATACATTGTTTGCTAATTGGAATCCTGTTTATTTTTCATCAGAAGAAGAGACAGAATCT GCAAGTTCTGAGGTTGCTTTTTGGACAGGTGGACAGGTGAACTGAAGAAGGTCTTAGATGGCTGTTGGAGAGAGGATTCAAAACTATTATAGATCTCAGAG CTGAGACTATAAAGGACAACTTCTATGAAAAATTGCTAGATGAAGCCATTTCCTCTGGAGATATTGAAGTACTGAAACTCCCTGTTGAAGTTGGCACAAC ACCTTCAGTGCAGCAGGTTGAGAAGTTTGCAGCACTGGTCTCTGATGTATACAAAAGACCCATATTTCTCCACAGTAAGGAAGGAGTTTGGAGGACATCA GCTATGGTCTCCAGATGGAGACAATACATGACTCGCTACACATCACACTTTGTGCCGAATGCTAGTAAGGATGTGACCTCCAGTGTGAACTCATTTTGTG GTAGTAGAGGAACACAGGAGGCAGGCACGCCAGTTAACTCAGAAGAAAAATAAAACTTCTACTTGTGAAGGTGTGTCTGCATCTGATCACACAAAAATGGGAC ACTACCTACAAGATCAAACAGCATAAATTCTGCTGGGAAACTTTTCAAACACAAATTCCTGAAGCTATAGAAACACAAAGACCTAAGTAAGAATGAAGCTGAT GATACTGTTGAATTCACTTGGAAAGGCACATTGCTTACTGCTGATGGTGGGGTAGTATCCTATAATAAACACAAACCCTCTTAAGTCGCAACTGCCTCCCC CTAAATTTTTCTCCCGGACTGAGATGTCTACATATTTCAGGAGTAGAAGGTTTCACCAGAGACATATTTCACTCATGGAAAGAAAAGATTGGAGGGGCT CCATGCTTCGAGGTATTACTACAAGAGAGTACCTAAAGGGAATGAAATCATAGACAGTTACACTGAGGACAGAGCTATGGATTCTAGAAATCCAAATGGG CCATCTAGTAATATGCGCTTATCAACGAAACCTTCGAATTCCTCTGCGAATATGGAAAAGTATGAGGGTCATAATGGCTCTGCAGTGCCAATTTTGAACA GATTCAACAATGGTGAAGTACATACCTCTGTTAAAAGCAGTGGTCTTATTGATGCAAGTAACGAGTTGGATGCTAATGCTGTGTCCTCAGCCACTGCCAT TGAGAGGAGGAACGTTGAGGCCCCCAGGCCTTCAGTTGATGATAACATGGAGCTAATTGAAGGAAATATGTGCGCTTCTGCAACTGGTGTGGTAAGATTG CAGTCCAGAAGGAAGGCAGAGATGTTCTTGGTTCGCACAGACGGGTTTTTGTGCAACAGAGAAAAGGTAACAGAAACTTCCTTGGCCTTCACTCATCCTA ACACCCAGCAGATGCTTTTATGGAAATCCACGCCAAAGACCGTACTACTTTTAAAGAAGCTGGGACAAGAACTCATGGAAGAAGCCAAAGAGGTTGC TTCTTTCTTGTATTACCAAGAGAAAATGAAAGTTCTTGTTGAACCTGAGGTGCATGATATATTTGCACGAACTCCAGGCTTCGGATTTGTTCAGACCTTT TATTGTCAAGATACCAGTGATCTTCATGAAAGTGTTGATTTTGTTGCCTGCCTAGGCGGAGATGGAGTTATACTCCATGCTTCTATAAATTATTTCGAGGTG CTATCCCGCCCGTTGTGTCATTTAATTTAGGATCCCTTGGATTTCTCACTTCCCATACATTGAAGATTATAAGAGGATCTAAGACAAGTCATTCACGG GAATAGCACTCTGGATGGTGTTTATATAACTCTGAGAATGCGTCTTCGATGTGAGATATTCAGAAATGGAAAGGCAATGCCTGGAAAGGTGTTCGATGTC CTAAATGAAGTTGTAGTTGATCGTGGTTCTAATCCATATCTCTCCAAAATTGAGTGTTATGAACATGATCGCCTCATTACCAAGGTGCAAGCCGATGGGA TCATAGTAGCCACACCAACTGGAAGTACTGCTTATTCTACGGCTGCTGGAGGTTCCATGGTGCACCCAAATGTTCCTTGCATGCTCTTCACACCAATCTG CCCACACTCGCTCTCATTCAGACCTGTCATTCTTCCGGACTCTGCAAAGCTGGAGCTAAAGATTCCAGAGGATGCACGCAACAATGCTTGGGTCTCCTTT GATGGGAAGAGACGGCAACAACTGTCTAGAGGAGATTCTATTCGATATGTATGAGTCAGCATCCACTTCCTACAGTCAATAAATGTGACCAAACAGGTG ATTGGTTTGGTAGCTTGATCCGTTGCCTGAATTGGAATGAAAGACTAGACCAGAAAGCACTC
9	>9c3ab348-2705-42b3-b4f8-31f1a0ec59f9 (sequence:mRNA) 1019 residues [Nitab4.5_0002045:184352-203199 + strand]

	[peptide]MTAYCPCKFDMGRKVAGGPIHFQDCQLRYVKISGFGIGFSYGYRRGRLRWVQRRRQKKLVVSAELSNAFSSNVGFDSQPRDILKLPWIGPL PGDIAEIEAYCRILRAAEQLHNTLMETLCNPVTGECSISYDVPSEDKHLLEDRIVSVLGCMVCLLNKGREDVLSGRSFIINSFSDFDVHVMDDKLPPLAF FRGEMKRYCESLHVALENFITPDDPTSINVWRKLQRLKNVCYDSGFPRGDDHPYHTLFANWNPVYFSSEEETESASSEVAFWTGGQVTEEGLRWLLERGF KTIIDLRAETIKDNFYEKLLDEAISSGDIEVLKLPVEVGTTPSVQQVEKFAALVSDVYKRPIFLHSKEGVWRTSAMVSRWRQYMTRYTSHFVPNASKDVT SSVNSFCGSRGTQEAGTPVNSEENKTSTCEGVSASDHKNGTLPTRSNSINSAGKLFKQIPEAIENKDLSKNEADDTVEFTWKGTLLTADGGVVSYNKTNP LKSQLPPPKFFSRTEMSTYFRSRKVSPETYFTHGKKRLEGLHASRYYYKRVPKGNEIIDSYTEDRAMDSRNPNGPSSNMRLSTKPSNSSANMEKYEGHNG SAVPILNRFNNGEVHTSVKSSGLIDASNELDANAVSSATAIERRNVEAPRPSVDDNMELIEGNMCASATGVVRLQSRRKAEMFLVRTDGFLCNREKVTET SLAFTHPNTQQMLLWKSTPKTVLLLKKLGQELMEEAKEVASFLYYQEKMKVLVEPEVHDIFARTPGFGFVQTFYCQDTSDLHESVDFVACLGGDGVILLH ASKLFRGAIPPVVSFNLGSLGFLTSHTFEDYKKDLRQVIHGNSTLDGVYITLLRMRLRCEIFRNGKAMPGKVFDVLNEVVVDRGSNPYLSKIECYEHDRLI TKVQADGIIVATPTGSTAYSTAAGGSMVHPNVPCMLFTPICPHSLSFRPVILPDSAKLELKIPEDARNNAWVSFDGKRRQQLSRGDSIRICMSQHPLPTV NKCDQTGDWFGSLIRCLNWNERLDQKAL
10	**>Nitab4.5_0001168g0130.1** **Nitab4.5_0001168:291056-300158** ATGGCAGCCTATTGTCCATGCAAGTTTGATATGGGTCGGAAAGTTGCTGGTGGGCCAATCCATTTTCAAGATTGTCAGCTAAGATATGTAAAAATTTCAG GGTTTGGTATTGGGTTTTCATATGGGTATAGGCGTGGGAGGTTGATGGGTTCAAAGGAGAAGACAGAAGAAGCTTGTTGTCGGTGCTGAGCTCTCTAG TGCTTTCTCCTCTAACGTTGGATTCGATTCTCAGGTTGTTTTTTATTTTTTCATTTACCTACTAGTGTCATTCTCAGTTAAAGCTTTGATTGAATATTT TGCCCTCTTATCTTGTTAATTTGTTGGTTGGAACTGTGCTGACAGAAAGGGATGTTCAACTGAACCTAATAGCTTTGACTCAAACCCAGTCCATGTGTTA AAATGTCTACTAAATAAATACAAATAATAGATTTTGAGCCTAGTAAGTCAAATGGGTTGTGATAAAATTTCGAACTCGAACCCATAAAGTTCAAATCTTG ACTCCCCCTCTGTGCACTAGTGATGACTTTACAAGCCAGTGAGTCTTCAGATAGAGCAGACTTGAGATTGTGTTTCGGAGCATTTGTTTCATGTAAGATAT GCCAAGGTTGAACCTTGATATGTCCTGGTATAGGGAGTAAAAGCCAGAGAGATGAATTCTAAGCAGAAGTGCAAGCTATTATATCCGATTCTAGCTGTTT TGTTTTCGCCCCAATAAGGTGACAGAGATATCTGATAGTTTTGGTGTCTTTTGTATTGACTGATATGAAAAACCAAGTGGTTCTTGC TTGAAACTCGAAGCGAAAAACAGCTTAGTTGATTAATTTAAATTTTAAGCAAACAAACGTTTGCACACAAATCTTTCAACTATGTTATGTTCTCATTACC CTTTTTTGTTGTTTCTAATTAGTTCTTAATTTGAAAGTTTTGCCTTTTGTTTTCGTATGGTTTGTTTACTTGATCCGGGATGGTTTCTAGTCAAAATTTA TGTTTTTATTTTCTGATATAAGTAACTCAATTAACAAAAAATGGGACTACCCCGTATACTAGAGGTATTATGAAGTTCTAGATAAATAGTTTCCGACTAT TAATCTCTGCCTGCAAGTTGTTGGTTTGAACTTTTTGTTATTAGCTGACAGCTTTCCTGAAATTTTGAACTCTCCAGCCCCGTGATATACTGAAGTTACC CTGGATTGGCCCTCTACCTGGCGATATTGCTGAGATAGAGGCTTATTGTAGAATCCTTCGTGCTGCCGAACAACTTCATAACACATTAATGGAGACACTA TGCAATCCAGTGAGTGGAGAATGTAGTATCTCATATGATGTACCCTCAGAAGATAAACATTTACTGGAGGATAGAATAGTTTCTGTTCTAGGATGTATGG TATGTCTTCTAAACAAAGGGAGCGAGGATGTCCTATCTGGGAGATCCTTTATCATCAACTCATTCAGTGATTTTGATGTGCATGTAATGGATGATAAGCT TCCTCCACTAGCTTTTTTCAGAGGTGAAATGAAGAGGTATTGTGAGAGCTTGCATGTTGCTCTTGAAAACTTTATAACACCTGATGATCCTACGAGCATT AATGTATGGAGAAAGCTGCAAAGACTGAAGAATGTATGTTATGATTCTGGATTTCCCCGTGGAGATGACCATCCCCACCATACATTGTTAGCTAATTGGA ATCCTGTTTATTTTTCATCAGAAGAAGAGACACAATCTGCAAGTTCTGAGGTTGCTTTTTGGACAGGTGGACAGGTAACTGAAGAAGGTCTTAGATGGCT GTTGGAGAGAGGATTCAAACTATTATAGATCTCAGAGCTGAGACTATAAGGACAACTTCTATGAAAAAGTGCTAGACGAAGCCATTTCCTCTGGAGAT ATTGAAGTACTGAAACTCCCTGTTGAAGTTGGCACAACACCTTCAGTGCAGCAGGTTGAGAAGTTTGCAGCACTGGTCTCTGATGTATACAAAAGACCCA TATATCTCCACAGTAAGGAAGGAGTTTGGAGGACATCAGCTATGGTCTCTAGATGGAGACAATACATGACTCGCTACACACCACTCTTTGTACCGAATGC TAATAAGGATGTGACCTCCAGTGTGAACTCATTTTGTGGTAGTAGAGGAACACACAAGAGGCAGGCACGCCAGTTAACTCAGAAGAAAATAAAACTTCTACT

TGTGAAGGTATGTCTGCATCTGATCACAAAAATGGGACACTACCTGCAAGATCAAACAGCATAAATTCTGCTGGGAAACTTTCAAACAAATTCCTGAAG
CTAGGGAACACAAAGGCCTAAGTAAGAATGAAGCTGATGATACTGTTGCAGTCACTTGGAAAGGCACATTGCTTACTGCTGATGGTAAAACAAACCCTCT
CAAGTCGCAACTGCCTCCCCCTAAATTTTTCTCCCGAACTGAGATGTCTACATATTTCAGGAGTAGAAAGGTTTCACCAGAGACATATTTCACTCATGAA
AAGAAAAGATTGGAGGGGCTCCATGCTTCGAGGTATTACTACAAGAGAAATACCTAAAGGGAATGCAATCATAGACACAGTTACACTGAGGACAGAGCTATCG
ATTCTAGAAATCCAAATGGCCACCTAGTAATATGGGCTTATCAACGAAACCTTCGAATTCCTCTGCGAATATGGAAAAGTATGGGGTCATAATGGCTC
TGCAGCGCCAATTTTGAATAGATTCAACAATGGTGAAGTACATACCTCTGTTAAAGCAGTAGTCTTATTGATGCAAGTAACGAGTTGGATACTAATGCT
GTGTCCTCAGCCACTGCCATTGAGAGGAGGAACATTGAGGCCCCCAGGCCTTCAGTTGATGATAACATGGAGCTAATTGAAGGAAATATGTGCGCTTCTG
CAACTGGTGTGGTAAGATTGCAGTCCAGAAGGAAGGCAGAGATGTTCTTGGTTCGCACAGATGGGTTTTTGTGCAACAGAGAAAAGGTAACAGAAACTTC
CTTGGCCTTCACTCATCCTAGCACCCAGCAGCAGATGCTTTTATGGAAATCCACGCCAAAGACCGTACTACTACTTTTAAAGAAGCTGGGACAAGAACTCATG
GAAGAAGCCAAAGAGGTGCTTACACAGTTATCACTTATCGATCTTGGACTCTGATATATGACTACTCTGTTAATGCTGGACTCTCCCTGTTGTTTTTCAT
GAGGATGCATACAGTTACCAATAGGTTTAATTTACTATAGCAAGGCTCACATTTGTGTTTTTTACAATTGAAGGAAGACCTGCAAAG
GACCTTCCTTGAGAGCAAGACATTTATTTTGGCTCATTTGAAATTTCTGTGCTACACGATAATCATCTTATCTGCTTTGTGAAATCTATGTATATACTTT
GAAACTCCAGATTGAGTGATCAAGTAGGCCTACTTGCTCCTGTTTTTTCTTTATCAACAAAAAGAAAAAAGAATATGCTCGACATATGTTCCTCCATCA
TACTAGCCCGACCTGTCAGCTTATTCAAACAAATTACAAAAGCTGCACTTCGTTTGTGAATGGACATTTAGTTACTGTCAATCATCTGTAATATGAATTC
CATCTACTTTCTGTCCAATTTTGTATCCACCAACCAACTTGAGAAAGGAATATATGACCATTAGTTTTCTCTACTCATTTTGGATCTTCTTTGCCATTGA
CTTTCATCCCTCTTACATCATGAATTTCAGTGTGTCATGCTTTAGGCCTTGAGATTTTTCCCCATTTGTTTCATTTTGAGAACCTTTAGTTTCAGCTTGT
TGTGCATGGCTCTATGATCCATTTATCTTTGAAGGTTCCATATTTAAGAAAATATGTAGTTCTAGCAAATTCGCTATTTTCTTTTGCAACA
TGCTTATGCAGTCGAAAACTATGTCCTTTTCGTTCTTTAGATAAGTATCTCTAGCAAATAGTTTACTATCTCATAATTTTAACTCTTTTGCCTTTTTCTT
TCTTTCTAATTCTTCTTTGAAGGTTGCTTCTTTCTTGTATTACCAAGAGAAAATGAATGTTCTTGTTGAACCTGAGGTGCATGATATTTGCACGAACT
CCAGGCTTCGGATTTGTTCAGACCTTTTATAGTCAAGATACCAGGTGAAATGACCCACAAACTGTACGCTCAGATAGCATTCTTCATATATTTGTTTGTT
TGTCTCGGTTCACTTCTTTTACTTCACTAAGATATGCTGACCTTGCCTGATGTACCTAATGCAGTGATCTTCATGAAAGTGTTGATTTTGTTGCCTGCCT
AGGAGGAGATGGAGTTATACTCCATGCTTCTAAATTATTTCGAGGTGCTATCCCGCCCGTTGTGTCATTTAATCTAGGATCCCTTGGATTTCTCACTTCC
CATACAGTATGCTTCTTGACTGAAAACATTATGATTTAGTTCTTAATTCGCTTGGGAAGTCTCTCCTCTTACTTTTCTTATTAATCCGTAAAGTTCGAA
GATTATAAGAAGGATCTAAGACAAGTCATTCACGGGAATAGCACTCTGGATGGTGTTTATATAACTTTGAGAATGCGTCTTCGATGTGAGATATTCAGAA
ATGGAAAGGCAATGCCTGGAAAGGTGTTCGATGTCCTAAATGAAGTTGTTGTTGATCGTGGTTCTAATCCATATCTCTCCAAAATTGAGTGTTATGAACA
TGATCGCCTCATTACCAAGGTGCGCTTTTTGTAAACCTAACTGCCAACACCCCGCTGTCCTTGCAGAAGAACAAAAAATTATCCCATTTTCACATTTC
AGGCTTCGTAGATCTTCCTAGCCAACTTTTTTATTTCAACAAGCGTCCGCTTGATTAAGGAGCTATAGGACATGGAGATACCTGTGACACCATTTTATGT
TATTTGGCAGGTGCAAGCTGATGGGATCATAGTAGCCACACCAACTGGAAGTACTGCTTATTCTACGGCTGCTGGAGGTTCCATGGTACCTATTATCCCT
AAGAGATATGAACCTTTTTAACTTACATGAGTAATTGAACATATGTGTCCTGAAATACCACAGGCAATGAATTGTATCCAAAGCAAATACTCCCTCCGTT
TCAATTTATGTGACATAGTTGATTCGGCACAGAGATACGAAAAAAAAGAAGAGAAAAAGAAGACTTTTGAAATAGTGGTCTTAAAAAGCTTAA
AGGGTAAAAGTTTTGTGGGGCCATAACATTTGTGTCGCTATAAAAGATTTTCATTAAGTCTAAACTGGGTAAAATGAAAAGTTTAAAGTTTTTC
CAAATATAGAAACGTTCCATTCTTTCTGAAATGGACTAATAAGGAAAGTGTGTCACATAAATTGAAGCGAGGTAGTAATAAGTTAAAAGCTGAATTACA
GAACTAGTGCAAGATGGTGAGGTTTATAGCTTGAACAGTTTGTTGTGGCATTGATGAACAGCTAAAAAAAAAGATCCACAGAGAAGTGATTGACTACCCCC
TCCCCTCTCATTAAGAAATAGGAGAAGAGACAGGAAAAGAAAGTAGACTGTATTGTGCACTTAATCGTAAAGATATAGAAGTTTCCTGACATATTTT

ACGAGTTATTGGTAGAAGCTAGTTGATTCCTCCTGTCATATCGTTACTATTGATGTGGAAATGGGGTGATGCAACAGACACTGGAATTGAAATAGATGGA
AAGGTTACTTTCGAAGATATTGGGGAAAAACTGCTAAGCTAGTCAAATTTGGACTCCTGATTGAAAACCTACTATCGGGGTTTTCTTTAAAGGTTGAAAG
CCTAATATTTAGCATTAAACAAAGCTGATAGCAAGACTTCTTTGCTTGGTTCTGTATTCAGTGGTGATTATTATGATCCTTAGCTGCGTTACATAATCTA
ATACTGATCAATTGATTTATTTTTTTCCACTTCCTATTTATTACTGAAAGTAATAAATGCATAAATCTTGTTTTCCATCTCCTATTTATTGTATGACT
GAGAATAGTGGAGATCGTAAAATTGTTTCACAAAATGTCACTCTAAGTGTAGATGTGTGCAACAAAGCAATAACTTGCAGTAGTCAATCAGGTTATCAGC
AAGTATGAGCCAATTTTTGGCATCGAAAGCAAATTGACGAAACTCCATATTAAACAAATTAGCTTTCAATGTACTGCGAAGCAGTTTTAGCTAGTTGTGT
GAAGATTTCTTTCTGATTTTTTTTCGTTCTACTTAGTTGCAATCTTGCAATAAGATGGTTTTGTTATTTGATCTTTATACACAGTTCT
TGATCTATATTTTTTGAAATTTTGAAAGTTGTGTTGTTTATTTGTTCATTTAATTAATGCTGAAACTCTAGCTGTGCTCCTTCTGGTCAATAAATTGTTGA
TTGTTTCGCCTTTACCAAGTGGATGCTAGAATAGTATATCACAAGGAAATTTATGTCATGCAGGTGCACCCAAATGTTCCTTGCATGCTCTTCACACCAA
TCTGCCCACACTCGCTCTCATTCAGACCTGTCATTCTTCCGGACTCTGCAAAGCTGGAGCTAAAGGTGACTTTCTCCATGAATATTTAAATAGTAATGGT
ATACATTTACTTTTTAACTACTAGTGCTAGTACCTTGATACAGTGTGTTAATACTCTTCTTTAGTAAATTTTGAAAGATCATCTTCTAGACAAATG
GAATAAGTTATTGTCCGGAAGGGTAGCACCCAAGGGTGTCGCCTAGTGGTCAATTAAGTTATTGTCGGGAAGGATAGCACCCAAGGGTGTGGCCTAGTGG
TCAATGATGTGGGTTGAGAACCATGAGGTCTCAAGTTCAAATCTTAACGGAGGCAAAAACACTAGGTGATTTCTTCTCATCTGTCCAAGTGTCGTCCATG
GACAAAGTTGTCCGATACCTGTGGCTAGTGGGAGGTAGCACGCAAGCTGGGTTGGACACCACAATTTTCAAGAAAAAGTTAGTGTCTGGAAG
GGTAGTTAATGTTGGATGTTAATGAACATGTTGGGAGATATAGAAACGAGAGGGAGAGTGAGCCAGATGAGGGTGAAAAAATGAGGAAGATGAAAG
TGAAAGGAGAGTTAGCTATGATGGCTATGATACATAATCTAGTTGAAGAAGAAATCTAGGGTATTTGGTTGAGAGAGGGGAATTCTTATTCAAAATTCTT
CCATTGAGAAGCTATAGCGAATAAGAAGAAACTTATAGAGGCAATAGAAGTTAAGGAATAATAATGTACAAGGAGAAGTAGTGCATGAGGTGATTGAG
GATTTATACATACAATTATACAGGGAAGGAGTGCTTTAGGGGTCGTTTGGTATCATGGTTAAAGAAAATAATCCCATGATAAAAATGTGGGATTATTTA
GTCCTGTGTTTTTTTTTAAAAATCTTCATAAAAAAGAAAACGAAAATGAAAGAGAAAAAAGGAGGACAATACGTGGCATAGTCCATGTTCTCTTTGTTGTT
GTCGTCTTGAGTCTATGACACACCTTTATTTACTACCTTTCTTTCTCTATGATAACAACACCACCTCAATTTCAAACTAGTTGGGGTCCGCTATATG
AATCCTCACTGACCATGTTCTCCATTTAAATTCATCTCAGGCCAGTATTATACAAAATAAAAATGAGAAGTACTAGAAGTTCTTGATATTTCCT
ACTAGCATAAGAATCTGTAATAAGACTAAAAAAACTCCTAGAAACATAGGCCTAAAGTAAACGTAATAGCGTGATATTGCACTGAATGGTAAGTGAAGTT
GTTAAAAGGAGACGAGATGGGCCTAAAACACATGAAAGAAGTTGTCTCAAAAGACTTAAAATCTCTTGGAATCAATGCTAACTTACTTCTATGTGATTTT
AGGTCTACCTCGTCTCTTTTTATCACCTTCACTATCCATGGTTTAACATCCACAGTGTACAGACCAGTGCATGTGGAGGTCAATGCATGACATGACCAAA
CCATGTAAACCGACATTCTCTGATTTTATCCTCATGCTACTTGTTTTATCCAGCGAATGTAGTTATTTTTAATCTTGTCCATTCTTATATGAACGCACAT
CCATCTTAAGTTTCGGATCTTTGAGACACTCATCTTGTGGATGTGTTGGATTTTAGCAGCCCAATATTCACATAAATATAACATGCTGCTCTTATAACT
GTTCAATAGAACTTGCTTTCACTTTGATACGTATCCTCCTATCACATAACACCTCAGTAGCACAAAATGTTCGTCGACACAACAAGATTGGAACCTGCACAG
GAAAATCCCAACAGATTTCAAGTCTTCTCCTTAACTGCTCGAACATATAGGCAGTTTTTTTAATAAGTTCTTATCTTTTCTCATAAAACAACAATCTTTTC
TATTTGTAGAATGTGTGTCCTTTTATTCTGCTACTTCTCTTCTACTCGTGATGAAGTTTACTAAAACAATTAAACTTGAAGTTGTGC
CTGATTTTGATAGGTCATCTTAGAATATGTTCTAGTCAATATTCTTAAAAGCAGACATGATTCCTCAATGAATTTAGGTTTCTTAGGATGCTTTAAGGTT
TTCCCTAATAGAGTAATGTCGGTGCAAAAGCCTGTATTTCAACTATCGTGTTTCATGATTGGCACCTTGATGAGTGCACGTATCCAATTTCTTGATGTTA
CACCATTATTGTTTGACTATTGGCACCTTGATGAATGCATATCCATTTCTTGATGTCAAATCATGAAATCTCTCGTGATTTCATGGCATGTGTCTAGCAG
GCAGCCAAGGCACAACTGCGTCTTTCTAAACTTTTTGCTCTTTTGTGGATTCATTTCATACCATTCAGTTAGAAGGTTAAGTTATTTGCTTATAAACAGA
TTCCAGAGGATGCGCGCAACAATGCTTGGGTCTCCTTTGATGGGAAGAGACGGCAACAACTGTCTAGAGGAGATTCTATTCGGATATGTATGAGTCAGCA

	TCCACTTCCTACAGTCAATAAATGTGACCAAACAGGTGATTGGTTTGGTAGCTTGATCCGTTGCCTGAATTGGAATGAAAGACTAGACCAGAAAGCACTC TGA
11	[Nitab4.5_0001168:291056-300158 + strand] [cdna]ATGGCAGCCTATTGTCCATGCAAGTTTGATATGGGTCGGAAAGTTGCTGGTGGGCCAATCCATTTTCAAGATTGTCAGCTAAGATATGTAAAAA TTTCAGGGTTTGGTATTGGGTTTTCATATGGGTATAGGCGTGGGAGGTTGAGATGGGTTCAAAGGAGAAGACAGAAGAAGCTTGTTGTCGGTGCTGAGCT CTCTAGTGCTTTCTCCTCTAACGTTGGATTCGATTCTCAGCCCCGTGATATACTGAAGTTACCCTGGATTGGCCCTCTACCTGGCGATATTGCTGAGATA GAGGCTTATTGTAGAATCCTTCGTGCTGCCGAACAACTTCATAACACATTAATGGAGACACTATGCAATCCAGTGAGTGGAGAATGTAGTATCTCATATG ATGTACCCTCAGAAGATAAACATTTACTGGAGGATAGAATAGTTTCTGTTCTAGGATGTATGGTATCTTCTAAACACAAAGGGAGCGAGGATGTCCTATC TGGGAGATCCTTTATCATCAACTCATTCAGTGATTTTGATGTGCATGTAATGGATGATAAGCTTCCTCCACTAGCTTTTTTCAGAGGTGAAATGAAGAGG TATTGTGAGAGCTTGCATGTTGCTCTTGAAAACTTTATAACACCTGATGATCCTACGAGCATTAATGTATGGAGAAAGCTGCAAAGACTGAAGAATGTAT GTTATGATTCTGGATTTCCCCGTGGAGATGACCATCCCCACCATACATTGTTAGCTAATTGGAATCCTGTTTTATTTTTCATCAGAAGAAGACACAATC TGCAAGTTCTGAGGTTGCTTTTGGACAGGTGGACAGGTAACTGAAGAAGGTCTTAGATGGCTGTTGGAGAGAGGATTCAAAACTATTATAGATCTCAGA GCTGAGACTATAAAGGACAACTTCTATGAAAAAGTGCTAGACGAAGCCATTTCCTCTGGAGATATTGAAGTACTGAAACTCCCTGTTGAAGTTGGCACAA CACCTTCAGTGCAGCAGGTTGAGAAGTTTGCAGCACTGGTCTCTGATGTATACAAAAGACCCATATATCTCCACAGTAAGGAAGGAGTTTGGAGGACATC AGCTATGGTCTCTAGATGGAGACAATACATGACTCGCTACACACCACTCTTTGTACCGAATGCTAATAAGGATGTGACCTCCAGTGTGAACTCATTTTGT GGTAGTAGAGGAACACAAGAGGCAGGCACGCCAGTTAACTCAGAAGAAAATAAAACTTCTACTTGTGAAGGTATGTCTGCATCTGATCACACAAAATGGGA CACTACCTGCAAGATCAAACAGCATAAATTCTGCTGGGAAACTTTTCAAACAAATTCCTGAAGCTAGGGAACACAAAGGCCTAAGTAAGAATGAAGCTGA TGATACTGTTGCAGTCACTTGGAAAGGCACATTGCTTACTGCTGATGGTAAAACAAACCCTCTCAAGTCGCAACTGCCTCCCCCTAAATTTTTCTCCCGA ACTGAGATGTCTACATATTTCAGGAGTAGAAAGGTTTCACCAGAGACATATTTCACTCATCATGAAAAGAAAAGATTGGAGGGGCTCCATGCTTCGAGGTATT ACTACAAGAGAATACCTAAAGGGAATGCAATCATAGACAGTTACACTGAGGACAGAGCTATCGATTCTAGAAATCCAAATGGGCCACCTAGTAATATGGG CTTATCAACGAAACCTTCGAATTCCTCTGCGAATATGGAAAAGTATGGGGGTCATAATGGCTCTGCAGCGCCAATTTTGAATAGATTCAACAATGGTGAA GTACATACCTCTGTTAAAGCAGTAGTCTTATTGATGCAAGTTGGATACTAATGCTGTGTCCTCAGCCACTGCCATTGAGAGGAGGAACATTG AGGCCCCCAGGCCTTCAGTTGATGATAACATGGAGCTAATTGAAGGAAATATGTGCGCTTCTGCAACTGGTGTGGTAAGATTGCAGTCCAGAAGGAAGGC AGAGATGTTCTTGGTTCGCACAGATGGGTTTTTGTGCAACAGAGAAAAGGTAACAGAAACTTCCTTGGCCTTCACTCATCCTAGCACCCAGCAGCAGATG CTTTATGGAAATCCACGCCAAAGACCGTACTACTTTTAAAGAAGCTGGGACAAGAACTCATGGAAGAAGCCAAAGAGGTTGCTTCTTCTTGTATTACC AAGAGAAAATGAATGTTCTTGTTGAACCTGAGGTGCATGATATATTTGCACGAACTCCAGGCTTCGGATTTGTTCAGACCTTTTATAGTCAAGATACCAG TGATCTTCATGAAAGTGTTGATTTTGTTGCCTGCCTAGGAGGAGATGGAGTTATACTCCATGCTTCTAAATTATTTCGAGGTGCTATCCCGCCCGTTGTG TCATTTAATCTAGGATCCCTTGGATTTCTCACTTCCCATACATTCGAAGATTATAAGAAGGATCTAAGACAAGTCATTCACGGGAATAGCACTCTGGATG GTGTTTATATAACTTTGAGAATGCGTCTTCGATGTGAGATATTCAGAAATGGAAAGGCAATGCCTGGAAAGGTGTTCGATGTCCTAAATGAAGTTGTTGT TGATCGTGGTTCTAATCCATATCTCTCCAAAATTGAGTGTTATGAACATGATCGCCTCATTACCAAGGTGCAAGCTGATGGGATCATAGTAGCCACACCA ACTGGAAGTACTGCTTATTCTACGGCTGCTGGAGGTTCCATGGTGCACCCAAATGTTCCTTGCATGCTCTTCACACCAATCTGCCCACACTCGCTCTCAT TCAGACCTGTCATTCTTCCGGACTCTGCAAAGCTGGAGCTAAAGATTCCAGAGGATGCGCGCGCAACAATGCTTGGGTCTCCTTTGATGGGAAGAGACGGCA ACAACTGTCTAGAGGAGATTCTATTCGGATATGTAGTCAGCATCCACTTCCTACAGTCAATAAATGTGACCAAACAGGTGATTGGTTTGGTAGCTTG ATCCGTTGCCTGAATTGGAATGAAAGACTAGACCAGAAAGCACTCTGA

12	>95ccf40b-2581-4141-90df-ec94047d1c54 (sequence:mRNA) 1013 residues [Nitab4.5_0001168:291056-300158 strand] + [peptide]MAAYCPCKFDMGRKVAGGPIHFQDCQLRYVKISGFGIGFSYGYRRGRLRWVQRRRQKKLVVGAELSSAFSSNVGFDSQPRDILKLPWIGPL PGDIAEIEAYCRILRAAEQLHNTLMETLCNPVSGECSISYDVPSEDKHLLEDRIVSVLGCMVCLLNKGSEDVLSGRSFIINSFSDFDVHVMDDKLPPLAF FRGEMKRYCESLHVALENFITPDDPTSINVWRKLQRLKNVCYDSGFPRGDDHPHHTLLANWNPVYFSSEEETQSASSEVAFWTGGQVTEEGLRWLLERGF KTIIDLRAETIKDNFYEKVLDEAISSGDIEVLKLPVEVGTTPSVQQVEKFAALVSDVYKRPIYILHSKEGVWRTSAMVSRWRQYMTRYTPLFVPNANKDVT SSVNSFCGSRGTQEAGTPVNSEENKTSTCEGMSASDHKNGTLPARSNSINSAGKLFKQIPEAREHKGLSKNEADDTVAVTWKGTLLTADGKTNPLKSQLP PPKFFSRTEMSTYFRSRKVSPETYFTHEKKRLEGLHASRYYYKRIPKGNAIIDSYTEDRAIDSRNPNGPPSNMGLSTKPSNSSANMEKYGGHNGSAAPIL NRFNNGEVHTSVKSSSLIDASNELDTNAVSSATAIERRNIEAPRPSVDDNMELIEGNMCASATGVVRLQSRRKAEMFLVRTDGFLCNREKVTETSLAFTH PSTQQMLLWKSTPKTVLLLKKLGQELMEEAKEVASFLYYQEKMNVLVEPEVHDIFARTPGFGFVQTFYSQDTSDLHESVDFVACLGGDGVILHASKLFR GAIPPVVSFNLGSLGFLTSHTFEDYKKDLRQVIHGNSTLDGVYITTLRMRLRCEIFRNGKAMPGKVFDVLNEVVVDRGSNPYLSKIECYEHDRLITKVQAD GIIVATPTGSTAYSTAAGGSMVHPNVPCMLFTPICPHSLSFRPVILPDSAKLELKIPEDARNNAWVSFDGKRRQQLSRGDSIRICMSQHPLPTVNKCDQT GDWFGSLIRCLNWNERLDQKAL
13	**>Nitab4.5_0001273g0050.1** **Nitab4.5_0001273:290222-307380** ATGGCGAAGCTGAAGACGGATTCTCCTCTCTCTCGTCGCATAGTCCTCTCCTTCTTGCACTTTCTAGATTCTGGTACGCCTCTGACAATTGCCCCACATA TATATACATGTGTGTGTTTACGTATATATAATTGTGTTGAAATGAATCCATACAAGATGTTTCTCTTTTCTGGACGTAGCTAGTCAAGATCAGTCTATAG GAGTACTTATAGATCAGAAATTTGTATGGATTAAAAGGAAAAATAAGCTTCACGGGGTATGCTGTTCAGCATTGAGACACATGGTGCTAAATCAGTGAAA ATCAGTTTATTTCATAAACTACTGGCTGATTCTAGGCAGGGATGATACTAGCTGTAGTGGACTGCATAGATAAAGTTGTCTCGCTGACTAAGTTTCCCTT TAAGGGGATTAGGATTAGAATCTTACAAGGGCAGCTCAATGGAATTGATGCTCTAAAGCCAAACTTCAATATGAGGCCTAAAATATGTTTCTTAAAATTC GTATATATTTTTATCTGATGTGACAGCTTGAGAAAAAAAGAATGTGAAAAGGAATGTGAAATAAAATTATTCAATGACCCTTGTTGTTGGCA AAACAGTGAGAGAAAGGGAAGGAAAGAGAGCTTGATTTGATAAAGTACAATAATTGAGTAAATTGAACGTTGTGCCACTAACTCGTGTGCTTGTCTGCCTC CCAGCTGAAACACGGGAAAAAACGTTTGGTTCATGTGAGGAAAAATAAATGACAGTAGGTTGGATTATAAGATAGGTTTAAGTGATTAAATGAATAGACAA ATAATCCGGCACTGAATCTTGTATTATATGAAGGGAAACTAAACTTCATACGGTACAAGTAGCTTTAAGTGGGGAGGAGGAGAGGGGTTTAGGATT AGACTGTTTTATTTGATTTAGGTACAAGTGATCTTCTCAGTGTATGCTGTTCAGACCTGAAACTCATCAGTGAACTGTGAAGCATTAGTTTCTTGTATTTG GTAATTGACTGGTGTTGAACCTGATTACAATCTTTGATGCCAGTAGCTTTAAAGGGATTTAGCATGAGAATCTTAAAGTGACGGAAGGGAAACTAATCTT TGAAAAGTATGTTAGGAGTATTTATTACTGACACATGTGGGTCGTGGGCACAACATTAGGGACTTGTGTATATAGCAACTGATGTGGTTCTAGGTTAGAAGG CATTGTTTATTGAGATTAAAGAGTTACATTTCTTTGAGAATGAGTATCCTACCAAGGTTTGAAGAGAACTGCAAATAGGAACTTCCTGTTGCCGATTT TTAGCATTTAGAATACTGTTGGAATTCCGAACATAGTTGTTAGGCGAGTAAATCACATATCTATCCTCGATTTTGAGCTAAGATACAATTTGACTGTATC CCTATGGTATGAACTTGGTGTTTGGGTGATCCATAGAGCAACTGTAAATGAATTCCAACGAGATTATGCTTGCATGTGAGCATCATAGATTAAAAAAGT GAAAAAAAGATCGTAAGCCAAAAAAAATTGAAAAAAGCTCCTGACCAACTGCTTTGAAAGAATAAACAGAAGTAAGCGAGAATCATAATGAGCAAATGATAT TACTAGAAAAAAGAATAAGCAAAAACATTAAGTGAAAAAGAAAAAGCCCTAACCATCACTGCTGTTGCCTCTCATCCCCCACTGCAAAGTCTGCTCAGCC ATATAACCATGCTTTTGTCATCCAGCCTTTGGGTGAGTAAAGGCTGGCCACCCGCAGCCTGGATATTTTTTTTTTTTGCATGTCCTTTCCGTTCCTAATTT TTTTCCCTAGGTGCTGGTCAATTTAGATTTGACATTATCAAATAATTGGCATATGTGTAGGGGCCAGGAGCACACTGGAAAAAGAAAAAGAACTTAATAGA AAATGTAAATGCATCATGCAACTTAAATGCAGCTTATTCTTATATAATTATCCTCTTTTTACCAACACAAAGATACAAATACAGTAACTGGAATGAAA

AACTAAACTGAAAAACAAAAATTTAAACTCAATATAATGCATCTTGCAATAGATTAAATGAAGACGACATGTTGCTTTTAAACTGTGCAAGTATCACTA
AAACATGATGCAGACGGAACACATAGGATGTTCTTTAAAGTAATCGTTTTCAGGCTTCTGGATCTCTTCTTAGAGCAATCACAGCAAGATCAACATTCAT
GGAGATTTCACTTGGTGAACCATGGCAGTTTTGGGAGCCTGAACCTCTGGAGAAAATTTTCAGAGCTTCGTGCTACAGACTTTGGGTGGAATAGAATTTC
ATTTGTCAGATTCTGCATAAATATTCTTTTTCTGTCCGTTCCAAATATCAAACTTCTGAGAAATACATTACTGTATAACACAAATCATAATAAATCCTGAC
CAAAGATTTAAGCTTTTTTAATCATTACTCAATAAAAGGATTAAACATTCTGACAAGCAAAAAGTGATATGAGCAGAAACTATTATACACCCTTCTTCTCA
TTCTTTTTGGGTGCAGATTGCCATGCAAAAGTATCAGCTAAAAGAGTAGTTTGCTAATGCAATTCAAGAATTCAAAGTTTAATGGATTGAACAATTGGCC
TACAATATTTATTAAGCGGGAAAAAAAGAAAACCAGACAAGATTTAAAAGGTCTGCCAAATTTGAGTGTAACAACTTTAGGAGGGCTTACTCTTGCACT
CTTTGACTGTAGCACATGCATTCCTGAAACTAAGCCTCCTTTACTAATAACATTTATCCATGATTTGAGATTTTTAGCTGAACAAATAAAGGTT
GAATCTTGTGCATTCGGAGTGCGACCCTCTAACCATAGAAAATCCATCAATAAACATCTACAACCTGCCCTCTCTCCCTCTCCCTAGCTCATGTGC
ACGTATTACTTCTCTTCTGTAGCCCCTTCACGATTGGGTGGAAGTCCTAGCACCGCTTGACTGGGGGAAGAATTGGGAAATACCAAATCTTTTTTGGAAA
CCTCATTGCTATGAATAAAGCCATTTAGATGTTATATTTGGTGCATGCTACTTGTTTCTATTAGTCATGGTTATTTTTCTGTTTTTATTCTGTTTAAA
TTGTGAAGTCTCCTCAAATCTTAGTTTCAATGTAGAGTGAGTGAATTTTGCATCTACTGACTGTGTATTCCTTGTTACATAGTTGAACCTGCTTCTGG
AGTTGATGATGAAGGACTTGAGGTTGTTAAGCAATGTCTATCGGAGGCCTTTAAGATTGATCCATCATCTGCTGGTTATGCCTCAAGTTCAGATTCCTTA
GTGGACATATTCAGTTCACGAGAAGCGATTGAGCAAAGCCAAATAAATTTGGATCTCAGGCATGATGTATCTTCATCAGATGCTCCTTGTTCATCATCAG
GACAGAAGGTAGCAGATGCTAAAGATACAGATGGATCTCAGTTTCTGGTCTAATTCTACCTTTTTCTCTTGCTATTTCTCTTAATTTTGTGTAGTTATC
TTTCCAGCTTTTGATTCTTTTTAATATTTTCTCCACAGTTCATGTATTTTAGGTCTATGTCATGTTTGGTTTCTCCTTTCTGT
CTTCTATATAGTGGTCCATTAGTTATGCAATCAGTAAACCTGTGAGGTTTAAATTCATTCTGCTGTCTACATCTATCTGCTCTTATTTTCTATCACCATA
TAGACAATAATAGTCTCCCTTCATGGTTTGGCAAAAGTGGGATATTTGGGGCATGGCCATGATATGTGCTTGATGGAATTTTCGGGATAGATTCA
GATTTTGCTTGGGAAAGGAAATTTTAACCAAATTCTGTTGCTAGACAATCTTTTCTTGAAAGAACTTTAGTTTTTAAGGATGGAGTTGGTTCAAGCTAT
CATTGAAATGATCATCAAATGCTTACATCTGGTAATACTGTGTGCTTGGCTGTACAGCCATGGCAGAGGACTTATTGTACATTCTCTGTTGATGTTGACA
TTTGTTATAATCATAACCTATTTAGATCAAACATTAGCTACTTTTAATCTCCTTCATATGTTGCAACCAATCTGTGAGTGAACTTTTAGTCTGTGGTAGC
CTTTTACTTTCCACATACTTGTACATGATTTACCTCTTACCTCATTTTGTTTTCTCTAAACAGGCTGATGAACGGACAAAAGAAGCTCCCACTTTGGTAA
GGCTAGGCTCCAAGTGCCTAGGTTTTTATAATATCTCAATGAAGGCCTCTGTATATTGTCCCTCTGATTCTTATCTAACTAGTGTACCACACTTGACCTT
CTCACTGTTACTCTTTCGTGTCCATACGTTTGTGCTGTCATTCAATTTTTATCTCATGTCATCTTCAGTGACGTTAAAGGCTTGCTTGTAGCATGCTTAA
GCTCTGAAGTTATGTGAAGCCCAAGTTAGAGGGTCCCCCTTACTTAGGTGGCTCATGACTGCAAGCTCCAAGGGACTATGATATGCCTCATCGCCTTG
AATTCTGTCTTTAAGAGGGCTACACTTGGCAAATAGATATAGTTGGCAAAACAATAGAGTAATTATTATAAATTGTGATTAAGGAACTCTTTTGGTGGTCA
GGAATATGAAATTAGAGAAGTGATATACAAGCCTAGTCACTAAAAACTTAAAAACTGTATTTTTACATCTAAATAAAAAATTTATAATGGTTTTTCAACT
TGCTTGACATGATTTTTATTTCATATATTTTGCCTAGTTCCTATTTTTTGTTAATTTGTTTATCCTAAATGATCATTGTTTGTTTATAGTGATTCTTAC
TTGCTACATATACTTTTATGGCAAGTAAATAATTATTCGCTTTTTCTTGCATATATTTATATAGTGTTTGTATTTTATTATTTC
TTTATTAGCAATTTCTTCCCTAAAGCGCACGAGTTGAATTGAGCTTTGCGTTGAATGCCCTAACATGCTTTTTGCCTTTCAAAACTCTGGTCAATCTTTT
TGGCAATGAGTTCATTAATACCTCAAATCTAGTTGCCTTTTACTTTAATTTTTTCTCCTAGTTTCACAGTAATTCTTTATACTGAAGCAAACCATGAAAA
ATAAGGTCAAAATTTATACATAGTTACATACATACATAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
GTATATATATGTATATATATATATATATATATGTATGTATGTATATATAAATATGTATGTACATATGTATATATGTAC
TAGTCTTAGGGTACGCGCTTTGTGCGTGTACCATATATTAATGAGTACAAAACTTGTATGAATCATATATTATTTTAAACAATATGCTTAAGTTATG

AAATAAAAATTTAAAAAATTATATTTCTTGAAAGTAATGAAGGTAGATCCCATTTAGCTAGTTCATTATGAAAAACATCGTGCCCTGTGGAAGTCTTTA
GATGCCTTATTGTGATTTTTGAGGACTTCTTTGAAAAATTATTGCTATTGTGTGACATAATACTCGAAACAAACAAGTGACACAAAAA
TAGTTCATTACAGTTCTCCAAAGATTTGTCAAATAATGCAAATTAAAAATTAAATCTTTTGGTCTTTTAGACATTTTGGAATGAAAAGGTTGGTTCTATG
GCTCATAGAAAAATAAACTAATAAGTCTAAAGTGACAATGTAAAGTCTAAAAAAAAAATCAATATTGAGTTAATATTTTACAAAAAAGAACTTAATATATAAAAAAAA
TAAGAATTTGTCATTTGTTGGACAAATGCAAAGGAAGGAAAAATAATTAATGATAACTCATTCAGAATATTTTTATTTTCTACTGTTTCATCCTTATTAC
TTTAAGTTTGCATTTTCACCAACTTGACTCTCAATACGATATTGTAACTAATTTCTGTTTAACTGTCAATGCAGTTCTTATGAAAATAAATTTACGTACA
CTACGAGTTTGCAATTACTTATATGATGAATTTGTAAAATAGTTGAATTGGATTTTCTAACTAATTAGTTGATTCAAAAACTTATTTTGTTAGATTAGTT
CAAAGATTTAATTATTCATTCTCAACATTAGAGGAAATAATTTTCCTTATTCAAATTCTTAATGTTTTAACTATCATCCCAAATTTTTAATGTTTAACTATAAC
TATATTATTATCCAATAAAAGTCTATTTTCAAAGGATTAAAAACCGATCTGACCTTTCACATAGAGAATTTTAGTGTTTTTAAAATTTTTAAATATTGG
AAAGATAATTAAATTATAAATCTTTTTCAATGTAAAATTTATCTGTAGCGTATAGAAATTACTGATCAACATATTATAATAATTTCTAAAATTTAAT
TATTTTTAATCACTTTATTAGATTGTAAAAATCCAAACCCCTCAATATATTTAAGTAAAATACCACACAATAAAAAAGATACTACTTTAAGTCAGAATGTTA
ACTACATACGAATATTAGAGAAAAAAATAAATGACTATTTTGTGTAGAGTAAACTTTATTTTTGAAGGGTAAAAAAGGCGAACGACATTTTGCTAAGGGC
CTTCGTGCTTTTGATATAGTATAGATAAATGTCATACAAGATATGTATAAGGGATTAATCACAATTCACAATGTGAGAATAGCAGGAGATATAGAGGAG
TTTCCAGCAACTGCAGATTTACGCCTGGGGTTTGCCTTGAACCTGTACTTGTTTGTATAGGTATGTTGAGTTAACCAATAGTATACATTGTGAGGTTTC
ATGGTGAATGCTAGTTGCCAGTGATATTCTGTCAAACAAGCGATGTAGTCAATAAAAAGTTTGAATTATGGAGAAGTACTCTTGAAGAATAACAATTTAG
ATTCTCTACTTCTTGTCTGCATCTTGTATATTGGAGACTTTCCGCCCTACTGCACGTCAATAAATTCACCTTAAAAGAGAGTAACGGTTTTAGATAAGTA
TAAATTATACAAACTCTATATGCTACAAATTAAACCCATAAGATAAATGAAGGTGAAGTCTCGATTAAACAAAATGAAGGTTAGACAAGA
TTGTGGTTTCTAAATGTAGAAACTCAGATATCTAGTTAAGTCTTCCAGCAGAATGACATTATAGATGAAGATGTGACACATTAGATTAAAATAGTATGAC
TAAAATGGAGTGCTTTGATAAAAAGATACTTAACGAAGGGAAAGGCAAGTTTCATAAAAAAAGGGTTATCATACTAGCAATGCTATGTAGGGGTGAACGT
TGGTCCTCTGAGTCCAACATATCCACAATATTGATGTCTCAAAAATGCGAATGTTAAGATGAATGTACGCTTATATACTAGATTGAACAAGACTAGAAATGA
CCACAACCAATAGAGGGTGTAAGTAGCACATAAAGGGATAAAATGATAGAAGCTCGTCTAAGATGGTCTGGCTATTTCCTACATAAACTGCGGTATGCA
TTGGTCTGTAGGTACTACGATAATTGAAGGTGCTAAAAGAATTACAAGGTAGATTTAAAATCAAATGGAGAAAGTTGTCTCAAGAGACCTACAACCTCTC
ATAATAAATACGGACTTAGCTAAGAACGGAAGAAATTTGAAGTAAAGAATCTATTGAGACTACTAACTAGTTGTATTAAGGTTTAGTTGTTGTAGCACTT
ACATTAGGTCTAGTGTTTGCGGGGAGTCTTTCTATTCTGGATAAGATGTGCATGCCTGTCTTTAGGAGAAGCAACTGTCACTCGTGACAATGGATGCCTG
CTTATATTAAACTATCCTCAATCTCTTCTTCTGTAAAAGCTAGATTCTTCTTCCTTCTCCTTCGACCCAAGCAAAGGAGCATGCATTTTTCTCATAGTG
TGAGAAAATGAGGTCGAAATCTAGTTCTATTTATCATTCAGAGAGTTAGAAAACTAGACATACTGTCAAAGGTTGGGTTCTCTAGGGGATCGACTCCTAT
GCTTTGCATTATGTAAGTTGGGCTGCTCCAGCTGAAGTGCATTAGAAACTTTTGTCCGTGGGATTGTCGGGGAAACACCCGATATATGACTCCTTTTAAT
CTGAATAAAGACGTACGCAAGTATAAGGATAAGTGTGATGTAATTGGTACCATATAGACAATTTTAGTATTTGAATGAAATGGGCCGAGTAGAGCGG
GGTGTTTATAGAGCTCATATAGATTACATCCAACTTGTTTAGAATTTGGGTGTAGTTGATTGATTGTTAGATGTAACGAAATCATTCTTGTAAGCTACAT
CCGTGTTGCTATTTTCTGGCCCATACAAAATCATTCTTCTAAGCTAACTGTCTTGCTATTTTCTGGCCCATCTTCTTTGTACTCCTATAATTCATTTGTA
TTTGGGTTAATTATTTCCCTTAGAATGTGGATGTGGGCCTTTAGTTAAAAATGACATACATGTAGGCTACACACATGGTTACAAAGTATAGCAAGCAGGG
GCACCTACTTTTCTTGTAATCCTTTTTTTCCCTGTTAGCTATCGATGCACGTTCTTGGGTGGCCTTGTCTTCGTCATATCAAATACATCATCTAAATGTT
TTAAACCTTGTTTTGCCTTCTTTCTTTGAGTATCTTACAAGATATGTTTGTGACAGTGTCAGGTAAAAGTGTCTCTCTTTCTGATTTGCCCTACGTTCTG
TCCTGCTTTATGTTTAATCAACTGATGCTGTTTATCATGCTTTGCAGGTATATCCAAGGATGAGATCTTTGGGCAGTTTTTTGATGCCCTTGAAAAAGCTC

ACTATTTTAGAAGCTTGCCAGATGGAAAAGATGATCAGGCTCAGCTGGATAGAGCATCACGGGTATTTCACACTGCTTTGGAGGTATATATTGACAACGA
GTTATTTCATTCATGTTTCTACTATTGTAGTGCATCAGTTTTCCATACTTGATGGAATTGGTTGAGAAAACACAGCATTTAATCTTCACCTGGTGGTTTC
CATTTTTTCTGTTTTTGAACAAAACTTTGCTTTTTCCAAACATTGATGACAGTATCTTCTGTTCCCCATATGATTAAAAATACTTTCT
AGAGTAGAATGGCTCAAAGTTCTTTTGAAAGAGCTCATACTCCTTTTGACATCATTGTGACCAGCTAAAAGGTGAATCTATGCAGTTGCATGAAGTCCTT
CTACTCCCTTATGACATCACTATGACTGAGGTTAAAGGGGAGACTCATTGTTAGCAAAGGCTCGCTTGACTGCGCTTAAGGATATTTTTGATTAGCTTAG
ACGAAGCTTCAATGCATTCACCAGGGCACTAAATACATGCGCTTCATCTCTATGCTAAACTCTCTCTAAAGGGCTCATATTTCAAATTGATGAATTATATTTTC
TATGAATAGGAATAGACATTGAGTATGTCACAGTTTGTTGTGTACATATACTAGCTACTGTTGCTTCTGTCATGATTGTCAAAATTATTGCAGTTCTTGT
TCGAATATCTTCCTACAACCTTTATCGTTTATGGTTTTAGGAGATGCAAAAATCAGGATGTGCTATGTTAAATCGAAACAACCTTGCTGAGACCTTAAAA
TCACAAGGTACTTTGTATTTACATGTTTCTTTTCTTTAAGAGTAAAGATATTTTATTGATGAAATGCTGAGATTACCAGTATCATAAACTAT
ATCTCGTTTTATGTAAGGAGTCAATTAGATGTAAGTCATTTATGTTTTGCATTTTACGCCAAAAAAAGAATATTGCAAAGAGCTATCCTTTATGCGTTGT
ATGACTTAAGTCTTCTTATCACAACATCTGTATTCCTTTCTCTCCATACATGCGCGAAGGCACTGTATTCCAGATGTTTGCATTGTCTTCCTTATTTTCT
TCCCTTTAAAAGCTAATCACTAGCTTTTGCTAGAATAATATCGTAAAGTCAGAGAAGAAGAATGCCATTATAGAGGGATCAGCGTGTAAACTTGAAG
GGCTAGGGAACTCCTTTAGCTCAATAATGAAAGTTAGGGTCGATCGTAGGAGGCCTTACCTCTTAGATGGTCCAAGCAAGTCATTAAGCTATTGGATGGA
TCCTCCTAACTGTAGTTACTTACATTCTATAGCTTGTTTCACAATGTTTGGTATCACCCATGGCGAACAGAATATATTTAGAAATAAACTCCAGCCTTGC
CGGCTAACAGGGAACTTTAACAGTAAGTGTTTTCCTCCGTTTCTTCACACTAGTAACAGCTGCTAACTAAATTGAGGCCTTTCTTCTGCAAGT
CCTCATTGTTTAAGCAGGCTTCATATGCTGCAATCTAAGTAATCAATGTTTTAACAGGGTTTTTGATTTTTCAGATTGTATTCTAAGGCCATCGTCTTCTTTC
TTTCACAGAACTATTAATACTTCATAGCAACATTTCACTGTAATTCACCCTCTTCTTTTTTAGAAAAAATCGTAATTTACTGACCTTGCTGTCCAAGCCG
AAGGGTCAGGTTCAATAATAACTTGTTAGCTCCTTCTAGCTTACTGAAGAAGTAGGTTATGCTTTCCCATTTCCCGATCTTTAATGTCTCTTGAGCCTGAT
GTCCCATTTATTCTCAATTATACAATCTTCAAATATAGGCCTTCTTGTCCAGTGCCAAACTGTATAGCTAAGGAAAGGCCTTCTCAGTATATTTTGCTATA
GCCAGAAATTTTCCTAGTACTTGATTTTCCTTCCATATCCAACTTTCAGCTTTACAAGATTGGTGAAATGTTCCTCTTTCTCAAAAAAAAGATTGGTGAA
ATGTTCATAGTCCCTTTATATGTTTCCAAACAGATATTAAGTTGTATGAATTCTCATTATTGCACCCAAGAGAATGGCCTATCAGTCGATGAAGCTGGAA
TAAGGACCCCAGTCAAGGCAAAAAAAAAAAGAAGTTGCGAAATCTCGGTGGAAATTTTGGAAGAACACGGGTAGGGCATGCTTGGAGTGGCCCACCGAGTT
GTTTAGTGTCATTTTTAGGATGAAGAACATGCCCGACGAATGGAGGTGGAGTACGATAATTCCGTTGTACAAGAACAAGGGTGATATTCAAAATTGCAAC
AACTATAGGGGTATCAAGCTGCTAAGCCATACTATGAAAGTTTGTGAGAGAGTGGTGAGTAGGAGTGTGCTTATTTCCGAGAACACTT
CTGATTCATGCCGGGTTGTTCGACTACGAAAGCCATTCACCTTGTTAGGAGATTGGTGGAGCAGTATAGGGAGAGGAAAAGGAATTTGCATATGGTGTTC
ATTGACCTGGAGAAGCATACGACAAAGTTCCGAGAGAGGTTCTATGAAGATGTTTAGAGGCTAGCACTGTGCCTGTAGCCTACATTTGGGTGATTAAGGA
CATGTATGGTGGAGCTAAGACCCGGATAAGGACAGTAGGAGGAGACTCGAAACACTTCTCTGAGAGTGATGGGGTTGCATTAAGGATCAACCTTAGCCCGC
CTGGCGATGGATGCACTGACGCGCCACATTTAAGGGGAGGTGCCATGGTATATATGTTATTTGCAGATTACATAGTTCTGATTGACGAGACGAGAGGTGGTG
TTAACGAGAAGCTGGAATGTTGGAGATAGACCCAAGAGAGGAAGTTTCAAGTACCTTGGATCAGTAATCCAAGGAAATGGGGAGATTAACGTAAATGTCG
CACATCGTATCGAAGCGGGGTAGATGAAATGGAGGCTCGCATCCGGTGTTTGTGTGATAAAAATGTGCCACTAAAACTTAAAGGTAAGTTCTACAAAGT
GGTGGTTAGACCGACTATGTTATATGGGGCAGAGTGTTGGCCAGTCAAGAACTTTAATGCGCAGAGAATGAAAGTAGCTAAAATGAAGATGTTGAGATGG
ATGTACGAGCATACCAGGTTAGATAAGACTGATATTCGGAAAAAGGTTGGAGTGGCCCTCATGGAGGACAAGATGCGTGAGGCGAGGCTAA
GATGGTTCGAGCATGTGAAGAGGAGAAGCACTGATGCTTCAGTTAGGAGGTGTGAGAGGGTGACTTTGGTGGTGTGCGGAGAGGTAGAGGTTGCATGCT
GTTTGATTGCTATCCTTTGCTTTAAGTTTTCTTAATATCTTGATGTTACAGCTTGTTCTCACTGCTTTTTTTTCATTCTTTCTTGAGCCGAGGGTCAATC

AGAAACAATCTCTCTGCTCTCCCAGGGTAGGGGTAAAGTCTGCGTACACACTACCCTTCACTGGGTTTGTTATTGTTATTGTTATAGAGCTAGGT
ATTTTCTTCCCATCTGCCTAAGCTTTAGTAGGCAGAGTTACTAGATAGATGCCTGCACTAGTGGAAGGATGCAGATACTCGGTGAAATAGTTGAGTTGCG
TGTGAACCGGTCCAAATACCGTTGTTTTCAGAAAAATGAAACTTCACTAATCATTTTGCTATATATTGGACCCATTTCATTCAAGGCTCACCTAGTTTCGA
GAATATTTAGGAAAAGACAACTAGTTGAATCACAATTACCACACATGAATTATCATTACTTCATTCTATAGGGTAACCCTGCAAAACATTTTAAGTTGCACT
CAGGCTACGCTAGTCTTGGCTCACAGTAAGCCCAACTTTCTTCCCGGCCTTAATTCAGCACATGGTCCATTTTGGTGGGCCGGAGGAGTGAGCTTTGAGC
TGCCTCAATTTAATTGGGAAGAATAATGTTGTCAAAAGGTCAGTTTCTAATTCTAAAGCAATTCATTGGCCTTGAATATCTCATAGATATGTACTCAAGG
TTGATGTCTACTTGCGTAAGTCTTTAGTTCCAATGTTGAAGTTCCCTTCTTTTTGAAAGCGTTTGCTTAAAGTTTTCTTCTATTTTTTTTTCTTTC
TGGGATATTATAATAGGTAATAAAGCCATGCAAGCTTTACTCTGATGCTATTGAGCTGTATACTTTTGCAATTGCACTTTGTGAAGATAATGCAG
TTTACTATTGCAACAGGTGCGTCATTATCACAATGATGGGAGAACTTGATTATTGCATTTTCCATGAAAAATTAGTACTTCCTCTGTTCCATTTCCTGTG
ACACTATTCCCTTTTCAATCTTCCGAAAAAAATGACACCCTTCAATATTTGGAAACAATTTATACTTCACATTTTACCCTTAATGACATGCTTTTTATAGC
CATAAATGTCATGACATTTTTGAAACCACAAATTCCAAAAGTCTTTCTTTGTTTCTTAGGGGTCATTTGGTAAAATGAATTGGATAAAATAATGCATGCA
CTAGCTTTGTGTATTGATAATACCTTGTTTGGTACACTTTTTGAACCTATTGTATTAGTTGCAAGCATGAGTTATACACCCTATTTGGTATTATCCTAT
GCATAACTAATGCATAGGAAACTATGGTATTTAACAATGCAATGGGTTTTAATGCATTGGCTTAGTTAAAGACAAAATTGCCCTTCAAAATTTATGC
TTGATTAAAATATGCTATCATATTAATGCAAGTTAGATGTAAATTTTTTGATACTTAAATTGCTCCCGCTATTTATATCTATACTCCGCCAATTTTTTCA
TTGTAAATTTTATACTTCCTCCGTTTCAATTTATATGAACCTATTTTCCTTTTTAATCCGTGCCAAAAGAATGACCCCTTTCATTATTTGGAAACAATTTA
CCTTTATGCAATGATTTATAGCCATACAAAATATATGAGACTCATTTTACACCACAAGTTCAAAAGTCCTCTATTTTTTTTAAACTCCGTGCCCAGTCAAA
TGGGTTCACATAAACTGACACGGAGGGAGTATATCTACTACTCTACTATGTACTGAAATGCACTTGTAATCAACTCATTTTCCATTTATTTTTTTAAT
TCAAACTTTCTTTGCCTAATAATCCAAGAAAGTGGGAAATAAAATTATATCCCTAGTAGATAAATAAATACTTAGCATATTTCCTTTTTTATAAATAAAT
ATTCAGTTATATACTACAATATAGGTAGACAAATCAAATAATTTTTTTTAAAACCTTTTTCATATAAAAACATTCCTCAACATATGTTTCTTTTAAAAGA
TAAAGTGATGGACTCGTTATGAGTCTATTTTTATAGACAAACATTTCTTTTAGAAATTGTGCAATGCTTTAATACATCAAATCAAACAATGGATAAGAAA
TATGTTAGTATAACTAATGCCAGCATAACTAATGCAAGCATAGCCAATACCAGTATTACTAATATACCGTATTCAACATTATTCTTATGCACCCTACCAA
ATGACCTCTTTAAACTCCATGCACAACTCAAACAGTGCCGCTTAAAATGAAATGGAGGGAGTTCTATTTTTATCTTTCTTTGAAAATTCTAATAGCCGTAAA
TTTGAATCTTTTCTAACCTGTAAATAGTTATCGTGGCCTATTTTGAGATGAGGTTTTGTTGATGTATCAGGGCAGCTGCTTTAACCCAAATCCACGAATACG
AAGCAGCAGTTCAGGACTGCCTAAAGTCTACTGCAATCAACCCCAATTATAGTAGATTGGGTTTTGTCTATTATGCTCAAGGGAAGTA
CCGTGATGCCATAGAGAAGGGTTTCACAAAAGGTTGGATAAGAGTATCTTTTTCCTTTACTACCTTATCAAAAAAATATCTTTTTCTTTACTAGTTATT
CAGTTTGTGTTGTTTATAAATTAACATCTTGGTATTTTGGAGTACATTATAGTTTTTAATTTATTGAGGATGGTCATCATTAAGCTCTGACTAACTTT
GATATGTGCGTACTAGGGTAGTTGTTGTAGTAGTCGTGTTGCTTCTCATGCTCTATAATTTGTTTCAGTTAATTTCATTAGTTCAGTATGAAATTTGATTTG
TAGTTCTTTCCATGACACATAGCAGTTCTAAAAAATGTTTTGGTGATTGCTAGCCAACTGTGGTTTGGGATTTGTCAATGAAACCATTATACATAGAATTC
TTTCAGGAAAGCTATACTCTGAGTCTCTCTTTTTCCTTTTTTATTGATTTCTATGTCCTTCAATATTTAAGATCTTTTTTTGATAAGGTTACAATATTTAGGT
TTTATTGGTTCTAATTTTTTAAAAATGTATAAAGATATAGTATTCACTTGAATTTGTTTTTTGTTGGAATTAACAAATATAGAAAAATAGCATGAAGTGGGTT
CCTGTAAAATGGCAAACAAGAAGAATGTAGTAAATATATGATGACATATTATGAAGATGACTGTAGCATGAAGAATGGCTTTATAGACCTCAGCTAACATTA
CTGCAACTTTCTGGAATCGTGCTTATTTGGAGTTTCTACAGCCTGATTTCACGTTTTCAACAGGTTTATTTAAGCCCTCTACTTGAAAGGAAAACAGAG
TGTTGGCTTAAATGTCAAAAACACAATTGGATGTCTGACAAAAACGGATTCTCCTACTCCATTAATGAAAAAATTAAAGCACCCTAAGGGTGTGGCCTAAT
GGTCAATTAATGGGGAGGTGAACCATGAGATTTTCAAATCCCAGTGGAGACCAAAAACGCTAGGTGATTTCTTACCATCTGCCTAAACCTTGGTTGGCAG

	GTTACTCGATACCTATGTTGGTGGAAGGTAGTAGGTATCTTGTAGTACCCACTAAAGCATGACTACGAAAGTCGACTAGAAGCAATTTAGAGAATAGAAA GAAAAACAAGTCAAAGAAACTGTTCTTGCCACAACCGATGTCCATCTCTAATGCCCCTCGCTTCCGTTATTTCTTTCTTCCTCCCTTCCCTCATCAAGAGA TAGTGAATGGAGTTCCATCCCCAAACAACCAATAAATGGATAAAAGAGAATCAAGACGGAACTGGCCTTACACATGCAAACTGGGGCATGGACTGATCTG TGGCTATGAGGGACTTGTTTTGGTGAAATGAAGATTTGTGTATATTATGTGTTGTATTGGTGCTCGGTAATGGGTCTGGCTACTTGATCGGTATTGGT CTTGTGAAGAAGTGTAAAGAGGTTTTGCACATTATCTCTGTCATCCGTGTTAGAAATTTTGATGAGCAGTACCAGAATACGAACATGTGAATTTTGTTTT CTAGTTAATCCCCCAAGAAGAAATCTATATCTGACGTAAACCCTTTTTTGTGAAGCATTGCAGTTGGATCCTAAATAATGATTCTATCAAAGAAAATATACG GGTTAGTTTCTATTATAGATTGCAGATGTTAAGTTTATTAGCTGGAAATGGTGCTTGCGATATATTCCAAGTATTACTCTCGTATTTTAAACTTGTAGGT AGCAGAACAGAAGTTAAAAGAAGAGGAGCAACAGAAGAGGGAGCATGATCAGGTAACTGACAATTGTCGCAGAAGGAGCTTTGTTAGATTTATTAAGTT GATAATTACAAGTTATGTAATTAACACCCAGATCATGGGCTTAGCTCAATACTACCTATAGTGATTTGTCTGTCCTCCTCAAAAAAATAAAACAAA TTTTTTGATATGTCTGTCCGTTGTAGCACCATCTAGTACCGCATTATTGTGTGAGAAAACTACAATGAACATGTATTGATGCTTATTCATTATCTTTCTG TTTAAAAAGAAAGAGAATGGAGCATGTCTCTGTTATTCAGCCCTCTGACAACGTGGTTATAATATTGGCTGCGCATATCTAAGATTTGATTCCATTTATC CATTTACCATCTAATTTAGCAACGGTGAGCATTTAACTAAATTTTCATTTTGTATCCTTTGGTCTTTTGCAGAGCTCAACTTCTGCGAGTCATGGACAAG AATCAAATCAGCAGCCTGCTGGTGCATCTAGAAGTCATGCCATCCCTCCACACCTTTGCATCATTCGACGGCAATGGCATTTCTGGTATTCCTGA TTTGACTAACGTGTTTATGAACATGACAAGAGATGCATTCCAAGGGCAGCATGGCCCAAACGGGTCTGAAGGAAGCAGCGATAGTGATTCAACTAACGAG CCCGGAATAAGATTAGGCAGGAACATCAATGTCAATTTTGGTGAACAGATGCCTGAGGAATTGACTGGGGCCTTGAGGTCTGTTATGGAGATGTTTTCAG GAGCACAACCTCCCAGGAATCCTCAGGACAGTACGAATGGAAGGTCAACACCAAATTGA
14	>949b169e-468a-46f6-9454-d37cbc6a3a9a (sequence:mRNA) 1323 residues [Nitab4.5_0001273:290222-307380 + strand] [cds]ATGGCGAAGCTGAAGACGGATTCTCCTCTCTCGTCGCATAGTCCTCTTCTTGCACTTTCTAGATTCTGTTGAACCTGCTTCTGGAGTTGA TGATGAAGGACTTGAGGTTGTTAAGCAATGTCTATCGGAGGCCTTTAAGATTGATCCATCATCTGCTGGTTATGCCTCAAGTTCAGATTCCTTAGTGGAC ATATTCAGTTCACGAGAAGCGATTGAGCAAAGCCAAATAAATTTGGATCTCAGGCATGATGTATCTTCATCATGCTCCTTGTTCATCATCAGGACAGA AGGTAGCAGATGCTAAAGATACAGATGGATCTCAGTTTCTGGCTGATGAACGGACAAAAGAAGCTCCCACTTTGGTATATCCAAGGATGAGATCTTTGG GCAGTTTTTTGATGCCCTTGAAAAAGCTCACTATTTTAGAAGCTTGCCAGATGGAAAAGATGATCAGGCTCAGCTGGATAGAGAGCATCACGGGTATTTCAC ACTGCTTTGGAGGAGATGCAAAAATCAGGATGTGCTATGTTAAATCGAAACAACCTTGCTGAGACCTTAAAATCACAAGGTAATAAAGCCATGCAGTCAA AGCTTTACTCTGATGCTATTGAGCTGTATACTTTTGCAATTGCACTTTGTGAAGATAATGCAGTTTACTATTGCAACAGGGCAGCTGCTTTAACCCAAAT CCACGAATACGAAGCAGCAGTTCAGGACTGCCTAAAGTCTACTGCAATCAACCCCAATTATAGCAAGGCTTATAGATTGGGTTTTGTCTATTATGCT CAAGGGAAGTACCGTGATGCCATAGAGAAGAGGGTTTCACAAAAGCATTGCAGTTGGATCCTAATAATGATTCTATCAAAGAAAATATACGGGTAGCAGAAC AGAAGTTAAAAGAAGAGCAACAGAAGAGGGAGCATGATCAGAGCTCAACTTCTGCGAGTCATGGACAAGAATCAAATCAGCAGCCTGCTGGTGCATCTAG AAGTCATGCCATGCCTCCACCCTTTGCATCAATTCGACGGCAATGGCATTTCTGGTATTCCTGATTTGACTAACGTGTTTATGAACATGACAAGA GATGCATTCCAAGGGCAGCATGGCCCAAACGGGTCTGAAGGAAGCAGCGATAGTGATTCAACTAACGAGCCCGGAATAAGATTAGGCAGGAACATCAATG TCAATTTTGGTGAACAGATGCCTGAGGAATTGACTGGGGCCTTGAGGTCTGTTATGGAGATGTTTTCAGGAGCACAACCTCCCAGGAATCCTCAGGACAG TACGAATGGAAGGTCAACACCAAATTGA
15	>949b169e-468a-46f6-9454-d37cbc6a3a9a (sequence:mRNA) 440 residues [Nitab4.5_0001273:290222-307380 + strand]

	[peptide]MAKLKTDSPLSRRIVLSFLHFLDSVEPASGVDDEGLEVVKQCLSEAFKIDPSSAGYASSSDSLVDIFSSREAIEQSQINLDLRHDVSSSDA PCSSSGQKVADAKDTDGSQFLADERTKEAPTFGISKDEIFGQFFDALEKAHYFRSLPDGKDDQAQLDRASRVFHTALEEMQKSGCAMLNRNNLAETLKSQ GNKAMQSKLYSDAIELYTFAIALCEDNAVYYCNRAAALTQIHEYEAAVQDCLKSTAINPNYSKAYSRLGFVYYAQGKYRDAIEKGFTKALQLDPNNDSIK ENIRVAEQKLKEEQQKREHDQSSTSASHGQESNQQPAGASRSHAMPPPFASMSFDGNGISGIPDLTNVFMNMTRDAFQGQHGPNGSEGSSDSDSTNEPGI RLGRNINVNFGEQMPEELTGALRSVMEMFSGAQPPRNPQDSTNGRSTPN
16	**>Nitab4.5_0001803g0120.1** **Nitab4.5_0001803:310317-343182** ATGGAAACAAAGACTTGTAGTAAGTCTTACCTATTAACCTTAGAAGATATCCACCTAATGACTAATAGTCATCCTCTTAGTTGTGGCATTT TGCCACGTTTTTGGGAGGAAAAATTAATCTTTATCTCATATTTTAATAATTAACTAGGTAATATCCCGTTACCCGGTAATTAACCAATTATCTCAAATTA CCTAAAAATACTACTCATTTTTAATATACTTTATATATTCATGGTTATATATGGTACCTTATATGGTACTAGTTCATAATTATCGGGTATTATCG CTCGACCCGTATTTTATCCCAAATTGGCCACTTTCAACGAAACTCGTTTTCTTTAATTCATGTACCTCTTTATCCTTTATGACACTTACTTATCGCTTGT TGTAAATAGCGTAAATACGTTAACCTTAACATAATCTCATCCCCAAGTCTATGTCGATTAACTGAAGGCGAAACTTTTAACGTACAAAAACGCGAGATGT AACATCCTCCTCCCTTAGAAACATTCATCCTCAAATGTTTAATTCCCCGGGATCTATATAACTTTGGCGAAGTCTCCTTTGTAACAACACTACTACCAAC TCTTCCTGTAGAAACTCAATAATTCAACGCCACACATAGGACCACACAATCATCAATAACACAATGGCCTCACACGACCAATGAAAATAACTAACACAAGAATCCA TACATGTACCTTAAGGTTATGGCGTCTCAGTCGGACCCTTTTCTAGAGGAGGAAATAAGTAGGGATATCTAGACTTCATGTCTTCCTCGACCTACCAAGT CATTTTTTCCACATTGTTGTTTCTCCAGAGTACTTTCATGGAGGCTACCTCCTTATTCCATAGCTTGCAAATTTATCAGTCTAGGATGGCTACAAGAATT TCCTCGTATGTCAAGTCCTTTGTAATCTGTACATCATCCGTGGGCACCACTCGGGTAGGATCGCCAATGTACTTCCGTAACATAGATATGTGAAAACCCG GATGAACAGACTCCAATTCCAAGGGCAATTCCAACTCATAAGCTACTTGGCCCACTCTCCAAATGATCCTATAAGGCCCAATATATCGTGGTCTAAGTTT GCCTTTCTTGCCAAACCTTATCACACCCTGCATAGGCGACACCTTTAAAAATACCCAGTCATTAACCTCGAACTCTAAATCTCATTGCCACATATCAGAA TATGACTTATGACGACTCTGAGCTGTCAACAGTCGATCCCGGATAAGCTTTAATTTTTCTATGGCTTGCTGAACCAGGTCTGGCCCATGTAATCCAGATT CTCCAACATCAAACCACACCCTATAGGCGACCTACACTTCCATCCGTAAAGAGCTTCGTATGGAGCCATCTAAATACTGGAATGGTAGCTATTATTATATGC GAACTCAATAAGCGGCAGATGATCATCCCAGCTACCCTTGAAGTCTATCACACAAGCCCTTAACATATCTTCAAGGGTCTGAATAGTACGCTCAGCCTGT CCGTCTGTCTGGGGATGAAATGTTGTACTAAGACTTACTTGAGTCCCCAATCCTTTTTGGAAGGACCTCCAAAAGTAAATTTGAGCTCCTATATCTGAGAT AATAGATATAGGGATACCATGCAGTCGTAATATTTCCTTAATATAAAGCCTTGCATAATCCTATGAGGAATATGTAGTTCTAATAGGCAAAAAATGTGCT GACTTTGTAAGCCTATCAACAATCACCCATATCGAATCGAACTTACGCTAGGTATGAGGTAAGCCTACGATGAAATCTATATTGATTACTTCCCATTTCT AAGTCGGAATCTCCCATAGCCTACAATAACCCATCGGGTTTATGATGCTCAATCTTAACTTGCTGACAGTTAGAACACTGAGCAACAAACTCTGCCATATC CTTTTTCATTCCGTCCCACCAATATACTTCCCTGATATCATGATACATATTTGTCGCTCCTGGATAGAATAACGAGAATAGTGAGTGTCTCCCATA ACCTGACGATGCAGCCCTGCAATATTAGGGACACATAATCATCCTTGATATCTGAGGATCCCATCTTCTGTAACCTCAAACAGTGTCTTCTCCTTCTGAG GGGTAGTATCCCTATAATGAACTAACACAGGATCCTCGTATTGACGTTCCTTTACTTTAGTTACTAAAGAGGATATTGTCGTATCCTTAATAGTAATTCC AATATCACCTAAGTCCAGTAACCGAACTCCAAGACTAGCTAGATAATGAACCTCATGGGCTATTCCCCTCTTTTCTAGCTGTAAATATGACCGGCTACCC ATAGATCTATGGTTGAGGGCGTCGGCTACTACGTTTGCCTTCCTCGGATTGTATAAAATATCAACGTCATAGTCTTTAAGTAGCTCCCAACTATCTCATTT GATGTAAATTTAATTCCTTTTGCTTGAAGATATACCGGAGGCTCTTATGATCCGTATATATATATCAACATGAATGCCATACAAGTAGTGCCTCCCACATC TTTAGTGCATGAATCATCGTGGCTAACTCTAAATCGTGGGTCGGGTAGTTCTTCTCATGCTTTCTTAGTTGTTTAGAAGCATAAGCTACAACCTTACCAT GCTACATCAACATACAACCCAATCCCACTGCCTGAAGTGTCACAATAGATCATATAACCATCGGTCCCTTCTGGAAGCGTTAGAACCGGTGCTGAAGTTAA TCTGTCCTTTAATGCCTGGAAACTCCGTTCGCAAGTATCAGTCCATTGAAACTTTGCTCCCTTCTGAGTCAACTTTGTCAAAGGTGCTGAAAGGGAAGAA

	AATCTCTCTACAAATCTCCTACAATAACCTGCCAAACCGAGAAAGCTACGAACCTCCGTCGGTGTTGTGGGTCTAGGCCAAGTCTTTACTGCCTCAATCT TTTGTGTATCCACCAGGATGCCTTCACCTGAAATGATATGCCTAAGGAAAGCTACAGAGTTCAACTAGAATCCACACTTAGAGAATTTTGCATACAACTT CCCTTTTTGTAGAACTCTGAGCATAGTACGCAAATGATCTACATGCTCAGCCTCTGAACGAGAATACACCGATATATCATCGATAAATATAATTACAAAT AGATCTAAAAAGGGTTTGAACATATGGTTCATCAAATCCATGAATGCTGCTGGGGCATTGGTCAATTCAAACGACATAACACGAAATTCAAAGTGCCTGT ATCCGATCCTGAATGTTATCTTCGGAATATCTTCCTCTTTAACCCTTACCCGATGGTACCCGGATCTTAAGTCTATATTTGAAAAATACTTGGAACCTTG CAACTGATCAAATAAATCATCAATCATCAGGAGTAGGTACTTATTCTTGTTTGTCACCTTATTCAACTGTCTATAATCAATACACATTCGTAAAGAACCA TCTTTCTTCCTTGCAAATAACACAGGCGCACCCCACGGTGATGTACTGGGTTTGATAAAGCCTTTTTCAAGCAAGTTCTTTAGTTGTTCCTTCAACTCTT TCAGCTCTACGGGGGCCATTCTATAGGGAGGAATAGATATTGGATGAGTATCTAGTAGTAGGTCAATAGCAAACTCAATTTCTCACTCTAGCAGAAGACT CGGAAGCTCATCGGGAAAAATAGCAGAAAAGTCATTAACACAGGGATGGACTAAATGGTTGGTGACTCTACTTCCATATCCTGAACCCGAACTAAGTGAT AAATACAGCACTTACTGATCATCTTCCTTGCCTTGAGATAGGAAATAAATCTACCCCTCGGCGATGCCATATCACCTTTCCACTCCAAAACAGGCTCCCC TAGAAAGTGAAATCGGACTATCTTTGATCTACAATCAAAGTTGTCATGACAAGAATCCAACTAATGCATAACCATTATAACATCAAATTCTACCAAATCT AACTCGATTAGGTTTGCTATGGTAGATTGACCATGGACTACTATTATACAACCCCTATATACTTGCTTAGCCATCACTGAGTCCCCAANNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNTCTAACTCG ATTAGGTTTGCTATGGTAGATTGACCATGGACTACTATTATACAACCCCTATATACTTGCTTAGCCATCACTGAGTCCCCAACAGGTGTAGACACCTCAA AAGGTTTAACCAATTCAGGTTTTATTTCAAACTTAGTGGCAACCAACGGAGTAACGTATGATAAGGTGGAATCTGGGTCAATCAGTGAATATATATTATA TAAGGAGACTAATAATATATATATGTAACAACATCAGGCAATGACTCCTGATCCTGTCATCCTGTCAGCTCCTAAATGCGGTTCTGAGGACTACTCGAGC TAGATGCTCTGCCTCTGCCTCTACTACGACTCATTGGTGCTTGTAGACCTTTCCTAGGGGGCATACTGATGATGATGAACCAGCTATAGATCCCGCTGGC TGAGCTATGCTTGCATCACTTCTCATCGAACAATCCCTCATAACGTGGCCCGGATAACCACAAGTATAACAAACATCTAATCCCATACAGCACTGCTCGG TATGCTGCTTACCACATTGAGCACATCGTGGCAAGGGCAGTCGCATCTGACTTGACTCACCCCTATATTGAGAACTTGAGGCCCTGAAATTCTGACCAGA CCCTGAATATGTGGAACGATCAAATCTCTTACTGTTACACTGAGGGGGCATGCTAGTCGATGGCTGGGCTGGATACCTCAGGTATTTTTACCTTTGACCA CCTCGAAACTCACTAGAAGGACCCGAAGACCTCGCTCTCTTATTCTAGGCCCTATCATGCTCACGATCAGCCCTCTGTTTCTGCTTACACTCCTCTACAC CCTGAGCTTANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN

	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNAGAACTGCTCTGATACCACTTTTGTCACGACCCAAACCGATAGG CCGTGACAAGTGCTCGAGTCCTATCTGTCGAACACCCTGAAGCATGCGACTAAGATATAAACATGAGTAATATCCGCTGCATTACGAGAGTAACATGCAC GAAGATAACCTGTCTAAAAGACACATATACATATATGTGCGGATTACAGTGGGCGAGTCAACAAGGCCGCGATAGGAAACTATATACCCAAAACTGAAAG CCGACAAGGCCACATACAATCCAACTATAAATAATTGTCTACAGACCTCTAATAGAAATACAATTGTACAAGGACGGACTGATCCACGTCATACCCATA TATATACATATATATATACAAGCATATCGTACCAAAATCAAAAATAGCTCCGGATCAAGTGGAGCACGCCACCTCTCGGTGATCAAGGATCCTAAGAAGG GGGACCGTCAGCTTGCCTACCTGCACCTGCGGGCATGAAACGCAGGCCCCATGGAATAGGGCGTCCGTACGAAAAATGTACTAAGCATGTAAGACATGAA AATCAATACGTAAAAGACATAGATGAAGCATGAAATAAAGAAATCCATCTGTAAATCTAAATATTTTTGTAAATTCTGAAACATTTATAATGTCATGCAC GTGCATATAAATATCGTGTCATGCATAGGTATAGGTGTACATAATATCATAAAGCCTTTGAGGGCATCCCATCATATCGTCTCAGCCACTATGGGCAACA TAATCAACATATACCAGCTGATCAGGTGGTGGTGCGTATATAATGTCGTAACCTTTTCCCATATCCCATATACATATATTTACATATATACGCGTATATA ACACTATTTGGTCATGGGTCAGTATAAATGAATGCAATGCGTGAAATGCATAAGAAATAAGTTAATAAGATTTTTATCAGAATGCCATAAAAATCAATAT ATCTTTCGGATAAACTTTATCAAATACGTATTTTTCTGAGACCCATGAACAAAATATATAATAAGAATTCACATGGGAAATCGAGAATATAGACACTCCT CGTATTTCTATGAATAGAGTCATTTATGAAAGTTGTGTATTGTGCTCATTTCATTTGTATTATTTGGATCATGCCAAAAAGAAATAAGGCATAACCTTAA CATATCTGGAGTAGGGAAAAATCCCTGATAAATGATAAAGACAGATGCAGGTTATATGATCTTGCCCGACTCTTAACAGCATTAACCTTGTGAAGAGCTA ACTTCTCATTTCAATAACGCTCCTATCAGTTTAATCTAATTTCCTTGAAACTGCGGCTAGTTATACAAATTTCGAATGTTCAACTTTCACACACAGACGG TAGAATACTTGAACCCACTCCACCCTGGGCGCTTGAAAGATAGATCCTTGGAGAAAGATGCTATATATTTTTGTTCCTTAAAACAATGTTTGTGTGCTTA TTAGACACAGAATTTAGCAACTACTACTAAGCAATCTTGGGCAAAGAATGTCATTTCTTGGAAAAGAGAGACTTATTTTACGACAAAAAAACCTTATGCT TGTTGTCCAAAAACGAGAAAATGGAAACAAAGACTTGTAGTAAGTCTTACCTATTAACCTTAGAAGATATCCACCTAATGATACAATGACTAAGAGTCAT CCTCTTAGTTGGTGGCATTTTGCCACTTTTTGGGGGAGCGGAAATTAATCTTTATCTTATATTTTAATAATTAACTAGGTAATGTTTCGTTACCCGGTAA CTAACCAATTATCTCAAATTATCTAAAAATACTACTCATTTTTAATATACTTTATACATTATACTACCGTGGTTATATGGTACCTTGCATGGTACTAGTT CATAATTATCGGGTATTATCGCTCGACCCGTATTTTATCCCAAATTGGCCACTTTCAACGAAACTCGTTTTATCATGTACCCCTTTATCCTTTATGACAC TTACTTATAGCTTGTTGTAAATAGCGTAAATACGTTAACCTTAACATAATCTCATCCCCGAGTCTATGTCGATTAACTGAACACGAAACTTTTAACGTAC AAAAACGCGAGATGTAACAAATGCGCTCACTAGGGGTATGCAGATTTCGGAGGGGCTCCTACAGCCCAAGCGCTATATCGTTGCGGCGTGCAGCCTGATC CATATATATATATATATATATACACCCAATCCAACATATCGTTGCAGCGTGCAGTCCGATCCATATATCGCTGCGGCGTGCAACCCGATCCATAAATATA TAATCCTCACAACCAGGACTCGGCCTCTCTCAGTCATCAACCTCACAATCAGGCCCTCGGCCTCTCTCAGTCATCAACCTCACAAGCCATTCGGGCTATC AGTAAAACAGGGTGTTCATCCCAAAATATCATTTTTATGCATCAAAGGAATAATAAATAATGAGTTATGAAATCAATAAAATATAACATGACTGAGTACA GATCTTAAGTCAAAACAGTGAGGAAATAATAAGAAAAAACCCCTAAGGGTCAACTTGGCACGAGGCCCAAATATGGCATTCCGTGCAAAACATAGTAATA TTTTCCAAAACACAAGTGTATCAAAGAGTTTTCAGTCAAATACGCAACTTAAAAGTATCTACGGGATGGACCAAGTCACAATCCCCAATGGTGCATGCCC CCACGCTCGTCATCTAGCATGTGTGTCACCTCAATATAGTGAAACGAAGTAGAATCCGGGGTTTCATACCCTCAGGACAAGATTTACAATCATTACTTAC CTCAAACTAGTCCAAACTCTAACCCGCGATGCTTTCGTCTCTCGACTCGGCCTCCAAATGCTCCGAATCTAACCAAAGTCAATATATTATCATCAATATA CGCTAAAGGAACAAAGCCCAAGTGAAAATAATCAGATTAATTCAAAAATTCCAAAATTGGCTCAAACCCGCCCCCGGGGCCCACGTCTCAAAATCCGATA AAAATCACAAAACTGAAAAACCCATTCACTCACGAGTATACCCATATCTAATTCACTAAAATCCAAAATTAAATGGTCATCCCAATGCCCAAAAATCAAC TCTCCAAATTTTCTTCTATTTTTTCCAAATTTTCACCCCAAATTCCCAAATTAAATGATAAGAATCAAGACATAATCATGGAATTTAACTAAATTTGAGT GAAGAACACTTACCCCAATCATTCCTCTGAAATTCCCCTCCAAAATCGTCTTCTCCCGAGCTCTCTAATGAATTTTGAGATTATGGACCAAAACCCTCGT TTTTGAACTTAAACATTCTGCCCAGACATTTCCTTAATCGCGATCGTTGAAACACCTTCGCAATCGCGAAGAAAAACCTTGTTGCCTCAACATTTTACTC

TACGCGACCGCGTTACTCCACACAACGCGATGCAGTGGTGCCTCTGACTTACGCGATCGCATTTGACCACGCGAACGCGAAGGGTAACGCCAGCTTT
CCTCTGTCGCCTTATTTCTTCTACGCGAACGCGGCCTGGCCCACGTGTTCGCGTTCGCAGCCTGACTGACCTACACGATCACATCCCATTCTCTGTGAT
CGCGAAGAATTAACCCAGCTAACTTCCCTACACGATCGTGAACCTCCTCCCGCGATCGCGATGAAGAACACACCAGCTGCAGAACACCAGCA
AACCAGCAAACTCCTTACGTATAAAAATAACTCGTTGAGCATCCGAAACACACCCGAGGCCCTTGGGACCTCAACACAAACACATGCCAACACATCCCATAAC
ATTATTCAAACTTAGTCGAATGTTCGAATCACTCAAAATAAAATCAAAACACCCAAATTAACCTCGGATTCAAGCTTAAGAACATCTAAATTTTCCAAATT
CCATAAACAACGCCGAAACCTATCAAACCTCGTTCGAATGACCTGAAATATTGCACACACTTCACAAATGAAACTACGAGCCTACTCCAACTTCCAAAAT
TCCATTCCGACCCCGATATCAAGATTTTCACTGCCGACCGGAAACTCCAAATTTCCAAATTTGCCAATTCAAGCCTAAATCTACCGCAAACCTCCAAATC
ATATTCCCGATGCGCTCCTAAGTCCAAAATCACCTAACGGAGCTAACCAAATCATAAAAATCCAAATCCAAGATCGAATACTAAAAAATCAAAACTTGGT
CAAACCTTTCAAATTTAAAGCTTCAAGCTGAGAACCATTCTTCCATATCTATTCCGATTAACCTGCAATCCAAAACCGACGATTTACATAAGTCATAATA
CATCATACGGGGCTAGTCATGCTCGAGAACTGGCGAGCGAAGTGCAAATACTCAAAACGACCGGTCGGGTCATTACACTAGATAAATTATCAATATTAAT
AAACTTTTAGGTAACATCTTGACATGTGAAATTATCAATAAACTGCAAGTTAATTATGATACATGCTTTTATCAATAAATTTCTACTTTATTGTTGCATT
CTTTTAATTGTGTAGTACAAACTGTAAAATAAAACAAGTAGTTTTTTCACATACATATTACTCCCGCTACAAGTTTTAATAATAGACGATATTAAATTTTAA
CCAACCGCTTTCGTGAATACTTTAGAAACTCAATAAATTTCTTTGTGACTTACTACAACACAATACCTTATTGATAAATCAATTTTATTTCGCCAAAACAT
TATAATTGGCTCTTCCTATGATGACCTGATAGGTCATCTAGAGTTTTAAACCTTAATTCTATGTTTCGAAGCCTCCAATAGCTCCTTTTAGCCTTCTGCA
AGTTGCGTGCGCATTCCGTGTCTTTTTTTCGGAAGTTTTTATGTGAAAAATTGGTTAAAATGTGAATTTGTACCTTAAAATCCATTTGAGTTAACTCTGG
TCAACGTTTTGAGCAAACAGACCCGGATTCATATTTTGACAACCTCGGTGGGTCCGTATCGTAATTTAGGACTTGGGTGTATGCCCATAATCGAATTTGG
AGGTCGCTAACTCGAGTTATCGCTTTTTGTCGAAAATTTGAAGTTTAAAGTTTAAAATTTTGACCAAAGTTCTTTATTGCTACCGGGTCC
GGATTTTGGTTCCGGATCTTGGTATAGGTTCATTACTATATTTATGACTTGTCTGCAAATTTGGTGCAAAATAGAGTTGGTTTGGTATGATTCGAACATC
CGGTTGTTAAAATAAAAAATTCTTAAGTTTCATTAAAAATTTCATTTGATTTGGTGTACGATTCGTAGTTCTAGGTTTTATTTTGGTATTTTGATCACGCG
AGCGAGTTCGTATGATATTTTTGGACTTGTGTACATGTTTGATTTAGAGCCCCGAGGGCTGAGGTGAGTTTCAAATAGGCTTTGGAGTGGTTTAGACTTA
AGAAAATAGCCGGTGCTCTAATGTTTCTGGTGGAGGATTCAAACCTCGCAATTGCGAGATCTGGCTTTGCATTTGTGATATGGAAGACTCTCTGTCAGTT
TCGCATTTGCGAGCTACTTCCTCGCATTTGCAAAGAATAGTTAGGCACCAAGGGTTCGCAATTGCGAAGTGGGACTGGGGAGCAGAGTTCTCATTTGCG
ATCAAAATGATCGCAATTGCGAGGAGTCCAAGTTCGCAATTGCAAACAAACCATCGCGGTGGTAGCAGAGATGGGGGACCTTCGCAATTGTGAA
GTATTTCTCGCATATGCGGGGTTCGCAATTGCGAATCCCAGGTTGCAATTGCGATATCTGCGCCTCGACTTAAGAAGGAAAAAACGGGATTTTAGCTCAT
TCTTTCAAATTCTTAACCCTAGAGACCTTAGAGGCAATTTTTTCAAGAGCTGTTCTTCTCCAATTCATTGGTAAGTGACTCAACCTACTTTATTTTAAT
TACCCATTTATATTTCACCAGATTTAACCTTAAATCTAGGATTTTCATGGTAGAAATTGGGATTTGGTAGAAATTTTTGTAAAATTGGAATTT
AGACCTCAAATTGAGGTTAGATTTGAAACAAATTACATAACCGAGCTCGGGAATGAATGGATAATCGCGTTTTGGTCAGAATTTCATGTTTGGACCAAGC
GGGCCCGAGAGTTAACTTTTTCAATAATGACCTAAATTGAACCTCTTTCATTCGTGGGTAGTTCCTAAGGCTTATTTTGAATCGTTTGGTTGATAATTGC
TAGATTTAGTTGGTTCAGAAGCTTGTTCGAAAGGTTGTTCGAAAGGCAAAGCCGTGGATGAGCTTTGAGTCGATTTTGGATCTAGGTAAGTGTCGTCGTT
AACCTTGACTTGAAGGATTAGAACTTGTTTGCCTATTTGCTACATGTTTAAATGTTTGGGTGCAACGTATATATGTGAGATGACGAGTACTTATACGTTGTT
GGGTTAAAGTATGCAGGTGAGACTTGTTTCTTGTAATTTATTTCCTTCTTTGATTATGATATCTATGTTTAGACTAGTTTATTACTTAATTGATCGTTCT
TTTCCGCATTTATGAACTATTTGTGATAATTGAGTATTTTTGGAAGTTGAGGCTTGGTATTTCGGAACCGTTGTTGACGTAAGACTTATTCTTGCTTATT
ATACCTTCCAAATGTTATATATTCATTGCTACATGGTAAAGGAGAGTGTTAAGTACGAAGGGTGATACCGTGCCTATTCTTATTTTATTTGATTAA
TACATAGTGAGAAAGAGAGTTAAAGCACAAAGGGTGATGTCGTGCCGTATTTATCGTTTCATGACGAAATTGAGAGTAAAAGCATGAAGGCTGATGTCGT

GCCGTTATTATTATTTATGGTGAGAGTAAAAGCACGAAGGGTGATGCCACGCCATTTTTATTCATTATGTTTACCCATTTCTTTGGTTGAGG
ATTTATTTAACCGTATTATATTGTCATTCCTATTGTATTGAGAGTAAAAGCATGAAAGGTGATGCCGTGTCGTTATTATTTATT
ATGGTGAGGTTGAGAGTAAAAGTACAAAGGGTGATCTCGTGTGCCATTTTTATTCATTATATTTATCTCTTTCCTTTGGCTGAGGATTTATTTAATTGTCTT
ATGTTGTCATTCCTATTGTAATTATCGTATCTTGTCCCCCTTCCTCCTGTCCCCTCCCATGTCCCCTCCCAATAGTACGTATTAAATCTCTATTTGTTG
TTATGTTGCACGTATATCTTTATTTGTACATGTTTAATTACATGGATGTCCAGTCATAGCCTCGTCACTGCCTCGTCGAGATTAGGCATGACACTAACAG
AGTATATTGGGTCGATTGTACTCATACTACATTCTGCTTTCTTGTGCAGATTTTGGTACTGGTCCTAGCTGTCTATGAGGTGCATCATCTCAGATTAGC
TTATTTGGAGACTCGAGGTAGATCTGCTGGCGTCCGTAGACTTTGAAGTCCCCATCCTCTTTCCTCAATTAATGTTTCATTTCATTCGAGACAGTTGTATT
TATTTTAGACTCTATTTGTAGAACTTATTGAAGCTCGTGTACTTGTGACTCCAGATCCGGGTTGTACTTAGTTAATTGGGTTATTATGCTACTTCGTACT
TTATTTTAGTTTCATTTTTGTTTAACTGGTTAAATTGTTTAAAGTGTTTAAAGAATTACTCTAACGTTGGCTTGCCCAATGAGTGAGTGAAATA
TTAGGCGCCATCACCATTCCGAAGGTGAGAATTTCGGGTCGTGACACTGGATCCTAACAACTACTAGCTAGGCAAGCCGACGTGAAATAGAAAATTAACT
GTTTTAACAATTGAAAATAAAATATTAACAATGTTAACGATATAAGTCTGAAGTAGAATAAAATATAATACCGGACAACAAAATCTAAACATAACTCATAAC
CTGGAGTCATGAGTACATGAACATCTAGAGGGCTACAAACAAGATCTGAACAAAATACACAGTTTTGAAACTCAATAAAACAGTAACGATAAAATGGAAG
GGGGCTTAGGATTATGGACGCCATACACTATACCTCAAGTCTCCGACACAGCGATTAAATCCAAGCAAAAGTTATTGAGCGCCGCTAGGATCAACTCCAGT
ATTTACAAAAGAAGTGCACAATATAGTATGAGTACAAGTGACCCCGTGTACTCTCTAAGTGTTGATCCTAACCTCGATGAAGTAGTGACGAGGCTAAAGT
GGGCCACTTATACACAGTAATAATAATAACAAGAAAAGAAGGAAGTGGAAGTAAAGCAGTTAACTCATGAAAATGAACTCAACCTCATTATTAG
AAAGCCTAGAATAACGCGTCTCGTAATCTCATATAAAAATGTCGAGGCCAACTCAACAATAACAACAAGTACACAACATATACAAACTGTTGCGGCATGTAA
CCCCATCCCACCATATCATCATTTATCAGTGTCAAGTATCACATATACAATCCGTTGGAGGCGCGCACCCGATCCCACCATCATCATTTATGAATGTC
AAAATCCATCCTTGTTCCACCATTTCAATTCATTACCTTTCGTTGCGGCATGCAACCCGCTCCACACAAACAATTTCATTTAACATAAGCAATCC
GATAGCCCGACAAGAATTTCTACAACAAAGAGTAAAGGCAAAAGGTAATAAAGGTCCACATAATAATCAAGAATGCAACAAATATTACACAAAATAATAG
GACAAGTAAAACATATGACAATCAAAACGTGCAAGTAAAACAATTAAGGCAAGTATTAAATAAGCATGGAGCCATGGAAGAGGAGAAAAATTTAGCATG
AGAATAATTAAATGACAAATAACAATTATGGCATAGAAGTCAAATAAAGCATGTAACAAGATAATTCATGAAATAGCGAAAAACAGGAAAACAGATATA
TTGACGACGTATATATACACTCGTCAGCTCGCATATATACCGTCTTCTCACATAAATCCACACACAACATAGTTCAAAAATTCCTAATTCCCTCGAGTCAAGG
TTAGATCTATCAATACTTATCTCACTTCGCAATCAAATCAAAGCTCAACGGGGTCTTTCCTCTAGAACTAGTCTCCAAACCAATCAAATCTAACCAAAT
ATAATTTAAACAATTCAAAATAAGCTTTAAAAATTAACCACGAGTGAAAAAAGATTCAACCTTTAATGATTTTAGAAAAAAGTCAACCCCGAGCCCACTTG
GTCAAAATCCGAGATTCAGACTAAAACCCAATTACTCAATCACCCCCAAGCCTGGTTATGTGATTAGTTTCGAAAATCCGACCTCATTTTGAGCTCTAAA
TATCAATTTCTCAAAACCCCAATTTCTACATAAATCCTCAATTTCTACCATAGCTCAAACTTGAAAATAGTGAAAAATGGGTTAAATCCCAACTTTGGAA
TATTTAAGTCACTAGGCATCAGGCCTTCTTCGCCTTCGCGAAGGGCCTGACGCGATTGCGAAGAGCAGTGGCCTAAATGGCCTTAGCGTTTGCGATGGCC
TGCCTCCCCTGGCCTTCTGTTCGCGAGCCTAGGCTCGCATTTCGCATTTGATCAACCCCCCCCCACTCCAGAATCCCATAACCCTATGTGTTCGCGAGAGGC
TGGTCTTGTTCGCGAAGGTTAACCCCTAGTGCTTCGCATTCGCGACCAAGACATCGCGTTCGCGATGAAGGAATCCGCGACACTAGATATCAGTAAAACC
AGCAGAACGTCTAAGTCCAAATCGAATCGTAGCCTATCCGAAGCCCTTGGGCTCCAAACCAAACATGTACACAAGTGTAATAACATCAT
ATGAACTCGCTCGCGTGATCAAAACACTAAAATAACACCTAGAACTATGAATCGAACATCAAAATGTATGAAATTTTCAATGAAACCTAAAAACTTTTAA
ATTCACAACCGAACGTCCGAATCACATTAAATCAACTCCGTTTTTGCACCCTGTTTTACAGAAAAGTCGTAAAATAGTAAATGAACCTATGCCAATTCTC
GAAACCAAAATCTAAACCTGGTAGCTATAAAGTCAACCTACGGTCAAACTTCAGATTCCTTAAATCTTTAAATTACTAGTTTTTAACAAAACACGTTAAA
TCAAGTTAGGGACTTTCGAATTCAATTCCGGGCATACACTCAAGTCCAAAATCATGATACAGAGCCATCGAGATCGTCAAAATACTGATTCGGGTCCGTT

TACACAAAATATTGAATGTAGTCAACTCAAATCATTTAAGGCTAAAATTCATATTTTCTTTCGTTTTTCACATAAAAAATTTTCGGAATAAGACACGACT
ATGCACATAAATCTAGAAATACTAAACAGAGCTAATCGAGGTCTCTGAACACAGAAATAGAGGCTAAAACTCAAAATGACCTATCGGGTCATCACATCCA
GAGCTCTCAATTAATTATTGTACTATCACATATTCAACATTCATATGGTGAATATCTCTTTCAAGGAGAAAGTAATACTATTATGGTCATGGTTAAAATC
GTTATAGGCAATATATGTTACTTCCATTACTTTATCCTGTAATACTCACATCCAGAGCTCTCTATTGCCACATCCACATCCATTAATA
CCTTTATGGTAATAATTTTATCTTCTTTCAGACTTATCACACATCTACTACCACATTTTATAAGTGAATGAGAATGAAACAAATTCAATGGGATGTATTTC
CACTTCTTGTGCTCACAACCATACATAAATTTGATCATGACAAGACATAAACATTTTATCAATTCCGATGGTTATAATTACTAACAAGCTCTATTAGAGT
TATAATAAAAAATTATTGTTTGCTTTATATTTAAAATCAATTTTAACACAATTTAAAAGAGAAAAATAAGGTAAAATACTCATCCTAAATCCGG
AATTTATTCCTGAACGAAGTCCATGAAATAATTGACAATCATTATGCTCAATATTATACAAGTTGAGCACCATGCATGTATTCCAAAGCTTGGTGACC
AAAGCAGATGCCTAGTATAGGAATATCAATACATTTATGCAAGAGCCTCACGCAAATTCCTATATCTTCTCCATGAGTGGTTTAATTTTTCAAACTTT
AAACAACTAACTAATAGTATATCTACAAATTAAAACAACGAAGAATATTGCAGAGAAAAGTATAGAGAATTCTAATTACAAGGGGATAAAACCAGGGACA
TCTCAATAAGTTTGGTGGCCTAAAGCCAAACTTTAATGTGAGGCCTTTTTAAAAAAATCATATATATTTTTTAGTTGACATAGATTTTTCTATCTTTTGAG
AATAAAATTACTTAAAATTTCTCGCATAAACCATTTAATATCTCATGTGCAGTATGATCTTAAGTAATAGAATATGATTCAAGAAGTTGATAACTTTATA
TTAAAAGAATTTTTTTGATGTTTCAATATACGTGTTGCAATGATGAAAACCTAAGGAAAAAAATAGAGTTTAAAAGTGTATTTACATATTTTCTAATGATTA
CTGTAATCCTTTTTTTTTATATAAAAATATTTTACTTTTCTACTTTGAAATACTTAATATTTTCTTCAAAGATTTGGGCCTCACGAGAAGAGTTATGGT
GGGATGGATAGCAAATCTCCTTCTTTAATCGGAAGTATTCAGTTTCATTATTTGGAACGAAAAAAATCCCTTATAAGGAACAATTCCCATTTTTAATGAAT
TTTACACGGTACGAATTTGAATTAATTAATGTCCCAAATTAAATATCGGACATCAAATAAGAATTGGAAAATAAAAATTGGGGCCCCCAAAATTCCGGGG
CCTAAAGCCACCGCTTTAGAGGCTTTACCCTCTGGCCGCCCTTGGATAGAACCCCTTTATTTATAGAAAAAGTGACTTAGTCGTAAAGTAACAAACTCT
AAAATCTCTTTAAAAAAAAAATAGACATTTTCTATAAATAAGATACTATTTTATAACAGTGATCATATTTTTAACGGGTAGTCTTAAAAGTAGCTATTTATGC
AAATGGCCACATGTATATAATAAGCTAGTAATAAAGAAGACCCGGTTCATAAGTGGTCTCCTCCGTCAATGGCGAAGCTGAAGACGGATTCTCCTCTCTCG
TCGCATAGTCCTCTCCTCTTCTTGCGCTTTCTAGATTCTGGTACGCTTCTGCCCATTATCCCACTTATATATGTGTGTGGATGTATGTATATAATTGT
GTTGAAATGAATCCATTTACAAAATGTTTCTCTTCTCTGGACATAGCTAGTCAAGATCAGTCTATAGGAGTACTTATAGATCAGAATTTTATATGGATTA
AAAGGAAAAATAAAGCTTCAATGGGTATGTTGTTCAGCATTGAGAACATGGTGCTAAATTAGCGAAAATCAGTTTATTTCATAAAAGGAAATCAAGTTAT
CGGGCCGCTAGTATTAAAGTGGAACTACTGGCTGATTCTAGGCAGGGATGATAATAGCTGTAGTGGACTGCATAGATATAGTTATTGCATTGACTAAGTT
TCCCTTTAAGGGATTAGGATTAGAATCTTACAAGGGCAGCTCAATGGATTTGATGCTCTACAGCCAAACTTCAATATGAGGCCTAAAATATGTTTCTTAA
AATTCGTATATATTTTTATCTGATGTGACAACTTGAGAAAAAAGTGTGAAAATAAAATTATTCAATGATCCTTGTTATTG
CAAAACAGGCAGATAAGTGAGAAAGGGAAAGAAAAGAGATCTGGATTTCAGGAAGTACAATAATTGAGTAAATTGAATGTTGTGTGCCACCTAACTTGGGCC
GTGGGCACAACATTAGGGACTTGTATATAGCAACTGATGTGGTTCTAGGTTAGAAGGCATTGTTTATTGAGATTACATTTTCTTTTGAGAAT
GAGCATCCTACCAAGGTTTGAAGAGAACTGCAAATAGGAACTTCCTGTTGCCAATTTTTAGCATTTAGAATACTGTTTGAATTCCTAACATAGTTGTTAG
GCAAGTAAATCACATATCTATCTTCGATTTTGAGCTAAGATACAATTTGACTGGATCCCTATGGTATGAAATTGGTGTTTTGGGTGATCCATACAGCAAC
TGTAAATGAATTCCAACGAGATTGTGCTTGCATGTGAGCATCATAGATTAAAAAAGTGAAAAGAAGATCATAAGCCAAAAAAATTTGAAAAAGCTCCTTA
CTAACTGCTTTGAAAGAATAAACAGAAGTAAGAGAGAATCATAATGAGCAAATGATATTACTAGAAAAAAGAGAATAAGCATTAAGTGAAAAAGAA
AAAGCCCTAACCATCATTGGTCATAACCATCACTATTGTTGCCTCTCATCCCCCACAAAGTCTGCTCAGCCATATAACCATGCTTTCGTCATCCAGCCTT
TGGGTGAGTAAAGGCTGGCCACCCGCTTCCATGAACCGTTGTTGTTAACCATTTACAACCACTATACCTCTCTCCAACCTTGTTGCTACATCCTCTGGCA
CCTCCCTGCCTTGTTAATTTATTCGCGATTCCTAGTCTGGCACCCTGTCACTAAGCCTCCTTACTAATAACATTTATCCATGATTTGAGTTTGAGACTT

TTAGCTGAACAAATAAAGGTTGAATCTTGTGCATTTGGAGTGCGACCCTCTAACCATAGAAAATCCATCGATAAACGTCTACAACCTGCCCTCTTCTCTC
TTTCTCTCTCACTAGCTCATGTGCACGTATTACTTCCTCTTCCGTAACACCTTCACGATTGGGTGGAAAGCACTGCTTGACTGGGGGAAGAATTGGGAAA
TACCAAATCTTTTTTTTGGAAACCTCAAAGCTATGAATAAAGCTATTTCGATGTTATATTATTTGGTGCATGCTACTTTCTTGTTCTATTAGTCATGGTT
ATTTTTTCTGTTTTTATCCTGTTTAAATTGTGAAGTCTCCTCAAGTCTTACTTTCAACGTGGAGTGAGTGAATTTTTCATCTACTGACTCTGTCTATTCCTT
GTTACATAGTTGAACCTGCTTCTGGAGTTGATGAAGGACTTGAGGTTGTTAAGCAATGTCTATCAGAGGCCTTTAAGATTGATCCATTATCGGCTGG
TTATGCCTCAAGTTCAGATTCCTTAGTGGACATATTCAGTTCACAAGAAGCGATTGAGCAAAGCCAAATAAATTTAGATCTCAGGCGTGATGTATCTTCA
TCAGATGATCCTTGTTCATCATCAGGACAGAAGGTAGCAGATGCTAAAGATACAGATGGATCTCAGTTTCTGGTCTAATTCTACCTTTTTCTCTTGATA
TTTCTCTTAATTTTGTGTACTTATCTTTCCACATTTTGATATTTTCTCCACAGTTCATGTATTTTAGGTCTATGTTATGTTACTTCTATC
TACATGTTTTGGTTTCTCCTTTCTGTCTTCTATATAGTGTCCCATCAGTTGTGCAATCAGTAAACCTGTGATGTCTAAATTCATTCTGCTGTCTACATCTA
CATCTGCTCTTATTTTCTATCACCATATTGACAAATAATAGTTTCCCTTCATGGTTTGGCAAAAGTGGGATATTTGGGGCATGACCATGATATGTGCTTGA
TGCTATGGAATTTTCTGGATAAATTCAGATTTTGGTTAGGAAAGGAAAATTTAAACCAAATTCTGCTGCTAGACAATCTTTCTTGAAAGAACTTTAGTT
TTTAAGGATGGAGTTCGTTCAAGCTATCATTGAAATGATCATCAAATGCTTACATCTTGTGTGCTTGGCTGTACAGCCATGGCAGAGGACTT
ATTGTACATTCTCTGTTAATGTTGACATTTGCTATAATCATAAACCTATTTAGATCAAACATTAGCTACTTTTAATCTCCTTCATATTTTGCAACCAATCC
ATGAGTGAACTTTCAATCTGTGGTAGCCTTTTACTTTCCACATACTTGTGCATGATTTACCTTTACCTCATTTGTTTTCTCTAAACAGGCTGAAGAACG
GATAAAAGAAGCTCCCACTTTTGGTAAGGCTAGGCTCCAAGTCCTAGATTTTTATAATATCTCAATGAAGGCCTCTGTATATTGTCTCTCAACTAGTGGG
CCACACTTGACCTTCTTACTGTTACTCTTTCGTGTCCATACGTTTGTGCTGTCATTCAATTTTTATCTCACGTCATCTTCAGTGACGTTCAAGGCTTGCT
TGTAGCATGCTTAAGCTCTGAAGTTATGTGAAGCCCAAGTTAGAGGGTCCCCCTTACTTAGGTGGCTCATGACTGCAGGCTCCAAGGGACTAAGATATAT
GCCTCATCGCCTTGAATTCTGTCGGCTACACCTGGCAATATATATAGTTGGCAAAACAATAGAGTTATTATTAAAAATTGTGATTAAGGAACTTTTTGG
TGGTCAGGAAATGAAACTAGAGAAGTGATATACAAGCCCAGTCACTAAAAACTTAAAAACTGTATTTTTACGTCTAAATAAAAAATTTTAAAATGGTTTT
TCAACTTGCTTGACGTGATTTTTATTTCATATATTTTGCCTAGTTCTATTTTTATTTTTTAAAATTCGTTTATCCTAAATGATCATTGTTTGTGTT
TATAGTGATTCTTACTTGCTACATATACTTTTATGGCAGCAAGTAAATAATTATTCGCTTTTTCTTGCATATATTTATATATGTAGGTCTGATATATGTG
TTTGTATTTTATTTTTCTTTATTAGCAATTTCTTCTCTAAAGCGCACGAATTGAATTGAGCTTTGCATTGAATGCCCTAACATGTTTTTTGCCTTTTCAA
AACTCTGGTCAATCTTTTTGGCAATGAGTTCATTAATACCTCAAATCTAGTTGCCTTTTTACTTTTTCTCATAGTTTCACAGTAATTCATTATAC
TTAAGCAACCATGAAAAATAAGGTCAAAATTTACACACACACATATGTATGTGATATGCATATAATATGTATGTATGT
ACGTATGTAAAGGATAGACTAGGGGTTCATTTTTTTTTGGTTCTGATTGTGGTGACACACAATCTTGTAATTATTTTTCTCATATATATAATCGTTACC
TTTCTAAAACAAAAAAAAAAAAAAAAAAAAAAATAGCAGGAGATATAGAGGAGTTTCCAGTAACTACGGATTTACACCTGGGGTTTGCCTTGAACCTTTACTTGTTT
GTATAGGTATGGATGAGTTAACCAATAGTATACATTGTGAGGTTTCATGGTGAATGCTAGTCGCCAGTGATATTCTGTCAAACATGCTATGTAGTCAATA
AAAAGTTTGAATTATGGAGAAGTACTCTTGAAGATAACAATTTAGATTCTCTACTTCTTGTATGCATCTTTTATATTGGAGACTTTCCTCCCTACTGCAT
GTCAATAAATTCACCTTAAAAGAGAGTAACAGTTTTAGATAAGTATATAAATTATACAAACTCTATACGCTACAAATTAAACCCATAAGATAAATGAAGGTG
AAGTCTCGATTTAAACGAAATGAAGGTGAAGTTGAGACAAGATTGTGGTTTCTAAATATAGAAACTCAGATATCTAGTTAAGTCTTCCAGCAGAAT
GACATTATAGATGAAGATGTAACACATTAGATTAAATTAGTATGACTAAAATGGAGTGCTTTGATAAAAGATACTTAACGAAGGGAAAAGCAAGTTTCAT
AAAAAAGGGTTATCATTCTGCAATGTTATATAGGGGTGAATGTTGGTCCTCTAAGTCCAACATGTCCACAATATTGATGTCCCAAAAATGCGAATGTTAA
GATGAATGTACAGTTATACTAGATTGAGCAAGACTAGAAATGACCACACAACCAATAGAGGGTGTAAGTAGCACACACAAAGGGGATAAAATGGTAGAAGCTCG
TCTAAGATGGTTTGGCTATTTCCTACATTAACTGAGGTATGCATCGGTCTGTAGGTACTATGATAATTGAAGGTGCTAAAAGAATTACAAGTTAGATTTA

AAATCAAATGGAGAAAGTTGTCTCAATAGACCTACAACATCTCGTAATAAATACGGACTTAGCTAAGAACGGAAGAAATCTGAAGTAAAGAAATCTATCGA
GACTACTAACTAGTTGTATTAAGGTTTAGTTGTTGTGCAGCACTTACATTAGGTCTAGTGTTTGCGGGGAGTCTTTCTATTCTGGATAAGATGTACATGCCT
GTCTTTAGGAGAAGCAACTGTCACTCGTGACAATGGATGCCTGCTTATATTAAACTATCCTCAATCTCTTTCTCTAAAAGCTAGATTTTTCTCTTCCCTT
CTCCTTCGACCCAAGCAAAAGGAGTATGCATTTCCTCTTAGTGTGAGAAAACGAGGTTGAAATCTAGTTCTATTTATCAATCAAAGAGGTAGAAAACTAG
ACCTACCGTCAAAGGTTGGGTTCTCTAGGGGTTAGACTCCTATGCTTTGCATTATGTAAGTTGGGCTGCTCCAGCTGAAGTGCATTAGAAACTTTGTCC
GTGGGATTGTCGGGGAAAAACCCGATATAGGACTCCTTTTAATCTGAATAGAGACGTACGCTAGTATACGGATAAGTGTGATGATGTAATGAACCTATAG
ATTGGTACCATATAGACAATTTTAGTATTTGAATGAAATGGGCCGAGAACAGCGGGGTGTTTATAGAGCTCATATAGATTACATCCAACTTGTTTAAGAT
TTAAGTGTAGTTGATTGATTGTTAGATGTAACAAAGTCACATCTGTGTTGCTATTTTTCTGGTCCATACGAAATCATTCTTCTAAGCT
AACTGTCTTGCTATTTTCTAGCCCATCTTCTTTGTACTCCTATAATTCATTTGTATTTGGGTTTATATTTCCCTTAGAATGTGGCTGTGGGCCTTTAGT
TAAAAATGTGGATGTGGGCCTTTAGTTAAAAATGACATACATGTAGGCTACACACATGGTTACAAAGTATAGCAAATATGGCACCTACTTTTCTTGTAA
TCCTTTTTTTTTTTCCCTGTTAGCTATTGATGCACGTTCTTGGGTGGACTTTTCTTCGTCATACCAAATGCATCATCCAAATGTTTTAAACCTTGTTTTTC
CTTCTTTCGTTGATTATCTTACAAGATATGTTTGTGACAGTGGCAGGTAAAAGTGTCTCTCTTTCTGATTTGCCCTACATTCTGTCCTGCTTTATGTTAA
TCAACTGATGCTGTTTATCATGCTTTGCAGGTATATCCAAGGATGAGCTCTTTGGGCAGTTTTTTTGATGCCCTTGAAAAAGCTCACTATTTCAGAAGCTT
GCCAGATGGAAAAGATGATCAGGCTCAGCTGGATAGAGGCATCAAGGGTGTTTCACACTGCTTTGGAGGTATATATTGACAATGAGTTTATTTTATTCATGC
TTCTACTAAAGTGTATCAGTTTTCCATACTTGATGGAATTGGTTGAGAAAACACAGCATTTAAATTTCACCTGGTGGTTTCCATTTTTTCTGTTCTACCA
TGAACAAAACTTTGCTTCAAAATGGTGCTTTTTCCAAACATTGATGACAGTATCTTTTTTTTTAATTTTTTTTGAAATGATAACATTTGAATATATTATA
AAATCAGTACAACCTTTGTACTGGAAACCATATTTTACAGGTAAATGCTAACTAGATTTGCTAGTGTGCAATCCTAGCAAGATCCTAATATGTCTATGATT
GACAATGTATTTTCTGTGTAAATCTGTTTACACCAAAACAAAAAGAAGAATACAATTGAGTTTGATCTTCTGTACTGAGTTACGTCTGTCTTCAAAACA
TCTAGCATTCCTTCCTTCCTTCCAAACTGTCCACCAGATGCATGCTGGGACAGTCCTCCATTTATCTCTGTTGGTTGCTTCAGTTCCATTCAGCTCCAGCTTC
TTCCCAACTAAAAAGAACTTCATTGATCCTATGGGCCATTGTCCATGAAATACCCCTAAGATTAATAAAATTTTTCCATAGTTGGTCTGTTATCTTGCAG
TGTAGAAATAGATGGCTAACTGTCTCCACATAGAAAACATCTTGAGCACAAGAGATTCCTCTTTTTATGAGATTATCCTGAGTTAATA
CAGCCTCCTTTGCTAGTAGCAAAGTGAAACAAGCTACCTTATAGGGAATCTTAGTTTTCCATATATGCTTCCATGGCCAATCTCTTATCTGTGGGTTATT
TTGGTTCAAAATCTTATATGCTGCCTTCACCTGATAGACCCCTCTGTCATGCCTCTGCCACCAAAGTGAGTCCACCCCTTCCTGTATTCCTTTAAATGTC
TCCAGCTTGTTATAAAGATCAGTTAATCTTGCTATCTCCCAGCCATTCAACTTCTTCTTAAGAATTAGTTTCCCATCCACATCTCAGCTACTGTCTTCTGT
TGGTCTTGTGCCAAACCAAACATATCAAGGTAGAGTTCCTTTAGGCATCCATGCCCCAACCAGTTATCATTCCAGAAATGAAGCTTTCCTTCCATTCTATT
CCCCATATGATTAAAAAATGCTTTCTTCTAGAGTAGAAGGGCTCGTACTCCTTTTGACCAGCTAAAATTATTGCAGTTCTTGTTCGAATATCTTCCT
TGCATGAAGTCCTTCTATTCCCTTATGACATCACTATGACAAGGTTCGCTTGAGTGCGGTTAAGGATATTTTCGATTAGCTTAGACGAAGCTTCAATGCA
TTCACCAGGGCACTAAGACATGCACTTCATCCCTACGCTAAACTCTCTAAAGGGCTCATATTTCAATTGATGAATTATATTTTCTATGAATAGGAATAGA
CATTGAGTTATGTTACAGTTTGTTTGTTCATATGCTAGCTACTGTTGCTTCTGTCATGATTTGTCAAAATTATTGCAGTTCTTGTTCGAATATCTTCCT
ACCACCTTTATCGTTTATGGTTTTAGGAAATGCAAAAATCAGGATGTGCTATGTTAAATCGAAACAACCTCGCTGAGACCTTAAAATCACAAGGTACTTT
CTATTTACATGTTTCTTTTCTTTAAGGGTAAAGATATTTTATTGATGAAACAGCAAAATGCTGAGATTACCAGTATCATAAACTATATCTCGTTTGATGT
AAGGAGTCAATTAGATGTAGGAGTGTCAAAATCATTTATGTTTTGCATTTTACACCAAAAAAAGAATATTGCAAAGAGCTATACTTTATGCGGTGTATGA
CTTAAGTCTTCTTATCGCAACATCTGTATTCCTTTCTCTCCATACATGCGAGAAGGCACTGTATTCCTTATTTCTTCCCTTTAAAACATTCTATTGCTT
GTTTCACAATGTTTGGTATCACCCATGGCGAACAAAATATATTTAGAAATAAACTCCAGCCTTGCCGGCTAACAGGGAACTTTAACAATAAGTGTTTTGC

AGTTTCCTACGTTTCTTCACACTAGTAACAGCTGTTAACTAAATTGAGGCCTTTCTTCTGCAAGTCCTCATTGTTCAGCAGGCTTTATATGCTGCAATCC
AAGTAATCTATGTTTAACAGGGTTTTGATTTTAGGGCCATTGTCTTCTTTCTTCACAGAACCATTAATACTTCATAGCAAGATTCC
ACTGTAATTCACCGTCCTCTTTTTAGAAAAATCGTAATTTACTGTACTTGCTGTCCAAGTCGAAGAGGTTCAATAGTAACTTGTTAGCTCATTCTAGCT
TTACCAAAGAAGTAGGTTATGCTTTCTATTTCCCAATCTTGAATGTCTCCTGAACCTGATGTCCCATTTATTCTCAATTATACAATCTTCAATATAGGCC
TTCTTGTCCAGTGCCAAACTGTTTAGCCAAGGAAAGGCCGTCTCAGTATATTTTGCTATAGCCAGAAATCTCCCTAGTACTTGATTTTCCTTTCATATCC
AACTTTCAGCTTTACAAGATTGGTGAAATGTTCCTCTTTCTCAAAAAAAGAAAAAAAGATTGGTGAAATGTTCAGCCCCTTTGTATGTTTCCAAACAGAT
ATTTAGTTCATGAATCCTCGTTATTGCACCCCAAGAGGATGGTCTAACCGTCGATGAAGCTGGAATAAGGACCCCAATCGAGGCAAAAAAAAAAGTTGCG
AAATCCCAGTGGAATTTTAGAAGAACGCGGGTAGGGCAGGCTTGGAGTGCTCACCGAGTTGTTTAGTGTCATTTTTAGGATGAAGAAGATGCCCGACGAA
TAGAGGTGGAGTACGATAATTCTGTTGTACAAGAATAAGGGTGATATCCAAAATTGCAACAACTATAGGAGTATCAAGCTGCTAAGTCATACTATGAAAG
TTTGGGAGAGGGTGTTAGAAGTAAGGGTGAGGAGGGAGTGTGCTTATTTCTGAGAACAGTTCTGATTCATTCCGTACCGAGTCGTTTGACTACGGAAGCCA
TTCACCTTGTAATAAGATTGGTGGAGCAGTATAGGGAGAGGAAGGATTTGCATATGGTGTTTATTGACATGGAGAAAGCATATGACAAAGTTCCGAGAGA
GGTTCTATAGAGATGTTTGGAGGCTAGCACTGTGCCTGTAGCCTACATTAGGGTGATTATGGAGATGTATGATGGAGCTAAGACTCGAATAAGGATAGTG
GGAGGAGACTCGAAACACTTTCCGCTAGTGATGGGGTTGCACCAAGGATCAACCTTAGCCCGTTTTCTATTTTCCTTGGCTATGGATGCACTGGCGCACCA
CATTCAAGGGGAGGTGCCATGGTATATGTTATTTACAGATGACATGGTTCTGATTGACGAGACAAGATGTGGTGTTAACGAGAGGCCGGAACATTGGAGA
TAGACCCTAGAGTGTAAAGGTTTCAAGTTGAGTAGGACCAAGACATAATACTTGGAGTGCAAGTTCAGCAACGTGACCAGTAAGCGGATATGGATGTGAG
GCTTGACACACAAGTCATCATTAAGAGTGGAAGTTTCAAGTACCTTGGATCAGTAATCCAAGGAAATGGGGAGATTGACGTAGATATCCCACATCGTATT
GAAGCGGGGTGGATGAAATGGAGGCTCACATCCGGTATCTTGTGTGGTAAGAATGTGTCCCTAAACTTAAAGATAAGTTCTACAAAGTGGTGGTTAGAC
CGACTATGTTGTATGGGGCAAGTATTGGCCAGTCAAAAACTTTCATGCCCAGAGAATGAAAGTAGCTGAAATGAGGATGTTGAGATGGATGTGCAGGCA
TACTAGGTTAGATCAGAGTAGGAATGAAGATATTCGAAAAAGATGGCAGTGGCCCTCGTGGAGGACAAGATGCATGAGGCGAGGCCAATATGGTTCGAGC
ATGCGAAGAGGAGAAACAGTGATGCTTCATTTATGAGGTGTGAGATGTGACCTTGGTGGATATGCGGAGAGGTAAAGGTAGAAGTAGGCCAAAGAAGTAC
TAGGGTAAGGTGATTAGGCAGGACATAACGCAACTTCATACTGCCGAGGACATGATCCTTGATAGGAGGGTGTGGAGGTCGGGGATTAAGGTAGAAGGTT
AGTAGGTAGTCGAACATTTTTCCTTTTCTTTCCCATTGCCGGTAGTATTAGTGTTACTATGGTATCTTCCTTATTCTTAGATTGCTATTACTACATATTG
TTTGATTGCTATCTTTTGCTTCAGTTTTCTTAATATCTTGATGTTGTTCTCACTGTTTTTTTTCCCCACTCTTTCTTGCGCCGAGAATCTA
TCTTAAATAATTTCTCTGCCCTCCCAGGGTAGGGTAAGGTCTGCGTATACACTACCCTTCACTGGGTTTGTTATTGCTATTGTTATAGAGCTAG
GTATTTTCTCCCTATCTGCCTAAACTTTAGTAGGAAAATGCCTGTACTGGTGGGAGGATAGGATGCAGATACTCAGGTGAAATAATTGAGTTGCGCGCAA
ACCGGTCCAAATACCGCTGTTATAAGAAAAATGAGACTTTACTAATCATTTTGCTATATTTGGACCCATTTCATTCAGGTTCACCTAGTTTCGAGAAAAA
TTTAGGAAAAGACAACTAGTTGAATCACAATTTCCACATGAATTATCATTATTTCATTCTATAGGGTAGCTCTGCAAAACATTTTAAGTTGCACTCAGGC
TAAGCTAGTCTTGGCTCACTGTAAGCCCAACTTTCTTCCCGGCCTTAATTCAGCACATCGTCCGTTTTGCTGGGCCGGAGGAGTGAGCTTTGGCTGCCTC
AATAATTGGGAAGAATAATGTTGTCAAAAGGTCAGTTTCTAATTCTAAAGCAATTGATTGGCCTTGAATATCTCACAAATATGTACTCAAGGGTGATGTC
TACTTGCACAAGTCTTTAGTTCCTTCCAATGGTGAAGCTCCCTTCTTTATGAAAGCGTTGGCTCAAATTTTCCTCTTATTTTTTTTTTCCTTTCTGGG
ATATTATAATAGGTAATAAAGCCATGCAGTCAAAGCTTTACTCTGATGCTATTGAGCTGTATACTTTTGCAATTGCACTTTGTGAAGATAATGCAGTTTA
CTATTGCAACAGGTGCGTCATTGTCACAGTGATGGGGGAAATTGATTATTGCATTTTCCATGAAAAATTAGTACTTCCTCTATTCCATTTCCTGTGACAC
TATTCACTTTTCAATCTTCCGAAAAGAATGACACCCTTTTATATTTGGAAACAATTTATACTTCACATTTTACCCTTAATGACATGCTTTTTATAGCCATA
AATGTCACGGCATTTTTGAAACCAAAAGTCTTTCTTTGTTTCTTAAACTCCATGCACAATCAAACAGTGCCGCATAAAATGAAATGGAGGG

	AGTTCTATTTTATCTTTCTTTGAAAATTCTAATAGCCGTAAATTGAATCTTTTTAACCTGTAAATAGTTATCGAGACCTATTTTAAGATGAGGTTTTTG
	TGATGTATCAGGGCAGCTGCTTTAACCCAAATCCACGAATACGAAGCAGCAGTTCAGGACTGCCTAAAGTCTACTGCAATCAACCCCAATTATAGCAAGG
	CGTATAGTAGATTGGGTTTTGTCTATTATGCTCAAGGGAAGTACCGTGATGCCATAGATAAGGGATTCACACAAAAGGTTGGATCAAGAGTATCTTTTTCCT
	TTACTACCTTATCAAAAAAATATCTTTTTACTTTACCAGTTATTCAGTTTGTGTTGTTTTTAAATTAACATCATCTTGGTATTTTGGAGTACATTATAGTT
	TTGAATTTATCGAGGATAGTCATCATTAAGCTCTGACTAACTTTGATATGTGTTGTACTAGGATTGTTTGTAGTAGTCGTTGTTTCTCATGCTCTATAA
	TTTTTTTCAGTTAATTTCATTAGTTCATTATGAAATTTGATTTGTAGTTCTTCCCATGACACATAGCAGTTCTAAAAAATGTTTTGGTGATTGCTAGCCA
	ACTGTGGGATTTGGGATTTGTCAATGAAACCACTATACATAGAATTCTTTCAGGAAAGCTATACTGAGTCTTACTTTTCCTTTTTATTGACTTCTCTGAT
	CTCTTTTGATAAGGTTTACAATATTTAGGTTTTATTGGTTCTAATTTTTTAAAATGTATAAGGTATAGTATTCACCTGAATTTGTTTTTGTTGGAATTAA
	CAAATATGAAATATAGTATGATGTGGGTTCCTGTAAAATTGCAAACACAAGGAGAAATGTAGTAAATATGATGACATATTATGAAGATGACTCGTAGCATTGA
	ATAATGACTTTATAGACCTCAGCTCACATTACAGCAACTTTCTGAAATCCGTGCTTATTTGGAGTTTCCACAGCCTGATATCACGTTTTTCAACAGGTTTA
	TTTAAGTCCTCTACTTGAAAAGGAAAACCGAGTGTTGGCCAAAATGTCAAAAACACAATTGGATGTCTGACAAAAACAGATCCTCCTACTCCATTAATGA
	AAAAATTAAAGCACCCAAGGGTGTGGCCTAATGGTCCATAAATCGGGAGGTGAACTATGAGATCTTCAAATCCCAGCGGAGACCAAAACTGCTAGGTGAT
	TTCTTCCCATCTACCTAAATCTTGGTTGGCAGGGTTACTCGATACCTATGTTGGTGGGAGGTAGTAGGTATCTTGTAGTACCCACTAAAGCATGACTACG
	AAAGTCAACTAGAAGCAATTTAAAGAATAGAAAGAAAAGAAGTCGAAGAAACTGCTCTTGCCACACAACCGATGTCCACCTCTAATGCCCTCCCTCCGT
	TATTTCTTTCTTCCTCCTTCCCTCATCAAGAGATAGTGAATGGAGTTCCCACCCCAAACAATCAATAATTGGATAAAAGAGAATCAAGACGGAACTGGCC
	TTACACATGCAAACTGGGGCATGGACTGATCTGTGGTTACGAGGGACTTGTTTGGGTGAAATGAAGATCTGTGTATATTATGTGTTGTATTGGTGCTC
	GGTAATTGGTCTGGCTACTTGATCAGTATTGGTCTTGTGAAGAAGTGTAAAGAGGTTTTGCACGTTATCTCTGTCATCCATGTTAGAAATTTTGATGAGC
	AGTACCAGAATACGAACATGTGAATTTTGTTGTCTAGTTAATCCCCAAGAAGAAATCTATCTCTGACGTAAACCCTTTTTTGTGAAGCATTGCAGTTGGA
	TCCTAACAATGATTCTATCAAAGAAAATATACGGGTTAGTTTCTACTGTAGATTGCAGTTAAGTTTATAGCTGAAATGGTGCTTGCATATATCCAAGTAT
	TACTCTCGTATGTTAAACTTGTAGGTAGCAGAACAGAAGTTAAAAGAAGAGCAACAGAAGAGGGAGCATGATCAGGTAACTGACAATTGTCGCATTGCAG
	AAGAGCTTTGTTAGATTTATTAAGTTGATAATTACAATTATATAATTAACACCCAGACAAGGGTTGAGCTCAATGCTACCGATAGTGATTTGTCTGTCT
	GTTGTTCTCAAAAAAATAAAATAAGAAATTTGATATGTCTTGTCTGTGGTAGCACCATCTAGGGTTGAGCTCAATGCTACCGATAGTGATTTGTCTGTCT
	GTTGTTCTCAAAAAAATAAAATAAGAAATTTGATATGTCTTGTCTGTGGTAGCACCATCTAGTACCGCATTATTGTGTGAGAAAACTACAATGAACATGT
	ATGATGCTTACTCATTATCTTTCTCTTTAAAAAGAAAGAGAATGGAGCATGTCTCTGTAGTTCAGCCCTCTGACAATGTGATTATAAAATTGGCTGCGCA
	TATCTAATATTTGATTCCATTTATCCATTTACCATCTAATTTAGCAACGGTGAGCATTTAACTGAAATTTCATTTTGTATCCTTTGGTCTTTTGCAGAGC
	TCAACCTCTGGGAGTCATGGACACAAGAATCAAATCAGCAGCCTGCTGGTGCATCTAGAAGTCATGCCATGCCTCCACCATTTGCATCAATGTCATTCGATG
	GCAATGGCATTTCTGGCATTCCTGATTTGACTAACGTGTTTATGAACATGACAAGAGATGCATTTCAAGGGCAGCATGGTCCAAACGGGTCTGAAGGAAG
	CAGCGACAGTGATTCAACTAATGAGCCCGGAATAAGATTAGGCAGGAACATCAATGTCAATTTTGGTGAACAGATGCCTGAAGAATTGACTGGGGCCTTG
	AGGTCTGTTATGGAGATGTTTTCAGGAGCACAACCTCCCAGGAATCCTCAGGACAGTACGAATGGA
17	>b2b2638e-710e-4167-819a-ed479515614 7 (sequence:mRNA) 1206 residues [Nitab4.5_0001803:310317-343182 - strand] [cds]ATGGAAACAAAGACTTGTATTGAACCTGCTTCTGGAGTTGATGATGAAGGACTTGAGGTTGTTTAAGCAATGTCTATCAGAGGCCTTTAAGATTGA
	TCCATTATCGGCTGGTTATGCCTCAAGTTCCTTAGTGGACATATTCAGTTCACACAAGAAGCGATTGAGCAAAGCCAAATAAATTTAGATCTCAGG
	CGTGATGTATCTTCATCAGATGATCCTTGTTCATCATCAGGACAGAAGGTAGCAGATGCTAAAGATACAGATGGATCTCAGTTTCTGGATGAGCTCTTTG

	GGCAGTTTTTTGATGCCCTTGAAAAAGCTCACTATTTCAGAAGCTTGCCAGATGGAAAAGATGATCAGGCTCAGCTGGATAGAGCATCAAGGGTGTTTCA CACTGCTTTGGAGGAAATGCAAAAATCAGGATGTGCTATGTTAAATCGAAACAACCTCGCTGAGACCTTAAAATCACAAGGTAATAAAGCCATGCAGTCA AAGCTTTACTCTGATGCTATTGAGCTGTATACTTTTGCAATTGCACTTTGTGAAGATAATGCAGTTTACTATTGCAACAGGGCAGCTGCTTTAACCCAAA TCCACGAATACGAAGCAGCAGTTCAGGACTGCCTAAAGTCTACTGCAATCAACCCCAATTATAGCAAGGCGTATAGTAGATTGGGTTTTGTCTATTATGC TCAAGGGAAGTACCGTGATGCCATAGATAAGGGATTCACAAAAGCATTGCAGTTGGATCCTAACAATGATTCTATCAAAGAAAATATACGGGTAGCAGAA CAGAAGTTAAAAGAAGAGCAACAGAAGAGGAGCATGATCAGAGCTCAACCTCTGGGAGTCATGGACAAGAATCAAATCAGCAGCCTGCTGGTGCATCTA GAAGTCATGCCATGCCTCCACCATTTGCATCAATGTCATTCGATGGCAATGGCATTTCTGGCATTCCTGATTTGACTAACGTGTTTATGAACATGACAAG AGATGCATTTCAAGGGCAGCATGGTCCAAACGGGTCTGAAGGAAGCAGCGACAGTGATTCAACTAATGAGCCCGGAATAAGATTAGGCAGGAACATCAAT GTCAATTTTGGTGAACAGATGCCTGAAGAATTGACTGGGGCCTTGAGGTCTGTTATGGAGATGTTTTCAGGAGCACAACCTCCCAGGAATCCTCAGGACA GTACGAAATGGA
18	>b2b2638e-710e-4167-819a-ed4795156147 (sequence:mRNA) 402 residues [Nitab4.5_0001803:310317-343182 - strand] [peptide]METKTCIEPASGVDDEGLEVVKQCLSEAFKIDPLSAGYASSSDSLVDIFSSQEAIEQSQINLDLRRDVSSSDDPCSSSGQKVADAKDTDGS QFLDELFGQFFDALEKAHYFRSLPDGKDDQAQLDRASRVFHTALEEMQKSGCAMLNRNNLAETLKSQGNKAMQSKLYSDAIELYTFAIALCEDNAVYYCN RAAALTQIHEYEAAVQDCLKSTAINPNYSKAYSRLGFVYYAQGKYRDAIDKGFTKALQLDPNNDSIKENIRVAEQKLKEEQQKREHDQSSTSGSHGQESN QQPAGASRSHAMPPFFASMSFDGNGISGIPDLTNVFMNMTRDAFQGQHGPNGSEGSSDSDSTNEPGIRLGRNINVNFGEQMPEELTGALRSVMEMFSGAQ PPRNPQDSTNG
19	**>Nitab4.5_0001961g0080.1 Nitab4.5_0001961:372401-373786** TAGACGGTGAAGAAATCCATTCAATGGACTATTCCAAAATGGATTCAACGACTGCAACTAACTTTGTCAAAGGAAAACATGTAGTCTGTCATTGGGTCCC ATAAATCAGGAATGGACATTGCCATCCACAGTTAATGGTAAGTACTGCTACTGATCTTCCATACACCTTCATTACCCGAACCTCTTTATTGTTTGTTTAT AAACATCAGTTTGATTCTGTTTGCACTAGTTTTTGCAGGGATTGAACGTCCATGCACAGTAGTAACCAAGACACCACATTGGAACATTCCCGATTATTTTC CATAGGGATTGCCTTTGGCATATCTCTACCTACATTGATTCTCCGAACTTATGGTGCATAAACCCGGCGAAGGCCTTCTTCTAAGTCTCCTCGCCACAAC TCTTTCACCTTTGGTAATTGTTTTCTTCGTGACTAATTAAACTTATATAAAAAGACTCAACTATGTGCAGCAACAATTAGTCATACTTCTCATTGAAACTA ACAAATGCATCTGCAAACAAAGCCAAATACTTTAAGGTTGAGTAATGTTCATTTCTAGCTGTGGATTGTACATCTCGTTCAAAACTTAACCATTGTAGAG AAACTAAGCTACATATAAGTTATATACCTGAATTTCAATAACCAAAGACTGAGTAGCTTTATAAACTAGAAAGATCTTGAATTTGGAAATGAAAAGTTAT TCCACCAATTGTTTAAAGAAAGGAAAAAAGTGAAGAAATGTTTTTTACATGACACCACTATGGAACCCAAAAGGATAAAAAGAGTAAAAAAGTGGAGTTT AACATAAGAATAAAATTACATGACTGTTACGTAACACAAATACATGTAAGTTAAGTGAAGAACAAGAAAAATAAATATTATTGTTGTTGTTGAA GCAACAATGGAGCAGCATGTGTCGACAAAGTTAACTACTATTCCTTTCTTCAAAATATTAGAATCTATCCCAAAGTTACTAGAAGTCGTCTTAAAGTTACTA CTCATATGCTCAAAGTTACAACAATGTGATTTTTATTCTTCCTCTCATAAAACTAATATGCAGTACTAATCGAAATCACTCATATTTTAGAGAGGTGGTC CTTCTCAAAATTTGTAGAAAGTCATATACAATACAAACTCGGGCTTGCAAAACATGGAACAATGCTAGAGTATAGTTTCTTAAACGAATTGAGTTCATGT TTAATTTCTATAGTGCATGAGGGATTCTACGATAGAGTTGGGGAAGGAAATATCAAGTTGCAGAAAGCAAAGAGCTTTGGATTCTCGAAAGAAGGTATCA TTCTCGAGGGTCAGACTGAACCAATAAAATCCGACTTAGTCAAACTCGCTACTGGATCATCATGCCCTCTTCTGCCAGCTTCAGAG
20	>c5f8af46-be33-4b4c-9de9-d8a9ee13ab59 (sequence:mRNA) 357 residues [Nitab4.5_0001961:372401-373786 + strand]	

	[cds]ATGGTGCATAAACCCGGCGAAGGCCTTCTTAAGTCTCCTCGCCACAACTCTTTCACCTTTGAGGTGGTCCTTCTCAAAATTTGTAGAAAGTCA TATACAATACAAACTCGGGCTTGCAAAAACATGGAACAATGCTAGAGTATAGTTCTTAAACGAATTGAGTTCATGTTTAATTTCTATAGTGCATGAGGGA TTCTACGATAGAGTTGGGGAAGGAAATATCAAGTTGCAGAAAGCAAAGAGCTTTGGATTCTCGAAAGAAGGTATCATTCTCGAGGGTCAGACTGAACCAA TAAAATCCGACTTAGTCAAACTCGCTACTGGATCATCATGCCCTCTTCTGCCAGCTTCAGAG
21	>c5f8af46-be33-4b4c-9de9-d8a9ee13ab59 (sequence:mRNA) 119 residues [Nitab4.5_0001961:372401-373786 + strand] [peptide]MVHKPGEGLLLSLLATTLSPLRWSFSKFVESHIQYKLGLAKHGTMLEYSFLNELSSCLISIVHEGFYDRVGEGNIKLQKAKSFGFSKEGII LEGQTEPIKSDLVKLATGSSCPLLPASE
22	**>Nitab4.5_0000783g0240.1 Nitab4.5_0000783:567132-574285**ATGGCAAATAATGATCGCAGAAAGAAGCAGGTGATAGCAATAATCGGAGCTGGTATAAGCGGCCTATTGTCCTGCAAATACAGCCTTTCAAAAG GTTTCGATCCCATAGTGTTTGAGTCCGAGGGTAGCATCGGAGGTGTGTGGACTAAGACAATTGAGAGTACCAAGCTGCAGACTCCAAGACCTGTTTACCA GTTCTCTGATTTTCCATGGCCTGATTCTGTAACCGACCAACAAACTGTGCTGAACTACATTGAATCCTACGCTCATCACTTTGATTTG CTCCGTCACATTCAGTTCAATAGCAAAGTGTTGAGCCTCAGCTATGAATCTGGTGACGACTCCGACGGAGAATGGAATTTGTGGGGCGGCACCGGCGAGC CTCTTAGCACTAAAGGGAAGTGGAATGTCACTGTACACGACACTCGAACCCTCTCCACACAAGTAATTAATTACTATGCCACTTTCTTTATCTCCT ATTTTAAAATGTTCTCACGCATTGGAGATTTGTGTCGGTGTCAATGTCCGATGATGACATTTGAGTAGTATCGTCTAAAGATCAAATAAATTATGCTGAA ATATACTTATTCTGTTTAGCTTGTTTGGCATCATATTTTTGGGTCAATTAATAGTATTTACAACGTAACTAGGCATTCTCTAGCATATACTAGCTAAATAC ATTAACGCATGACTCAGCTTCATGCCTTTGTAGGATTTGTATCTAACCGAACTTGTCCTATATATTGACCACCAGGCCAATACTTCATTAGTAAGAGTAG AAGAGCTGATAAGCTAGTATTACTAACACTGTTTTAATATAATTCCCCAATAATTACTAGAATAGCTAATCTAATGCAGTGTGTAATATGATTGTTTGTA ACCCAAAGAGAGCTGCATAAACGGTTCTCTGTAACTGTGACAGAGTGCTATTTCAAAAACTTCTACAAATTATTTTTACAAGTTCAATTCAGTATTTA TTGTTGGGTTAATATCTGAGTTGACTCCTACTTGATAGTTGATTTTTAGAATCAAATGGAGGATAAGCATATCTCTCATGGCTGCA GGTATACCAAGTTGATTTTGTAGTAGTTTGCCTGGGAAGATTCAGCCAAGTTCCAAATATCCCTCTAAACACAAAGGTCCACACAAGCTTTCGAC GGTGAAGTAATCCATTCAATGGACTATTCCAAAATGGATTCAACACAACTGCAACCAACTTGGTCAAAGGAAAACATGTAGTCGTCATTGGGTCCCAGAAAT CAGGAATGGACATTGCCATGGAGTGCTCCACAGTTAATGGTAAGTACTGCTACTGATCTTTCATATACCTTCATTCCCTGGACCTCTTTATTGTTTGTTT ATAAACATCAGCTTGATTCTGTTTGCACTAATTTTGCAGGGATTGAACGTCCATGTACAGTAGTAACCAGGACACCGCATTGGAACATTCCCGATTATTT TCCATGGGGATTCCCTTTGGCATATCTCTACCTAAATCGATTCTCCGAACTTATGGTGCATAAACCCGGTGAAGGCCTTCTTCTAAGTCTCCTCGCCACA ACTCTTTCACCTATGGTAATTATTTTCTTCCTGACTAATTTGACTCATAAAAAAACTGAACTATGTGCAGCAATGATTAGTCATACTTTCTCATTGAAAC TAACAAATGTATCTAAAGCAAAGCCAAATACTTTGAGGTTGACTAGAGGTTTCTTTATCTGAGATGAACATATTAGGAAACAAGTACTTTTCTATTTCTG GCCGTGGACTGTACATCTCGCTCAAAACTTCAGCATTGTAGAGAAACCGGGCTAAATATACTCTAAGTTATATACCTGAATTTCAATAACCAAAGACTGA GAATCTGAGTAGCTTTCTAAATTGGAAAGATCTTGAATTTGGAGATGAAAGTTATTCCCACCAACTGTTATGGAAAAAGGAAAAAAGCAAAGACTCTTTA CATGACACCACTAGGGAATCCAAAAGGATAAAACAAGTAAATAAGCGGAGTTTAACTGGATAGTTACAAATTTAAGGTTATATGTCAGCTATTTAGATCA TAAGAATCAAAATACATGACTGTTACGTAACACAGATACATGCAAGTGTGCAAGTTAAGTGAAGAAATAAGAAAAATAAATACAATTGCTGCTGAAGCAA CAACGGAGCAGTATGTGTCGACAAAGTTAACTACTATTCCTTCTCAAAATATCAGAATCCATCCCAAAGTTACTAGAAATCTGTCTTAAAGTTACTACTT ATATGCTCGAAGTTACACACAATGTGATTTCTATTCTTCACAAAACTAACATGCAGTACTAATCGAAATCTATCATATTTTACAGAGGTGGCCTT CTCAAAATTTGTAGAAAGTCATATACACAACACAACTCGGGATTGCCAAAACATGGAATAGTGCCAGAGCATAGTTTCTTAAATGAATTGAGTTCATGTTTA

ATTTCTTTAGTGCCTGAGGGCTTCTACGATAGAGTTGGGGAAGGAAGTATCAAGCTGAAGAAAGCAAAGAGCTTTGGATTCTCGAAAGAAGGTATCGTTC

TCGAGGGTCAGGCTGAACCAATAAAATCCGACTTAGTCATACTCGCTACTGGCTTCAAGGGAATTGATAAGCTCAAACACATTTTTGAATCACCAAGATT

TCAGGACTTCATTGCAGGCACAGATGATTCTGCAGTTCCTCTCTATAGGTGAGTATTAAAAGCCTTGTGCAGTGGTGGGAATTCTTTTGCATACATTCAT

CATATTCATCCAAATGTATATAACCTCATCTCATGTGAAAGTAATGTCATTCGGATGATTTCAGGGAATGCATTCATCCCCGAATTCCACAACTGGCAA

TAATTGGATTCTCAGAAAGCATAGCAAATCTATACACCTCGGAAATAAGATGCCGATGGCTGGCAGAACTTCTGGATGGAAAATTTAAGCTACCTAGCAT

TAAGGTCATGGAAGAAGACATAGCAGAATGGGATACATACAAGAAGAGATATTCATATTTGAAGAACTACAGGAGATCATGCATCGGTGCATTGCATATT

TGGTACAACGACCAACTGTGCAAGGACATGGGATGGAATCCTAAAAGGAAAAAAGGGTCTTTTGGCTGAGTGGTTCGATCCTTATGGACCAATGGATTACA

CCGGCTAACGAGTATGAGAGTATCATCAGGTCCTCTTCTGACAGCTTCAGAGAGGTCTCCTTTTCTTTGTTCTCGCTTGTACTCAAAATACTATTGTATG

TTTCGTATTTTCCTTCGTGTTGATGGCCGGGTGCCTTAGCTTGTCACTCTTTTATCATAGCATTCGCACAATAAGCTTGTGATCCACCTGTGGAAGCCCT

GATTTATACTAGCTTGACGTTCATGCTGGCTCAGCTTTATGAATCTCTACTGAATGACAAGCAGTTTGTGTAGATTCAATGGCACAAGAATATCAGATGG

ATCTACGAAACGGTGTGTTGTATTCTTGTTCATTAACCTTGAACAATAAGGCTAGCATTTTTAAATTTGCTATACTTTTTGTTCCGAAAGTGTCTTCCTT

CCATTTCTGGTCAATCTAAACAGTGTTCATTTGGTTGGAAAAGGAAGTGTTCATTTTCCTTGAGACATCCAGTACAATGAACCAAGGGAAAAAAGTATG

CCAGATCCTCTTTTCTCAGTACATCTATTCTGTTCTCTAGAAACTGTAGTCAATGTATCATTGCAAATGCTCCCAAGCACAATTACTGCTGAC

AATTTACATTTCTAGCTCGTCTAAGAGATTGGCCAAACTGCCAAAAAATGCTTCTTTTGCAATTCACTTTTACCAAAATGGATTTTCAAAATAGTAAATT

TTGGAGCGTAACAGTTTGTGTTTGGCCAATCCCTTTGAGAATTGTTTTTTACATTTATTTGAGAAGCAAAACCAATCTTTATAAAAAGTGCTTTTGAGTAT

TAAATTACTATCAAGGTTCAGTTTACTATTATTTAATTAAGTTTTTTATTTCTATTTAATTTAAAATATAAACTAATAATAAAAATTAAATATTT

TAATCAATATAAAATATGGTTATACCAAATTAATAATATTTCACATCTGGTATATTACAAAGAAACAAAAAGAAATCATACCTTATGAAAAAAGTGTTTA

AATGTGGTGATTAGCTTAGTTATTTTCAAAAAAGGAATGGGAGTAGAAATAGGAGATTTTTTTTTGGGCAAATTTCCTTCTCCAACCCATTGAACTT

TAGGCATCCCATTAAACTTGGCTTAATTTTTCATTTTGGCACTCTAATCCAATCTTGTTTCATTTTGGCCCTCTAACCTCATTTCTTATGTTTCACTTTA

ACACAAAAGCTGACTAGATTCAGAGTGTGAGTTGTCTCACCCTGTAGGCGCGTGAGGGAGATTTTTTTTGGGCAAATTTCCTTCTCCAACCCATTGAACTT

GGCTCCATTTTTTCATTTTGACACTCTATCTCAACCTTGTTTCATTTTGGCCCTCCAACTCTATTTCTTATGTTTCATTTTAACACGAAAGCTGACTAGT

CACATAGTGTGAATTGGCTCACATGATAAGTGCGTGAGACAGATTGTTTATTGCAAAATTCCTTCCCAACAGTAATAAAATAATTGTTGGCCCCATTTTT

AATCTCCTAACTTCCTTCCCCCATATTTCAACTTCATTTATTTCCTTCCTCATTTACTGTTTATATAAAACCAACCGATTTCAGTCGGTTTCTTAAAAAA

TAAATTTTGCGGGACACTAAAAGAGTTTCCTATATTTTTGCGCAAAAGAAAAAACGACCGAAGTTGGTCGATTTCGTAAAAAAAATTAAAAAATAAAATA

TTTTGAAATAAATTTTATATTTTGTGAAACATGTGTGCATGCATATCCTGAATTACCCACAAATTTCATAAACTTTGAATCGCTAAATTTTCGAACACCG

TAAACTTTCGAACCCGTAATGGAACAAGGTTGAGATAGAGTGCCAAAATGAAACATAAGAAATAGAGTTGGAGGGCTAAAATTGAACAAGATTGAGATAG

AGTGCCAAAATGAAACATAAGAAATAGAGTTGGAGGGCCAAAATGGAACAAGGTTGAGATAGAGTGCCAAAATAAAAAATGAAGCTAAGTTCAATGGGTT

GTCTATGTATTTGGCCAAGCAAGTATGTATACGCGCTGCATTTTTACGTCGGATAATAGTTTTTGTGTTAAAGTGAAACATAAGAAATGGGGTTGGAGGG

CCAAAATGGAACAAGATTGAGTTATAGTACCAAAATAAAAAATTGGGCCAAGTTTAATGGGTTGTCTATGTATTTGGCCAAAAATCTATTTATAGTACAA

AGACTATTACAAAGAAGAGTAGGAGAAGGACAATAATACGGTGAAACAAAAAGAGAAAGAGGATGGCATAAATAAAAAAAGGAAAATAAAAAGG

TTAAATTAATGAAAGATAATTTTGAAAATAGATATTTTGTCTTTAAAAAGTTTAGTTTTCTACTTCAACTTATTCTGTTTCTTCGAAAAAAACTAGAATTG

CTTTTGATTATGCTAAAAAGCACTTTTCCGTCTTGGCCGGACACCTCAATATCCAAAAAAATTATTTTTTCAGGAAAAATACTTTTGACGCCCAGAAGTT

TGACCAAACAAACTATAAAATGGGAGTCCATGTATAGCCGTGTTGTTGGCTGATCCTCCTGCCAAGTTTGTTGAAATACCACACTCTTGACCTATTGTGATA

TTATTCTACCGGATATATGAAGTTTAATTTAAAATGTTAGTTATTTGATTTTGATAGGTATATAAGAACAAAACTGTTGGTCAATTAGCCATCTGTAGGA

	TGGGTTACTGCTGTATACAATAGTTTTTTCTTCAGCTCTAGTCTAGACAGATTAGCGTTCTTTCTTATTGGACAGACTTGAAGAAAGAAAAAGGAAAGGG CAAAAAGAGTAGAAGATCCATTAGTAGAAACTTATCAGCATCCGAGTGTTTATATCAAATTTTATGTACTTATCACGATTAAGTGATTAATAATGTGAA AATTAAATAATCATAATTAAGTGGAAAAAAATGTAAGTGCAAACACATGTCATGCATGTATATAACATTTGACTCCATTTTAGTGGATGTAAAGAACAAT TTGATTCTGATTTTATTTGGAGTCACTTGATATTTTGTTTTTGAAAAAATTGAAAGAAAAAAATGTTAAAGATTAACTACAAGTCGTATAAAGATTACTA ATCAGATGTTAATTTACTTGTACGATTCCTGCTCTTCCTCTCACAAAGTGACACAAAACTAACTAATTAGTCGATTTCACTATCATATTTTACAGAGGTGGG CCTTCTCAAAATTTGTAGAAAGCCATATAAAACACAAACTGGGGCTTGCAAAACATGGAATGGTACCAGATCATAGTTTCTTAAGCGAATTGAGTTCATG TTCAGTTTCTAAAGTGCCAGAGGGCTTCTACGACAGAGTTGAGGAAGGAAGTATCAAACTCATCAAGAAGAGAATTTGGATTCTCTAAAGAAGGTATCG TTCTCGAGGTCGAGGGTCAGGCTGAACCAATAAAATCCGACTTAGTCATACTCGCTACTGGCTTCAAGGGAATTGATAAGTTCAAACACATTTTTGAATC ACCAAAATTTCAGGAATTCTTAACCGGCTCAGATGAATATTCTGCACTTCTCCTCTATAGGTGAGTAATAAATTAAATTAAAAAGGCTTGTGCAGCGGTA ATTAGAATTCTTATATATTTTCAGGGAATGCATTCATCCCCGAATACCACACAACTGGCAATAATTGGATTCTCAGAAAGCGTATCAAATCTATACACCTCG GAAATAAAGTGCCGTTGGCTGGGAGAGCTTCTTGACGAAACATTTAAGATAATGGAGAAAGACATAGCAGAATGGGATAAATACA AGAAGAGATATTCATATTTGAACAACTACAGGTACAGGAGATCATGACCAATTGTGCAGAGACATGGGATGGAATCCTAAAAGG AAAAAGGGTATTTTAGCTGAGTGGTTCCATCCTTATGGACCAATGGATTACGCCGGCTAA
23	>3e2f1a0b-637d-477c-ace4-d89ab73a949f (sequence:mRNA) 724 residues [Nitab4.5_0000783:567132-574285 + strand] [peptide]MANNDRRKKQVIAIIGAGISGLLSCKYSLSKGFDPIVFESEGSIGGVWTKTIESTKLQTPRPVYQFSDFPWPDSVTDVFPDQQTVLNYIES YAHHFDLLRHIQFNSKVLSLSYESGDDSDGEWNLWGGTGEPLSTKGKWNVTVHDTRTLSTQVYQVDFVVVCLGRFSQVPNIPQFPLNKGPQAFDGEVIHS MDYSKMDSTTATNLVKGKHVVVIGSQKSGMDIAMECSTVNGIERPCTVVTRTPHWNIPDYFPWGFPLAYLYLNNRFSELMVHKPGEGLLLSLLATTLSPMR WAFSKFVESHIQHKLGIAKHGIVPEHSFLNELSSCLISLVPEGFYDRVGEGSIKLKKAKSFGFSKEGIVLEGQAEPIKSDLVILLATGFKGIDKLKHIFES PRFQDFIAGTDDSAVPLYRECIHPRIPQLAIIGFSESIANLYTSEIRCRWLAELLDGKFKLPSIKVMEEDIAEWDTYKKRYSYLKNYRRSCIGALHIWYN DQLCKDMGWNPKRKKGLLAEWAFSKFVESHIKHKLGLAKHGMVPDHSFLSELSSCSVSKVPEGFYDRVEEGSIKLIKKRILDSLKKVEGQAEPIKSDLVI LATGFKGIDKFKHIFESPKFQEFLTGSDEYSALLLYRECIHPRIPQLAIIGFSESVSNLYTSEIKCRWLGELLDETFKVPNIKIMEKDIAEWDKYKKRYS YLNNYRDMGWNPKRKKGILAEWFHPYGPMDYAG
24	>3e2f1a0b-637d-477c-ace4-d89ab73a949f (sequence:mRNA) 724 residues [Nitab4.5_0000783:567132-574285 + strand] [peptide]MANNDRRKKQVIAIIGAGISGLLSCKYSLSKGFDPIVFESEGSIGGVWTKTIESTKLQTPRPVYQFSDFPWPDSVTDVFPDQQTVLNYIES YAHHFDLLRHIQFNSKVLSLSYESGDDSDGEWNLWGGTGEPLSTKGKWNVTVHDTRTLSTQVYQVDFVVVCLGRFSQVPNIPQFPLNKGPQAFDGEVIHS MDYSKMDSTTATNLVKGKHVVVIGSQKSGMDIAMECSTVNGIERPCTVVTRTPHWNIPDYFPWGFPLAYLYLNNRFSELMVHKPGEGLLLSLLATTLSPMR WAFSKFVESHIQHKLGIAKHGIVPEHSFLNELSSCLISLVPEGFYDRVGEGSIKLKKAKSFGFSKEGIVLEGQAEPIKSDLVILLATGFKGIDKLKHIFES PRFQDFIAGTDDSAVPLYRECIHPRIPQLAIIGFSESIANLYTSEIRCRWLAELLDGKFKLPSIKVMEEDIAEWDTYKKRYSYLKNYRRSCIGALHIWYN DQLCKDMGWNPKRKKGLLAEWAFSKFVESHIKHKLGLAKHGMVPDHSFLSELSSCSVSKVPEGFYDRVEEGSIKLIKKRILDSLKKVEGQAEPIKSDLVI LATGFKGIDKFKHIFESPKFQEFLTGSDEYSALLLYRECIHPRIPQLAIIGFSESVSNLYTSEIKCRWLGELLDETFKVPNIKIMEKDIAEWDKYKKRYS YLNNYRDMGWNPKRKKGILAEWFHPYGPMDYAG
25	**>Nitab4.5_0014046g0010.1** **Nitab4.5_0014046:3726-5355**

	TTCAAGCTCAAACAACTGTTTCTGTTTTTGAACCCAAAGAAAATACTAGGAAGAAGGAATTTTTCTCTTTTTTCCTTTCATTCTCTTAAAAACTGAAAA
	ATGGCCGGAATCGTTGTAGTTTTTGACTTCGACAAGACGATTATCGATTTGGACAGCGATAATTGGGTGGTGGATGAGTTAGGTGCTACTGATTTGTTCA
	ATCAACTTCTCCCTACCATGCCCTGGAATTCTGTCATGGTTGGTTTTTCTTTTCATATTTACACAGTCTATTATGTTCAATTTAATGACGTGAATTCGGG
	TGTATCGATCAAATCAAATAGTTTAACCGATGTTTACAATTTCAAATGACACAATATGCTCATGTAAGATAACGTGTCTTGCATTTATTAGTTTGATTTTT
	CCTAATTTTCCGGCTTTCACACAAATTACAGGACAAGATGATGAAGGAGCTTCATGCACAAGGCAAAACTATTGCGGACATTGAAGAAGTACTTAAACGG
	GCACCAATAATTCCCCGCGTAGTACCAGCCATTAAAGAGGCTCATGCTTTAGGGTATGTAGATCTCTAAACTTCAGTATTTTTATTGGCATATATTTCCT
	CTGTGTTTTTTTTTTTTTAAAATAGTGACTCAAGACATTAAATGCAATGCTTGTTATTTTCAGTTGTGATTTGAGGGTAGTAAGTGATGCAAATATTTTC
	TACATCGAGACAATATTGAAGCATCTGGGAATAATGGATTGCTTCTCAGAGATCAACTCAAACCCAGGATATATTGATGAGGAAGGAAAACTTAGAATCC
	AACCTTACCATGATTTTCACACTTCATCTCACGGTTGCAGTCTTTGTCCTCCCAACATGTGCAAGGTAAAACACAACAAAAAACTTAATCTTACTCGATG
	TCTACACAAATTCAACGATGTTAAAACACGTGTTTTTACATATACTGTTATTAATTAATTATGATCGTGGTAAATTATATGATTTTGGTGTAAACATCT
	GATTCTAATCTCATAACTTTCTTTTTATGGCCCAAGCTAGTTTCCGACTTGCATGTTAGATGGTACTAGAGTAGTTTTTCCTAATGTTTTTCGAATAT
	GCTGAAATACAGAATTTCTCTTTTTGATAATGTAGTATACTCTCATATCAAGCCTATGGCCAGCAAAATTTCTCTTTATATTTTTCTTTTAATCTTTT
	ATATGTATAATATTTGGATTTATAAACAATTAATATTCTAATGGATGGGTTTCTTTGTTAATAACAAAACAGAGGCCCTCATTATAGAAAGAATACAAG
	CTTCTCTGGCTACGGAAGGGAAGAAAGAATGATTTATCTTGGTGATGGAGCAGGTGATTTCTGCCCCAGCTTGAAACTCAAAAAGGAAGATTTTGTAAT
	GCCAAGGAAAGAATTTTCCGGTCTGGAAATTAATAAATGAAAATCGTGAACTGGTAAAAGCAGAGATTCATGGGTGGAAGGATGGAGAAGAGTATGAACAT
	ATTCTGCTTCAGATAATCAAAACACAATAACAATGGAGGAGAACCAATTGTTATCAGTTGATTGCAAGTTTCAGACACAATTCCCATTGAGGCCTTACCAAAAG
	CTGTCCCTGTGCCTTACTAATATTGAAGCT
26	>d382c2a7-4702-42b1-87da-23ec29e8ceba (sequence:mRNA) 807 residues [Nitab4.5_0014046:3726-5355 + strand]
	[cds]ATGGCCGGAATCGTTGTAGTTTTTGACTTCGACAAGACGATTATCGATTTGGACAGCGATAATTGGGTGGTGGATGAGTTAGGTGCTACTGATTT
	GTTCAATCAACTTCTCCCTACCATGCCCTGGAATTCTGTCATGGACAAGATGATGAAGGAGCTTCATGCACAAGGCAAAACTATTGCGGACATTGAAGAA
	GTACTTAAACGGGCACCAATAATTCCCCGCGTAGTACCAGCCATTAAAGAGGCTCATGCTTTAGGTTGTGATTTGAGGGTAGTAAGTGATGCAAATATTT
	TCTACATCGAGACAATATTGAAGCATCTGGGAATAATGGATTGCTTCTCAGAGATCAACTCAAACCCAGGATATATTGATGAGGAAGGAAAACTTAGAAT
	CCAACCTTACCATGATTTTCACACTTCATCTCACGGTTGCAGTCTTTGTCCTCCCAACATGTGCAAGGGCCTCATTATAGAAAGAATACAAGCTTCTCTG
	GCTACGGAAGGGAAGAAAGAATGATTTATCTTGGTGATGGAGCAGGTGATTTCTGCCCCAGCTTGAAACTCAAAAAGGAAGATTTTGTAATGCCAAGGA
	AAGATTTTCCGGTCTGGAAATTAATAAATGAAAATCGTGAACTGGTAAAAGCAGAGATTCATGGGTGGAAGGATGGAGAAGAGTATGAACATATTCTGCT
	TCAGATAATCAAAACACAATAACAATGGAGGAGAACCAATTGTTATCAGTTGATTGCAAGTTTCAGACACAATTCCCATTGAGGCCTTACCAAAAGCTGTCCCT
	GTGCCTTACTAA
27	>d382c2a7-4702-42b1-87da-23ec29e8ceba (sequence:mRNA) 268 residues [Nitab4.5_0014046:3726-5355 + strand]
	[peptide]MAGIVVVFDFDKTIIDLDSDNWVVDELGATDLFNQLLPTMPWNSVMDKMMKELHAQGKTIADIEEVLKRAPIIPRVVPAIKEAHALGCDLR
	VVSDANIFYIETILKHLGIMDCFSEINSNPGYIDEEGKLRIQPYHDFHTSSHGCSLCPPNMCKGLIIERIQASLATEGKKRMIYLGDGAGDFCPSLKLKK
	EDFVMPRKDFPVWKLINENRELVKAEIHGWKDGEEYEHILLQIIKTITMEENQLLSVDCKFQTIPIEALPKAVPVPY
28	**>Nitab4.5_0001023g0200.1 Nitab4.5_0001023:588126-589192**

	ATGATATTATCTAGTGGAAATGCCGTGGGTAATGTGAATGTTACAACAAAGAGCAATGGAAGGAAGTGAAGGACAACAGAGTTAGAAATGTAGTAAATC AAAATAGTCTCACTAATAATGAAATGGAGATGGCACCAGCTAAGCAGTTTGAAAATAAAGAACAAGGAGGGTACAAGCACTGTAGAGAGACAAGTAGAA GCACAAAGCAACAATGATAAGGATTTGGTGGCAGTAGACAATGAAGATAATGATCATGTTCAAGTAGCTAACAAGTTTGCAATATTACAAGGGATGGATG AAGAGGAAGAACCTGTTAATCAATTGGCAATGGTGGCAGCTAATGTAGCAACAAACATTTCAGCACTTCATAACCAAGCATCTACAAAGCAAAACAAAA AAATATGGTCAATAATTGGGGAAAGTTAATTCCAGCAGCACAAGCTTTTCATCTTAACTCATCGGGGATTATTTCGAATAATGGTCTAGTCAATGGAAAG GAGGCATCAAAATCAAAAAACGATACTGCACAGTGGGTAGAAAGAAACTTCAAGGATAAAACAGGAGAAGGAATTATGAAGATCAATAACCATGCGAGG AAATCCCATCACAAGATACATTAGTGAACAAGGTACTTTATAAAAATCCAAATTCTCAAGCAGAAGGGTCAAAATCGTCTACATTTAAAGAAAGAGTGCA AGAATGTGGAGGCAAGCTATGGGAGGCACAAAGGGAATACGATTCAGAGGAGAACGAAGTACCAATAGGAGCACAAGCTGATGATGAACCCAATGAGAAT GACAAAGAAGAAGATAAGCAAAGTGTAAAATGGAGAGATTCATGCCAATGACAACAATACACTGGTATTCATTTAATAAAGGAAGGATCAAAAAATGTTG AGCACATCAATAATTTTCAGACAGCAGAAGTCACAGAAATTAGCAACTATGAAGGAAGAGGCCCAGAGGTAAATGATCCTGGAGGAACAGAAGAAGATCA ACTTCAGAAAGTACAAGAAGATCCACAAGGACACATCACACAGAGAAGGAGAAACACCAAAATAA
29	>d133fb37-6339-4064-a95c-3b26afd231a (sequence:mRNA) 338 residues [Nitab4.5_0001023:588126-589192 + strand] [peptide]MILSSGNAVGNVNVTTKEQWKEVKDNRVRNVVNQNSLTNNEMEMAPAKQFENKRTRRDLVAVDNEDNDHVQVANKFAILQGMDEEEEPVNQ LAMVAANVATNISALHNQASTKQKQKNMVNNWGKLIPAAQAFHLNSSGIISNNGLVNGKEASKSKNDTAQWVERNFKDKTGEGIMKINKPCEEIPSQDTL VNKVLYKNPNSQAEGSKSSTFKERVQECGGKLWEAQREYDSEENEVPIGAQADDEPNENDKEEDKQSVNGEIHANDNNTTGIHLIKEGSKNVEHINNFQT AEVTEISNYEGRGPEVNDPGGTEEDQLQKVQEDPQGHITTEKEKQPK
30	>d133fb37-6339-4064-a95c-3b26afd231a (sequence:mRNA) 338 residues [Nitab4.5_0001023:588126-589192 + strand] [peptide]MILSSGNAVGNVNVTTKEQWKEVKDNRVRNVVNQNSLTNNEMEMAPAKQFENKRTRRDLVAVDNEDNDHVQVANKFAILQGMDEEEEPVNQ LAMVAANVATNISALHNQASTKQKQKNMVNNWGKLIPAAQAFHLNSSGIISNNGLVNGKEASKSKNDTAQWVERNFKDKTGEGIMKINKPCEEIPSQDTL VNKVLYKNPNSQAEGSKSSTFKERVQECGGKLWEAQREYDSEENEVPIGAQADDEPNENDKEEDKQSVNGEIHANDNNTTGIHLIKEGSKNVEHINNFQT AEVTEISNYEGRGPEVNDPGGTEEDQLQKVQEDPQGHITTEKEKQPK
31	**>Nitab4.5_0000021g0870.1** **Nitab4.5_0000021:2726211-2729260** GGAAAGAGAGTAGAGATGACGAAAGCGTTGTTGTAAGGCCAAACGTGGCGGCGCTGTGTATCCGGCGACTGAGCAGCTCCGTTCAGCCACGTGTACCTT CATTGCTCAAATGCGTCTCTATTTCTGCTCGTTCCTACAGGCCGATGAAGCTCGTTCCCCAAATAGGTTCAAAGATTGAGCTTTTCGCTGGTCGTCGGAC TTTCTCTACTCGTGCAACTAGCGATACTGGCTCGATTGATTCTCCTCTTATGGAGTCTATGCAGAAAAAGGTTAATCTTATCGACTCTCTTTTTTAGTATG CGTTTATTGCCCCTTTGTGACCATAAAACCTGTAATTTGTTTAATTGGGATTGTATTATCTGCTATGTGGTTTCTATATCTTGGGCATTGCTAGTATC TCGTTCTAAAGCTTAAAACTTTATGGATTACTGAGGCTTTGACCCCAAGAAATGTTCTTGGAAGCATATTTCTTTTTGCTACTGACGCTGTTACTTTTTT CTTTTTTAACTATAATGCATCTGAATACCGAGTCCAAATGAGCCTCACTACTAGATTAGAATAGTTATGAAAGATAAGTAGAGTGCACTGAATAGTGTTG ATCCAAGAAAAAAAATGTAGAGTGAATAGAATAGTACGTTAACTATCAAAATTCAAAGGAAGAAGTTGGAGAGGAGGGATTTAGAGACCAAAAAGAA GAACTTGAGGTCTAAATTTAAGGACGAAAAAAGAGGTGTAAATTCAAGGGGCTTTAGGTACGGGGTTCAATTCGTTGGTAGGGACGTAGAGGAGCGAGAA ATTCAGTTTGGGAAGAAACATCCTTGAGTACGTGCCATCTTACAATACAATTTATATAGAAGGAACCTGATAATGAGTATGTGCCATCTTGAAGCT ACAAAATATTCATGATATTTATTACAATGGTATATCATAGAGATGTTATGATTTGAAAAGAAAAAAAGGATCAAGAAAAACAGTGAATAAAGCA

	ACCTCTGGATGTGCAGTGAAGCATGCAATGAAACCCCTTGCTCTTAGGACTTTGGTCTAGTGATAAGAGCGCAGCTCGTGATGTGCGTCTAACCGTGCTG CAAACATAAGCTTGGTACTTAAGTGGAGAAGGGTAGAGGGGCAGGCCAATTATCCACCGAATTTCAAACCGTGCGCCACTAGCCTCTGGAATTTTATCGG TTAAGTTGAACTAACCTCCCTAGGATCACAGATGTTGCATATTATATCAAATAGCTACTGGCATGGTGCAGTGCAACACCTTGATCCATGTGCAGTAAAC TGTAAAGTTATGAGAAAAATTGAATTTATAGGCACCAGAGCTGTGATGAGTTTGAGTGGTTTCCTGGATGCGGGAAGTGGCTGAATGAGAAAGGAAGAC AGATATAGATGCTGAAAACTACTTGAACTATGTAAAGGGAAGAAAAATATAGAGGGAGTGTACATTTGTTGACGATAGTGTTAGTGGGTGGGAAAGGA TTATTTTGCAAGCAATTATCATATATTAGGAGAATCTCCCTTATCAAAAAATTAGGAGAATCTCACATTATTCGTCGCTGTGTCGTATTGATGTTAATAG AACTATGCTTATAAAAAGACTTCATCGCACCAAATAAGAAAGCACGGTCTGGATTTTGGATTTACGTGAATCCCTTCTAGAGGATCCTTACATGTTAGTT TACAATATTCTCAATTACTCAGATATAGTTTGCTTTTTTCTTCTTTTCCTCGTAGTAAGTATAAGTTTTTTGTTTGCAAACTGGTTCTGGTGCAACGCA GGTGGTTCATATTGATCAAATGCCATTAACGTCGGAGTCATATAAGAACGAACTGTAGAATGCTGCTTCTTTATGGCATATATTAGCACCTACTTTAACT AAAATATAGCAGCAACAGCTCTTAAGCGCAGCATGTCCACAATTTCAGTTCTGGTAAAGAACTACTCTTCCTTGTCCCATCTTTCGCAATCAGTTTTGCT GTAGCTCATTGTCCACCATTTCCAAGCGTGATCTCTATGTATGTGGCCGTGCCTAAGGGCATATTGGTTGAAGTAGATTTTTCTTCAAGAATTCCATT CAGATGTGACATTACAGACACTTTTGATGATTGACCTTTTTACATGTTTGTCACATTTTTGTCACAGTTACATAAACAAGGCTGTAGCCTACTATTACATTA ATAAAGATAGCAGAATTTGATTGGCTTATGCACATTTAGCTTTATAAATTCTATGCTCCTATATTTCCAATAGTTTAGGCTACATTTCTACTGGGGACAT CAATGCGCTTACCTTCATGTATTTTGTACTTCTGCCTAAAATGGTCCATTTGAAATTTAAATATATGTCAACAGATCAAGGAACAGTTGAATGCGGATTC AGTCATCGTTAAAGATGCTTATGGTGATGGTCAGGTGAGGTTTCTTCTCTATTAGATGGTGCCTAAATCAAGTGGTGAGTATCTGTTGCT AGTTCTGTAAAGTCATTGTTATTCTTCTCTTTGCAGCATTGATGTAGTCTCTTCAGCTTTTGAGGGACAATCTGCTGTGAATAGGCAGAGGATGGTCTAC AAAGCCATATGGGAAGAGCTTCAGAACACTGTGCATGCAGTTGACCAGATGACTACAAAAAACACCGACCGAAGCTGGGAAGTGATCAAGACACAACCGAG GAGTATCAAAGCACCACCTCCATCAACTACTCTATTTTCAAGTCACAGCTTTTCAATTGATATCAGATATGTTACAATGAACAAAGATAACTTTGAGTGA TGTCAACTCTAAAGTAGTGTTTGCTTTACATTGTTAGTATGAGACAGGGCATTGAATGCCTGCCATTTTGTAACAGCAGTTGGCTCATTTTCAAAATCTA GTTCTTCCGTACTGGTGGTTGAAGAACAATAGTCGAATATGCCCCATCCAAATGAAGTCAAGTTAAGGCCTTTGCGCAATGTTTCTGAAATTCTGATGAA AAATAAGAAGCTTGAGAGAAATTGGATTTCACTGCTGGCTGAATATATACAGAAATTTGGAGTTGGAAAATTATCAGCTCTCTGTCGTTGCCAATATATT AGCACGAAATATCTACTCTTGTTTGATTGTTTTTTAGATATCCATTGAATTCCAATGGAAATTGCATTTGAAA
32	>e47ab5b1-bcdf-4f17-92d0-702c741ab47d (sequence:mRNA) 486 residues [Nitab4.5_0000021:2726211-2729384 strand] + [cds]GGAAAGAGAGTAGAGATGACGAAAGCGTTGTTTGTAAGGCCAAACGTGGCGGCGCTGTGTATCCGGCGACTGAGCAGCTCCGTTCAGCCACGTGT ACCTTCATTGCTCAAATGCGGTCTCTATTTCTGCTCGTTCCTACAGGCCGATGAAGCTCGTTCCCCCAAATAGGTTCAAAGATTGAGCTTTTCGCTGGTCGT CGGACTTTCTCTACTCGTGCAACTAGCGATACTGGCTCGATTGATTCTCCTCTTATGGAGTCTATGCAGAAAAAGATCAAGGAACAGTTGAATGCGGATT CAGTCATCGTTAAAGATGCTTATGGTGATGGTCGGCATGTCAGCATTGATGTAGTCTCTTCAGCTTTTGAGGGACAATCTGCTGTGAATAGGCAGAGGAT GGTCTACAAAGCCATATGGGAAGAGCTTCAGAACACTGTGCATGCAGTTGACCAGATGACTACAAAAAACACCGACCGAAGCTGGGAAGTGA
33	>e47ab5b1-bcdf-4f17-92d0-702c741ab47d (sequence:mRNA) 161 residues [Nitab4.5_0000021:2726211-2729384 strand] + [peptide]GKRVEMTKALFVRPNVAALCIRRLSSSVQPRVPSLLKCVSISARSYRPMKLVPQIGSKIELFAGRRTFSTRATSDTGSIDSPLMESMQKKI KEQLNADSVIVKDAYGDGRHVSIDVVSSAFEGQSAVNRQRMVYKAIWEELQNTVHAVDQMTTKTPTEAGK	

Suitably, the protein for use according to the present invention may be encoded by a polynucleotide sequence from *Nicotiana tabacum*.

In one aspect the present invention provides a method of decreasing the alkaloid content of a plant or part thereof or cell (e.g. plant cell), the method comprising modifying said plant by decreasing or inhibiting the activity or expression of at least one FAD synthetase.

In one aspect the present invention provides a method of decreasing the alkaloid content of a plant or part thereof or plant cell, the method comprising modifying said plant by decreasing or inhibiting the activity or expression of at least one FAD synthetase comprising the amino acid sequence shown as SEQ ID No. 3, or a sequence which has at least 80% identity thereto or wherein the at least one gene encoding a FAD synthetase comprises a nucleotide sequence as set out in SEQ ID No. 1 or 2, or a functional variant or functional fragment or orthologue of SEQ ID No. 1 or 2, or a nucleic acid sequence which has at least 80% identity to SEQ ID No. 1 or 2. Suitably, the protein may comprise a FAD synthase domain.

In one aspect the present invention provides a method of decreasing the content of a TSNA precursor in a plant or part thereof (e.g. leaf), the method comprising modifying said plant by decreasing or inhibiting the activity or expression of at least one FAD synthetase. Suitably, the method comprising modifying said plant by decreasing or inhibiting the activity or expression of at least one FAD synthetase comprising the amino acid sequence shown as SEQ ID No. 3, or a sequence which has at least 80% identity thereto or wherein the at least one gene encoding a FAD synthetase comprises a nucleotide sequence as set out in SEQ ID No. 1 or 2, or a functional variant or functional fragment or orthologue of SEQ ID No. 1 or 2, or a nucleic acid sequence which has at least 80% identity to SEQ ID No. 1 or 2. Suitably, the protein may comprise a FAD synthase domain.

In one aspect the present invention provides a method of decreasing the content of a TSNA precursor in a plant or part thereof (e.g. leaf), the method comprising modifying said plant by decreasing or inhibiting the activity or expression of at least one FAD synthetase.

In one aspect the present invention provides a method of decreasing the content of a TSNA in a processed leaf, such as a cured leaf, the method comprising:

modifying a plant by decreasing or inhibiting the activity or expression of at least one FAD synthetase;

harvesting a leaf from said plant;

and processing e.g. curing said harvested leaf.

Suitably, the method of decreasing the content of a TSNA in a processed leaf may comprise: modifying said plant by decreasing or inhibiting the activity or expression of at least one FAD synthetase comprising the amino acid sequence shown as SEQ ID No. 3, or a sequence which

has at least 80% identity thereto or wherein the at least one gene encoding a FAD synthetase comprises a nucleotide sequence as set out in SEQ ID No. 1 or 2, or a functional variant or functional fragment or orthologue of SEQ ID No. 1 or 2, or a nucleic acid sequence which has at least 80% identity to SEQ ID No. 1 or 2. Suitably, the protein may comprise a FAD synthase domain. The term "decreasing" or "inhibiting" (e.g. inhibiting the activity or expression of a FAD synthetase) as used herein means that the activity or expression of the gene encoding the FAD synthetase protein is lower or decreased compared with the activity or expression of the gene in a comparable product.

In one aspect, the present invention provides a method of increasing the alkaloid content of a plant or part thereof or cell (e.g. plant cell), the method comprising modifying said plant by increasing or enhancing the activity or expression of at least one gene encoding a FAD synthetase protein.

In one aspect, the present invention provides a method of increasing the alkaloid content of a plant or part thereof or plant cell, the method comprising modifying said plant by increasing or enhancing the activity or expression of at least FAD synthetase comprising the amino acid sequence shown as SEQ ID No. 3, or a sequence which has at least 80% identity thereto or wherein the at least one gene encoding a FAD synthetase comprises a nucleotide sequence as set out in SEQ ID No. 1 or 2, or a functional variant or functional fragment or orthologue of SEQ ID No. 1 or 2, or a nucleic acid sequence which has at least 80% identity to SEQ ID No. 1 or 2. Suitably, the protein may comprise a FAD synthase domain. In one aspect the present invention provides a method of increasing the content of a TSNA precursor in a plant or part thereof (e.g. leaf), the method comprising modifying said plant by increasing or enhancing the activity or expression of at least one FAD synthetase.

In one aspect the present invention provides a method of increasing the content of TSNA precursor in a plant or part thereof (e.g. leaf), the method comprising modifying said plant by increasing or enhancing the activity or expression of at least one FAD synthetase comprising the amino acid sequence shown as SEQ ID No. 3, or a sequence which has at least 80% identity thereto, or wherein the at least one gene encoding an a FAD synthetase comprises a nucleotide sequence as set out in SEQ ID No. 1 or 2, or a functional variant or functional fragment or orthologue of SEQ ID No. 1 or 2, or a nucleic acid sequence which has at least 80% identity to SEQ ID No. 1 or 2.

The term "increasing" or "enhancing" (e.g. increasing the activity or expression of gene encoding a FAD synthetase protein) as used herein means that the activity or expression of the gene encoding the FAD synthetase protein is higher or increased compared with the activity or expression of the gene in a comparable product.

According to the present invention, the activity or expression of a FAD synthetase is modulated.

In one aspect the present invention provides a method of modulating (i.e. increasing or decreasing) the alkaloid content of a plant or part thereof or cell (e.g. plant cell), the method comprising modifying said plant by modulating (i.e. increasing or decreasing) the activity of at least one FAD synthetase.

The term “activity” refers to any functionality of the FAD synthetase protein. Examples of activity include the ability of the enzyme to catalyse the adenylation of flavin mononucleotide to form flavin adenine dinucleotide coenzyme. Other examples of activity may relate to the localization of the FAD synthetase protein.

Modulation of the activity of a FAD synthetase may entail increasing or decreasing the activity of the FAD synthetase.

Increasing the activity of a FAD synthetase refers to enhancing or improving the ability of the FAD synthetase to carry out a particular function in comparison to a FAD synthetase in a plant that has not been modified in accordance with the invention.

Decreasing the activity of a FAD synthetase refers to reducing, inhibiting or disrupting the ability of the FAD synthetase to carry out a particular function in comparison to a FAD synthetase in a plant that has not been modified in accordance with the invention. The activity of a FAD synthetase may be reduced to such an extent that the activity is prevented or eliminated.

In some embodiments the activity of a FAD synthetase be modulated (i.e. increased or decreased) by at least about 10% 20% 30%, or 40%, suitably at least about 50%, 60%, 70%, more suitably at least about 80%, 90%, 95% or 100% in comparison to the activity of a gene encoding a FAD synthetase in a plant (e.g. a tobacco plant) which has not been modified in accordance with the present invention.

In some embodiments the modulated FAD synthetase exhibits increased or decreased activity compared to an unmodified FAD synthetase. The modulated FAD synthetase may exhibit at least about 1%, at least about 3%, at least about 5%, at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80% or at least about 90% increased or decreased activity compared to an unmodified FAD synthetase.

Techniques are known in the art for measuring protein activities. For example, assays are known for measuring the enzymatic activity of a protein and the localization of a protein can be identified using microscopy techniques.

In particular, the ability of a FAD synthetase can be measured using techniques known in the art, such as colorimetric or fluorometric assays.

In one aspect the present invention provides a method of modulating (i.e. increasing or decreasing) the alkaloid content of a plant or part thereof or cell (e.g. plant cell), the method comprising modifying said plant by modulating (i.e. increasing or decreasing) the expression of at least one FAD synthetase.

The "expression" of a gene refers to the degree to which the information encoded in the gene is converted to a functionality. The level of expression of a gene may be equated with the amount of the product of that gene present in a cell or organism. A modification that modulates (i.e. increases or decreases) the expression of a gene is one that increases the amount of the product of that gene in a plant or cell in comparison to an unmodified plant or cell.

In some embodiments the expression of a FAD synthetase gene is modulated (i.e. increased or decreased) in comparison to the expression of a gene encoding a FAD synthetase in a plant (e.g. a tobacco plant) which has not been modified in accordance with the present invention.

In some embodiments the expression of a FAD synthetase gene may be modulated (i.e. increased or decreased) by at least about 10% 20% 30%, or 40%, suitably at least about 50%, 60%, 70%, more suitably at least about 80%, 90%, 95% or 100% in comparison to the expression of a gene encoding a FAD synthetase in a plant (e.g. a tobacco plant) which has not been modified in accordance with the present invention.

In some embodiments, the modulated FAD synthetase protein exhibits increased or decreased expression compared to an unmodified FAD synthetase. The modulated FAD synthetase protein may exhibit at least about 1%, at least about 3%, at least about 5%, at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80% or at least about 90% increased or decreased expression compared to an unmodified FAD synthetase.

Typically, genes are transcribed to mRNA, which is translated to protein, the final gene product. Proteins may be sequestered in cellular stores and/or degraded. The expression of a gene may be modulated by modulating any or all of these steps. Accordingly, in some embodiments the modification modulates expression of at least one gene encoding a FAD synthetase in one of the following ways:

modulating transcription from the at least one gene encoding a FAD synthetase;

modulating translation of the mRNA from the at least one gene encoding a FAD synthetase;

modulating release of the FAD synthetase from intracellular stores;

modulating the rate of degradation of the FAD synthetase and/or

introducing a mutation which modifies the amino acid sequence of the FAD synthetase to decrease or increase its activity e.g. to decrease FAD synthase activity.

The expression of specific genes encoding a FAD synthetase can be measured by measuring transcription and/or translation of the gene. Methods for measuring transcription are well known in the art and include, amongst others, northern blot, RNA-Seq, in situ hybridization, DNA microarrays and RT-PCR. Alternatively, the expression of a gene may be measured indirectly by measuring the level of the gene product for example the protein encoded by said gene. For example, the expression of a FAD synthetase may be determined by measuring the presence of the protein using an antibody specific for the FAD synthetase (for example antibodies specific for a FAD synthase domain) by western blot.

Modifying

The plant or cell may be modified in any way that modulates activity or expression of at least one FAD synthetase. Types of modifications to plants and cells that modulate activity or expression of genes, as well as techniques to achieve those modifications, are known in the art.

In some embodiments the present invention provides a method of decreasing the alkaloid content of a plant or part thereof or cell (e.g. plant cell), the method comprising modifying said plant by decreasing or inhibiting the activity or expression of at least one FAD synthetase as described herein.

In some embodiments the present invention provides a method of decreasing the content of a TSNA or a precursor of a TSNA in a tobacco plant or plant part thereof, the method comprising modifying said plant or a cell culture by decreasing the activity or expression of at least one FAD synthetase as described herein.

Any method known in the art for decreasing or inhibiting the activity or expression of a gene or ptoein may be used in the methods according to the present invention.

Suitably, the activity or expression of the gene encoding a FAD synthetase may be reduced, partly inactivated, inhibited, eliminated, knocked out or lost such that the protein activity, expression or function of the gene encoding a FAD synthetase may not be detectable.

In one aspect, the at least one gene encoding a FAD synthetase is knocked out. In other words, the gene encoding a FAD synthetase has been rendered completely inoperative.

By way of example, the present method may comprise:

- providing a mutation in a nucleic acid sequence which encodes a protein comprising the amino acid sequence shown as SEQ ID No. 3, or an amino acid sequence which has at least 80% sequence identity thereto;
- providing a mutation in a regulatory region (e.g. a promoter or an enhancer) which contributes to controlling the expression of a protein comprising the amino acid

sequence shown as SEQ ID No. 3, or an amino acid sequence which has at least 80% sequence identity thereto;

- providing an antisense RNA, siRNA or miRNA which reduces the level of nucleic acid sequence encoding a protein comprising the amino acid sequence shown as SEQ ID No. 3, or an amino acid sequence which has at least 80% sequence identity thereto.

Each of the above approaches results in the reduction or prevention of activity or expression of a protein comprising the amino acid sequence shown as SEQ ID No. 3, or an amino acid sequence which has at least 80% sequence identity thereto or wherein the at least one gene encoding a FAD synthetase comprises a nucleotide sequence as set out in SEQ ID No. 1 or 2, or a functional variant or functional fragment or orthologue of SEQ ID No. 1 or 2, or a nucleic acid sequence which has at least 80% identity to SEQ ID No. 1 or 2.

As used herein, the term "mutation" encompasses a natural genetic variant or an engineered variant. In particular, the term "mutation" refers to a variation in the nucleotide sequence encoding the amino acid sequence or in the amino acid sequence compared to the sequence shown as SEQ ID No. 3, or an amino acid sequence which has at least 80% (preferably at least 85%, preferably at least 90%, preferably at least 93%, preferably at least 95%, preferably at least 98%, preferably at least 99%) sequence identity thereto.

In one embodiment the mutation decreases the alkaloid content of a plant. In another embodiment, the mutation decreases the content of at least one TSNA precursor in a plant or part thereof, or leaf such as a harvested or processed leaf. In one embodiment the mutation decreases the content of one or more TSNAs selected from NNN, NNK, NAT, NAB, preferably NNN and/or NNK content is decreased in a processed leaf. Suitably, the TSNA content is reduced in relation to a comparable product.

In one embodiment, a method according to the present invention may comprise providing a nucleic acid sequence to a plant or part thereof or plant cell, wherein said nucleic acid results in the reduction or elimination of the activity or expression of at least one FAD synthetase.

In one embodiment, a method according to the present invention may comprise providing a nucleic acid sequence to a plant or part thereof or plant cell, wherein said nucleic acid results in the modification of the nucleic acid sequence of at least one FAD synthetase.

Suitably said nucleic acid sequence may be introduced to the plant or part thereof or cell. Suitably an endogenous nucleic acid sequence in the plant or part thereof or cell may be modified to encode the polypeptide according to the present invention (e.g. by gene editing). For example, an endogenous nucleotide sequence may be modified to decrease the activity or expression of at least one FAD synthetase.

In a preferred embodiment, each copy of a nucleic acid sequence encoding a protein comprising a sequence shown as SEQ ID No. 3, or a sequence which has at least 80% sequence identity thereto or wherein the at least one gene encoding a FAD synthetase comprises a nucleotide sequence as set out in SEQ ID No. 1 or 2, or a functional variant or functional fragment or orthologue of SEQ ID No. 1 or 2, or a nucleic acid sequence which has at least 80% identity to SEQ ID No. 1 or 2, which is present in the plant is modified e.g. mutated as defined herein (e.g. each genomic copy of a gene encoding said protein in a plant is mutated). For example, each copy of the gene in the allotetraploid genome of *Nicotiana tabacum* may be mutated.

In a preferred embodiment, some or all of the homologues of the FAD synthetase as described herein are modified e.g. inhibited or mutated. Suitably, some or all of the homologues listed in Table 1, or corresponding sequences which have at least 80% sequence identity thereto are modified e.g. inhibited or mutated.

In some embodiments the plant or plant cell according to the present invention is homozygous. Suitably, the plant or plant cell may be homozygous for the modification e.g. inhibition or mutation.

In some embodiments the plant or plant cell according to the present invention expresses only the modified e.g. mutated nucleic acid encoding at least one FAD synthetase. In other words, in some embodiments no endogenous (or endogenous and functional protein) is present in the plant according to the present invention. In other words, if any endogenous protein is present it is preferably in an inactive form.

In one embodiment the present method may comprise providing a mutation in the nucleic acid sequence shown as SEQ ID No. 1, or 2, or a nucleic acid sequence which has at least 80% identity thereto, or a homologue of SEQ ID No. 1 or SEQ ID No. 2.

The mutation may alter the plant genome such that a nucleic acid sequence encoding a protein comprising the amino acid sequence shown as SEQ ID No. 3, or an amino acid sequence which has at least 80% sequence identity thereto, or a homologue of SEQ ID No. 3, is completely or partially deleted or otherwise modified to inhibit or eliminate the activity of the FAD synthetase. In some embodiments the mutation does not alter the level or expression of the protein but reduces inhibits or eliminates the activity of the FAD synthetase.

Suitably, the mutation may be in a FAD synthase domain of the FAD synthetase. Suitably, the mutation may be in an active site. Exemplary active sites of SEQ ID No. 3 are amino acid residues 100-103, 106, 109, 200, 230, 252-254, Suitably, the FAD synthetase may comprise multiple mutations. Suitably, a mutation may be at an amino acid position which corresponds to at least one of amino acid residues 100-103, 106, 109, 200, 230, 252-254 of SEQ ID No. 3.

The mutation may interrupt the nucleic acid sequence which encodes a protein comprising the amino acid sequence shown as SEQ ID No. 3, or an amino acid sequence which has at least 80% sequence identity thereto, or a homologue of SEQ ID No. 3.

The interruption may cause the nucleic acid sequence to not be transcribed and/or translated.

The nucleic acid sequence may be interrupted, for example, by deleting or otherwise modifying the ATG start codon of the nucleic acid sequence such that translation of the protein is reduced or prevented.

The nucleic acid sequence may comprise one or more nucleotide change(s) that reduce or prevent expression of the protein or affect protein trafficking. For example, expression of the protein may be reduced or prevented by introduction of one or more pre-mature stop codons, a frame shift, a splice mutation or a non-tolerated amino acid substitution in the open reading frame.

A premature stop codon refers to a mutation which introduces a stop codon into the open reading frame and prevents translation of the entire amino acid sequence. The premature stop codon may be a TAG ("amber"), TAA ("ochre"), or TGA ("opal" or "umber") codon.

A frame-shift mutation (also called a framing error or a reading frame shift) is a mutation caused by indels (insertions or deletions) of a number of nucleotides in a nucleic acid sequence that is not divisible by three. Due to the triplet nature of gene expression by codons, the insertion or deletion can change the reading frame, resulting in a completely different translation from the original. A frameshift mutation will often cause the reading of the codons after the mutation to code for different amino acids. The frameshift mutation will commonly result in the introduction of a premature stop codon.

A splice mutation inserts, deletes or changes a number of nucleotides in the specific site at which splicing takes place during the processing of precursor messenger RNA into mature messenger RNA. The deletion of the splicing site results in one or more introns remaining in mature mRNA and may lead to the production of abnormal proteins.

A non-tolerated amino acid substitution refers to a mutation which causes a non-synonymous amino acid substitution in the protein which results in reduced or ablated function of the protein.

Any method known in the art for providing a mutation in a nucleic acid sequence may be used in the method according to the present invention. For example, homologous recombination may be used, in which a vector is created in which the relevant nucleic acid sequence(s) are mutated and used to transform plants or plant cells. Recombinant plants or plant cells expressing the mutated sequence may then be selected.

In one embodiment the mutation introduces a non-tolerated amino acid substitution in a protein comprising an amino acid sequence shown as SEQ ID No. 3, or a sequence which has at least 80% sequence identity thereto, or a homologue of SEQ ID No. 3.

In some embodiments, the FAD synthetase domain may contain a mutation which decreases the expression of the at least one gene encoding a FAD synthetase or decreases the activity of a FAD synthetase.

The mutation may be a deletion, a splice mutant or codon encoding a non-tolerated amino acid substitution.

In one embodiment, the nucleic acid sequence encoding the FAD synthetase may be wholly or partially deleted. The deletion may be continuous, or may comprise a plurality of sections of sequence. The deletion preferably removes a sufficient amount of nucleotide sequence such that the nucleic acid sequence no longer encodes a functional FAD synthetase. The deletion may be total, in which case 100% of the coding portion of the nucleic acid sequence is absent, when compared to the corresponding genome of a comparable unmodified plant. The deletion may, for example, remove at least 50, 60, 70, 80 or 90% of the coding portion of the nucleic acid sequence. Suitably, at least part of the protein may be deleted. The deletion may, for example, remove at least 10, 20, 30, 40, 50, 60, 70, 80 or 90% of the coding portion of the protein.

The deletion may remove at least 10 amino acids (such as at least 20, at least 30, at least 40, at least 50, at least 60, at least 70, at least 80, at least 90 amino acids) from the FAD synthetase. Suitably, the deletion may remove at least 10 amino acids (such as at least 20, at least 30, at least 40, at least 50, at least 60, at least 70, at least 80, at least 90 amino acids) of the FAD synthetase wherein the sequence of the FAD synthetase is aligned with SEQ ID No. 3. Suitably, the deletion may remove at least 10 amino acids (such as at least 20, at least 30, at least 40, at least 50, at least 60, at least 70, at least 80, at least 90 amino acids) from the FAD synthetase domain wherein the FAD synthetase prior to deletion comprises an amino acid sequence set forth in SEQ ID No. 3, or a sequence which has at least 80% sequence identity thereto, or a homologue of SEQ ID No. 3.

Suitably, the protein for use according to the present invention may comprise a truncated FAD synthetase. Suitably, the truncated protein may be a truncated version of an amino acid sequence set forth in SEQ ID No. 3, or a sequence which has at least 80% sequence identity thereto, or a homologue of SEQ ID No. 3. Suitably, the truncated protein lacks at least 10 amino acids (such as at least 20, at least 30, at least 40, at least 50, at least 60, at least 70, at least 80, at least 90 amino acids, at least 100 amino acids, at last 110 amino acids) from the FAD synthase domain of the FAD synthetase.

The deletion may remove at least part of a FAD synthase domain. The deletion may for example, remove at least 10, 20, 30, 40, 50, 60, 70, 80 or 90% of a FAD synthase domain. Suitably, the deletion may remove at least 5 amino acids, at least 10 amino acids, at least 15, at least 20, at least 25, at least 30 amino acids, at least 40 amino acids, at least 50 amino acids, at least 60 amino acids, at least 70 amino acids, at least 80 amino acids of a FAD synthase domain. Suitably, the deletion may remove 5 amino acids, 10 amino acids, 15, 20 amino acids, 25 amino acids, 30 amino acids, 40 amino acids, 50 amino acids, 60 amino acids, 70 amino acids, 80 amino acids of a FAD synthase domain.

The deletion may remove at least part of a FAD synthase domain. The deletion may, for example, remove at least one or at least two amino acids from a FAD synthase domain. Suitably, the FAD synthetase domain may be completely deleted.

Methods for deletion of nucleic acid sequences in plants are known in the art. For example, homologous recombination may be used, in which a vector is created in which the relevant nucleic acid sequence(s) are missing and used to transform plants or plant cells. Recombinant plants or plant cells expressing the new portion of sequence may then be selected.

Plant cells transformed with a vector as described herein may be grown and maintained in accordance with well-known tissue culturing methods such as by culturing the cells in a suitable culture medium supplied with the necessary growth factors such as amino acids, plant hormones, vitamins, etc.

Modification of the nucleic acid sequence may be performed using targeted mutagenesis methods (also referred to as targeted nucleotide exchange (TNE) or oligo-directed mutagenesis (ODM)). Targeted mutagenesis methods include, without limitation, those employing zinc finger nucleases, TALENs (see WO2011/072246 and WO2010/079430), Cas9-like, Cas9/crRNA/tracrRNA, Cas9/gRNA, or other CRISPR systems (see WO 2014/071006 and WO2014/093622), meganucleases (see WO2007/047859 and WO2009/059195), or targeted mutagenesis methods employing mutagenic oligonucleotides, possibly containing chemically modified nucleotides for enhancing mutagenesis with sequence complementarity to the gene, into plant protoplasts (e.g., KeyBase® or TALENs).

Alternatively, mutagenesis systems such as TILLING (Targeting Induced Local Lesions IN Genomics; McCallum et al. (2000) Nat. Biotech. 18:455, and McCallum et al. (2000) Plant Physiol. 123, 439-442, both incorporated herein by reference) may be used to generate plant lines which comprise a gene encoding a protein having a mutation. TILLING uses traditional chemical mutagenesis (e.g. ethyl methanesulfonate (EMS) mutagenesis, which produces random mutations) followed by high-throughput screening for mutations. Thus, plants, seeds, cells and tissues comprising a gene having the desired mutation may be obtained.

The method may comprise the steps of mutagenizing plant seeds (e.g. EMS mutagenesis), pooling of plant individuals or DNA, PCR amplification of a region of interest, heteroduplex formation and high-throughput detection, identification of the mutant plant, sequencing of the mutant PCR product. It is understood that other mutagenesis and selection methods may equally be used to generate such modified plants. Seeds may, for example, be radiated or chemically treated and the plants may be screened for a modified phenotype.

Fast neutron deletion mutagenesis may be used in a reverse genetics sense (i.e. with PCR) to identify plant lines carrying a deletion in the endogenous gene. See for example Ohshima et al. (1998) Virology 213:472-481; Okubara et al. (1994) Genetics 137:867-874; and Quesada et al. (2000) Genetics 154:421-4315 which are incorporated herein by reference.

In another approach, dominant mutants may be used to trigger RNA silencing due to gene inversion and recombination of a duplicated gene locus. See for example Kusaba et al. (2003) Plant Cell 15:1455-1467 (incorporated herein by reference).

Modified plants may be distinguished from non-modified plants, i.e., wild type plants, by molecular methods, such as the mutation(s) present in the DNA, and by the modified phenotypic characteristics. The modified plants may be homozygous or heterozygous for the modification. Preferably modified plants are homozygous for the modification.

In one embodiment the method of reducing or preventing the activity or expression of a protein comprising the amino acid sequence shown as SEQ ID No. 3, or an amino acid sequence which has at least 80% sequence identity thereto does not comprise treating the plant with a chemical (e.g. an agrochemical).

Other ways of reducing or preventing the expression will be apparent to one skilled in the art and include the use of virus-induced gene silencing (VIGs), micro RNA silencing, RNAi, antisense, tDNA insertions, or dominant negative constructs (or antimorphic mutations).

In one embodiment the expression of a gene encoding a protein comprising the amino acid sequence shown as SEQ ID No. 3, or an amino acid sequence which has at least 80% sequence identity thereto may be reduced or eliminated by virus-induced gene silencing.

In one embodiment the expression of a gene encoding a protein comprising the amino acid sequence shown as SEQ ID No. 3, or an amino acid sequence which has at least 80% sequence identity thereto may be reduced or eliminated by microRNAs.

In one embodiment the expression of a gene encoding a protein comprising the amino acid sequence shown as SEQ ID No. 3, or an amino acid sequence which has at least 80% sequence identity thereto may be reduced or eliminated by RNAi.

In one embodiment the expression of a gene encoding a protein comprising the amino acid sequence shown as SEQ ID No. 3, or an amino acid sequence which has at least 80% sequence identity thereto may be reduced or eliminated by antisense suppression.

In one embodiment the expression of a gene encoding a protein comprising the amino acid sequence shown as SEQ ID No. 3, or an amino acid sequence which has at least 80% sequence identity thereto may be reduced or eliminated by sense suppression.

In one embodiment the expression of a gene encoding a protein comprising the amino acid sequence shown as SEQ ID No. 3, or an amino acid sequence which has at least 80% sequence identity thereto may be reduced or eliminated by tDNA insertions.

In one embodiment the expression of a gene encoding a protein comprising the amino acid sequence shown as SEQ ID No. 3, or an amino acid sequence which has at least 80% sequence identity thereto may be reduced or eliminated by dominant negative constructs (or antimorphic mutations).

In one embodiment the expression of a gene encoding a protein comprising the amino acid sequence shown as SEQ ID No. 3, or an amino acid sequence which has at least 80% sequence identity thereto may be reduced or eliminated by a targeted mutagenesis based system.

In one embodiment the expression of a gene encoding a protein comprising the amino acid sequence shown as SEQ ID No. 3, or an amino acid sequence which has at least 80% sequence identity thereto may be reduced or eliminated by a gene editing e.g. CRISPR based system.

In one embodiment the expression of a gene encoding a protein comprising the amino acid sequence shown as SEQ ID No. 3, or an amino acid sequence which has at least 80% sequence identity thereto may be reduced or eliminated by zinc finger nuclease, TALENs, meganucleases, mutagenic oligonucleotides or TILLING.

In some embodiments the present invention provides a method of increasing the alkaloid content of a plant or part thereof or cell (e.g. plant cell), the method comprising modifying said plant by increasing or enhancing the activity or expression of at least one FAD synthetase.

Any method known in the art for increasing or enhancing the activity or expression of a gene may be used in the methods according to the present invention.

In some embodiments the method may comprise overexpressing at least one gene encoding a FAD synthetase. Suitably the method may comprise expressing one or more additional copies of the at least one gene encoding a FAD synthetase in the plant or cell. Suitably the method may comprise modifying the endogenous copy of the at least one gene encoding a FAD synthetase such that its expression is increased. The method may comprise mutating the

coding sequence of the at least one gene encoding a FAD synthetase. The method may comprise mutating a regulatory sequence that regulates expression of the at least one gene encoding a FAD synthetase.

Suitably the method may comprise transforming a cell of a plant (e.g. a tobacco plant) with a genetic construct which encodes at least one FAD synthetase comprising an amino acid sequence as set out in SEQ ID No. 3, or a functional variant or functional fragment or orthologue thereof, or a sequence which has at least 80% identity to SEQ ID No. 3; or wherein the at least one gene encoding a FAD synthetase protein comprises a nucleotide sequence as set out in SEQ ID No. 1 or 2, or a functional variant or functional fragment or orthologue of SEQ ID No. 1 or 2, or a nucleic acid sequence which has at least 80% identity to SEQ ID No. 1 or 2; or which comprises a nucleotide sequence which encodes a protein which is capable of promoting or augmenting at least one endogenous FAD synthetase. It will be appreciated that each of these options would result in an increased activity and expression of the polypeptide encoded by the at least one FAD synthetase. The method may comprise regenerating the plant from the transformed cell. There is provided use of genetic construct which is capable of increasing the activity and/or expression of a polypeptide encoded by at least one gene encoding a FAD synthetase for increasing the alkaloid content (e.g. nicotine content) in a plant or part there of or cell transformed with the construct.

The genetic construct may encode a polypeptide comprising the amino acid SEQ ID No. 3, or a functional variant or functional fragment or orthologue thereof, or a sequence which has at least 80% identity to SEQ ID No. 3; or wherein the at least one gene encoding a FAD synthetase comprises a nucleotide sequence as set out in SEQ ID No. 1 or 2, or a functional variant or functional fragment or orthologue of SEQ ID No. 1 or 2, or a nucleic acid sequence which has at least 80% identity to SEQ ID No. 1 or 2.

In another embodiment, the invention relates to a method of increasing the alkaloid content of a plant or part thereof or a cell, comprising modifying said plant or cell by increasing the activity of at least one FAD synthetase.

In one embodiment the activity of at least one gene encoding a FAD synthetase may be increased by introducing (or providing) a mutation to at least one gene encoding a FAD synthetase.

Suitably, the activity of at least one gene encoding a FAD synthetase may be increased by introducing a mutation to at least one gene encoding a FAD synthetase which comprises an amino acid sequence as set out in SEQ ID No. 3; or a functional variant or functional fragment or orthologue thereof, or a sequence which has at least 80% identity to SEQ ID No. 3; or wherein the at least one gene encoding a FAD synthetase comprises a nucleotide sequence as

set out in SEQ ID No. 1 or 2, or a functional variant or functional fragment or orthologue of SEQ ID No. 1 or 2, or a nucleic acid sequence which has at least 80% identity to SEQ ID No. 1 or 2.

In some embodiments a modification which increases the activity or expression of at least one FAD synthetase and thereby increases alkaloid content by one of the following:

modulating transcription from the at least one gene encoding a FAD synthetase;

modulating translation of the mRNA from the at least one gene encoding a FAD synthetase protein;

modulating release of the FAD synthetase protein from intracellular stores; and/or

modulating the rate of degradation of the FAD synthetase protein.

Alkaloid content

In one embodiment the present invention provides a method of modulating the alkaloid content of a plant (e.g. a tobacco plant) or a part thereof, the method comprising modifying said plant by modulating the activity or expression of at least one FAD synthetase.

The term "modulating" is used herein to mean either increasing or decreasing.

The term "increasing alkaloid content" is used herein to mean that the alkaloid content in the product of the present invention (e.g. plant, part thereof (e.g. leaf), processed leaf or a product made from the plant (e.g. a tobacco industry product)) is higher compared with a comparable product which has not been modified in accordance with the present invention.

The term "decreasing alkaloid content" is used herein to mean that alkaloid content in the product of the present invention (e.g. plant, part thereof (e.g. leaf), processed leaf or a product made from the plant (e.g. a tobacco industry product)) is lower compared with a comparable product which has not be modified in accordance with the present invention.

In some embodiments, the modulation of alkaloid content refers to an increase in alkaloid content wherein the activity or expression of at least one FAD synthetase is increased (or example the protein is overexpressed).

In some embodiments, the modulation of alkaloid content refers to a decrease in alkaloid content wherein the expression of at least one FAD synthetase is decreased or inhibited or eliminated.

In a further aspect, the alkaloid content is measured from leaves. In one aspect the alkaloid content is measured from green leaves. In a further aspect, the alkaloid content is measured from cured leaves, e.g. air-cured, flue-cured, fire-cured or sun-cured leaves. In a further aspect, the alkaloid content is measured from flue-cured leaves. In a further aspect, the alkaloid content is measured from air-cured leaves.

The term “alkaloid content” is used herein to mean the concentration and/or total amount of the entire group of compounds classified as alkaloids or the concentration and/or total amount of one or more compounds classified as alkaloids. Alkaloids typically present in tobacco include nornicotine, PON, anatabine, anabasine, nicotine, and myosmine. In some embodiments the content of one or more alkaloids, such as two or more alkaloids, such as three or more alkaloids, such as four or more alkaloids, such as five or more alkaloids, such as all six alkaloids, selected from nicotine, nornicotine, PON, anatabine, anabasine and myosmine is modulated. In some embodiments the content of one or more alkaloids, such as two or more alkaloids, such as three or more alkaloids, such as four or more alkaloids, such as five or more alkaloids, such as all six alkaloids, selected from nicotine, nornicotine, PON, anatabine, anabasine and myosmine is increased. In some embodiments the content of one or more alkaloids, such as two or more alkaloids, such as three or more alkaloids, such as four or more alkaloids, such as five or more alkaloids, such as all six alkaloids, selected from nicotine, nornicotine, PON, anatabine, anabasine and myosmine is decreased. In some embodiments the total alkaloid content of the plant or cell is modulated. In some embodiments the total alkaloid content is increased. In some embodiments the total alkaloid content is increased.

Any method known in the art for determining the concentration and/or total content of alkaloids may be used. One preferred method for analysing alkaloid content involves the analysis by gas chromatography-flame ionization detection method (GC-FID) or by reversed phase high performance liquid chromatography with tandem mass spectrometry (LC-MS/MS).

In one embodiment there is provided a method for producing a plant (e.g. a tobacco plant) or part thereof, a plant propagation material (e.g. a tobacco plant propagation material), a cell (e.g. a tobacco cell), a leaf (e.g. a tobacco leaf), a harvested leaf (e.g. a harvested tobacco leaf), a cut harvested leaf (e.g. a cut harvested tobacco leaf), a processed leaf (e.g. a processed tobacco leaf), a cut and processed leaf (e.g. a cut and processed tobacco leaf), a product comprising said plant or part thereof (e.g. a tobacco industry product) or combinations thereof obtainable or obtained by a plant of the invention which has modulated alkaloid content, the method comprising modifying said plant to modulate the activity or expression of a FAD synthtase. The modulated alkaloid content may be determined by comparing the alkaloid content in the plant (e.g. tobacco plant) or part thereof, plant propagation material (e.g. tobacco plant propagation material), a cell (e.g. a tobacco cell), leaf (e.g. tobacco leaf), harvested leaf (e.g. a harvested tobacco leaf), cut harvested leaf (e.g. a cut harvested tobacco leaf), processed leaf (e.g. processed tobacco leaf), cut and processed leaf (e.g. cut and processed tobacco leaf), a product comprising a plant or part thereof of the present invention, e.g. a tobacco industry product, or combinations thereof with a comparable product.

Suitably the alkaloid content may be modulated in a plant, e.g. a tobacco plant e.g. modified tobacco plant. Suitably the alkaloid content may be modulated in a leaf (e.g. a tobacco leaf e.g. a tobacco leaf from a modified tobacco plant). Suitably the alkaloid content may be modulated in a harvested leaf (e.g. a harvested tobacco leaf from a modified tobacco plant). Suitably the alkaloid content may be modulated in a cut harvested leaf (e.g. a cut harvested tobacco leaf from a modified tobacco plant). Suitably the alkaloid content may be modulated in a processed leaf (e.g. a processed tobacco leaf e.g. a processed tobacco leaf from a modified tobacco plant). Suitably the alkaloid content may be modulated in a cut and processed leaf (e.g. a cut and processed tobacco leaf e.g. a cut and processed tobacco leaf from a modified tobacco plant). Suitably the alkaloid content may be modulated in a cured leaf (e.g. cured a tobacco leaf from a modified tobacco plant). Suitably the alkaloid content may be modulated in an extract of a green leaf (e.g. a green tobacco leaf from a modified tobacco plant). Suitably the alkaloid content may be modulated in a product comprising the plant of the present invention or part thereof (e.g. a tobacco industry product, for example a tobacco industry product produced from a modified tobacco plant or part thereof). Suitably the alkaloid content may be modulated in any one of the above products or combinations thereof. Suitably the modulation of alkaloid content described above may be an increase in alkaloid content. Suitably the modulation of alkaloid content described above may be a decrease in alkaloid content (e.g. a decrease in nornicotine and/or PON content).

In one embodiment the content of one or more alkaloids selected from nornicotine, PON, anatabine and anabasine is decreased. In one embodiment the content of nornicotine is decreased. In one embodiment the content of PON is decreased. In one embodiment the content of anatabine is decreased. In one embodiment the content of anabasine is decreased.

In one embodiment the nicotine content of a modified plant (e.g. tobacco plant), plant propagation material (e.g. tobacco plant propagation material), leaf (e.g. tobacco leaf), harvested leaf (e.g. harvested tobacco leaf), cut harvested leaf (e.g. cut harvested tobacco leaf), processed leaf (e.g. processed tobacco leaf), cut and processed leaf (e.g. cut and processed tobacco leaf) or tobacco industry product from a modified tobacco plant is not substantially decreased. Suitably, the nicotine content is at least 85% (such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99%) of the nicotine content of a comparable product.

In one embodiment the alkaloid content of a plant (e.g. tobacco plant) or part thereof may be modulated by at least 0.5, 1.5, 2, 3 or 4 fold when compared to the alkaloid content of a plant (e.g. tobacco plant) or part thereof, respectively, which has not been modified to modulate the activity or expression of at least one gene encoding a FAD synthetase and which has been

grown under similar growth conditions. Suitably the alkaloid content may be modulated by about 0.5 fold to about 4 fold. Suitably the alkaloid content may be modulated by about 4 fold. Suitably the modification may be an increase or a decrease in alkaloid content. Suitably the modulation may be of one or more alkaloids selected from nicotine, nornicotine, PON, anatabine, anabasine and myosmine. Suitably the modulation may be of one or more alkaloids selected from nicotine, nornicotine, PON, anatabine and anabasine. Suitably, the nornicotine content may be reduced. Suitably, the PON content may be reduced. Suitably, the anatabine content may be reduced. Suitably, the anabasine content may be reduced.

In one embodiment of the invention the alkaloid content of a plant (e.g. a tobacco plant) or part thereof may be modulated by at least 1%, 2%, 5%, 8%, 10%, 12%, 15%, 20%, 25%, 30%, 40%, 50%, 60 %, 70%, 80%, 90% or 100% in comparison to a plant (e.g. a tobacco plant) or part thereof which has not been modified according to the present invention. In one embodiment the alkaloid content may be modulated by at least 30% in comparison to an unmodified plant or part thereof. In one embodiment the alkaloid content may be modulated by at least 40% in comparison to an unmodified plant or part thereof. In one embodiment the alkaloid content may be modulated by at least 50% in comparison to an unmodified plant or part thereof. In one embodiment the alkaloid content may be modulated by at least 60% in comparison to an unmodified plant or part thereof. The modulation may be an increase or a decrease in alkaloid content when compared to an unmodified plant (e.g. a tobacco plant) or part thereof.

Suitably the modulation may be of total alkaloid content. Suitably the modulation may be of one or more alkaloids selected from nicotine, nornicotine, PON, anatabine, anabasine and myosmine. Suitably the modulation may be of one or more alkaloids selected from nornicotine, nicotine, PON, anatabine and anabasine. Suitably the modulation may be of nornicotine content, such as decrease in nornicotine content. Suitably the modulation may be of anabasine content, such as decrease in anabasine content. Suitably the modulation may be of PON content, such as decrease in PON content. Suitably the modulation may be of anatabine content, such as decrease in anatabine content.

Suitably the modulation may be of more than one alkaloid, such as two or more alkaloids, such as three or more alkaloids, such as four or more alkaloids, such as five or more alkaloids, such as all six alkaloids, selected from nicotine, nornicotine, PON, anatabine, anabasine and myosmine.

In some embodiments the alkaloid content of the plant may be modulated by between about 5% and about 100%, by between about 10% and about 90%, by between about 20% and about 80%, by between about 30% and about 70%, by between about 40% and 60%, by between about 40% and 50%, or by between about 50% and 60%.

Tobacco-specific nitrosamine (TSNA) content

In one embodiment the present invention provides a method of decreasing the content of at least one TSNA precursor in a plant (e.g. a tobacco plant) or a part thereof or in a tobacco cell. Suitably, the method may comprise modifying said plant by modulating the activity or expression of at least one FAD synthetase. In one embodiment, the present invention provides a method of producing a processed leaf with decreased TSNA content (e.g. relative to a comparable product). The method of producing a processed leaf with decreased TSNA content may comprise:

modifying a plant by decreasing or inhibiting the activity or expression of at least one FAD synthetase;

harvesting a leaf from said plant;

and processing e.g. curing said harvested leaf.

The TSNA may be measured in a processed tobacco, e.g. cured tobacco or reconstituted tobacco. In one embodiment the TSNA content is measured and/or modified (e.g. reduced) in a cured tobacco plant or part thereof (e.g. in cured tobacco leaf).

The term “tobacco-specific nitrosamine” or “TSNA” as used herein has its usual meaning in the art, namely a nitrosamine which is found only in tobacco industry products or other nicotine-containing products. Suitably the at least one tobacco-specific nitrosamine may be N'-nitrosonornicotine (NNN), 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK), N'-nitrosoanatabine (NAT) or N-nitrosoanabasine (NAB).

The term “precursor thereto” when used in relation to at least one tobacco-specific nitrosamine refers to one or more chemicals or compounds of a tobacco plant that give rise to the formation of a tobacco-specific nitrosamine or are involved in the nitrosation reaction leading to tobacco-specific nitrosamine production.

In one embodiment the TSNA may be one or more of group selected from: N’-nitrosonornicotine (NNN), 4-(methyl nitrosamino)-1-(3-pyridyl)-1-butanone (NNK), N’nitrosoanatabine (NAT) and N’-nitrosoanabasine (NAB). Suitably the at least one tobacco-specific nitrosamine may be NNK or NNN. In one embodiment the tobacco-specific nitrosamine is NNN. In another embodiment the tobacco-specific nitrosamine is NNK.

In one embodiment the precursor of the TSNA is one or more of the group selected from nornicotine, anabasine, anatabine, and an oxidised derivative of nicotine such as pseudooxynicotine (PON).

In one embodiment the TSNA is N'nitrosonornicotine (NNN) and/or the precursor is nornicotine. In one embodiment the content of NNN is decreased. In one embodiment the content of nornicotine is decreased. In one embodiment the content of NNN and nornicotine is decreased.

In one embodiment the TSNA is 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) and/or the precursor is PON. In one embodiment the content of NNK is decreased. In one embodiment the content of PON is decreased. In one embodiment the content of NNK and PON is decreased.

In one embodiment the TSNA is N'nitrosoanatabine (NAT) and/or the precursor is anatabine. In one embodiment the content of NAT is decreased. In one embodiment the content of anatabine is decreased. In one embodiment the content of NAT and anatabine is decreased.

In one embodiment the TSNA is N'-nitrosoanabasine (NAB) and/or the precursor is anabasine. In one embodiment the content of NAB is decreased. In one embodiment the content of nornicotine is decreased. In one embodiment the content of NAB and anabasine is decreased.

The precursor of the TSNA (e.g. NNK, NNN, NAB and/or NAT) may be measured in green tobacco leaf, e.g. prior to processing, e.g. prior to curing. In one embodiment the precursor of the TSNA (e.g. NNK, NNN, NAB and/or NAT) is measured and/or modified (e.g. reduced) in a green tobacco leaf, e.g. prior to processing, e.g. prior to curing.

In one embodiment carrying out a method and or use of the invention results in a reduction of at least one TSNA or a precursor thereto in the modified tobacco plant (or part thereof) when compared to a tobacco plant (or part thereof) which has not been modified in accordance with the present invention.

The terms "reducing at least one TSNA or precursor thereto" or "reduction of at least one TSNA or precursor thereto" are used herein to mean that the concentration and/or total content of the at least one TSNA or precursor thereto in the product, method or use of the invention is lower in relation to a comparable product, method or use. For example, a comparable tobacco industry product would be derived from a tobacco plant which had not been modified according to the present invention, but in which all other relevant features were the same (e.g. plant species, growing conditions, method of processing tobacco, etc.).

Any method known in the art for determining the concentration and/or levels of at least one TSNA or precursor thereto may be used. In particular a method such may comprise the addition of deuterium labelled internal standard, an aqueous extraction and filtration, followed by analysis using reversed phase high performance liquid chromatography with tandem mass spectrometry (LC-MS/MS) may be used. Other examples for determining the concentration and/or level of a precursor to a tobacco-specific nitrosamine include a method such as the one detailed in CORESTA recommended method CRM-72: Determination of Tobacco Specific

Nitrosamines in Tobacco and Tobacco Products by LC-MS/MS; CRM being developed into ISO/DIS 21766 or Wagner et al. (2005) Analytical Chemistry 77(4), 1001-1006 all of which are incorporated herein by reference.

Suitably the concentration and/or total content of the at least one tobacco-specific nitrosamine or precursor thereto may be reduced by carrying out a method and/or use of the present invention. Suitably the concentration and/or level of the at least one tobacco-specific nitrosamine or precursor thereto may be reduced in a tobacco plant of the invention (e.g. obtainable or obtained by a method and/or use of the invention) when compared to the concentration and/or level of the at least one tobacco-specific nitrosamine(s) or precursor thereto in a tobacco plant which has not been modified in accordance with present invention.

The concentration and/or total content of the at least one tobacco-specific nitrosamine(s) or precursor thereto may be reduced in a tobacco leaf, harvested leaf, processed tobacco leaf, tobacco industry product or combinations thereof obtainable or obtained from a tobacco plant (or part of a tobacco plant or a tobacco cell culture) of the invention when compared with a tobacco leaf, harvested leaf, processed tobacco leaf, tobacco industry product or combinations thereof obtainable or obtained from a tobacco plant (or part of a tobacco plant or a tobacco cell culture) which has not been modified in accordance with the present invention.

Suitably the concentration and/or total content of the at least one tobacco-specific nitrosamine or precursor thereto may be reduced in a processed tobacco leaf.

Suitably the concentration and/or level of the at least one tobacco-specific nitrosamine or precursor thereto may be reduced in a tobacco industry product.

In one embodiment the at least one tobacco-specific nitrosamine or precursor thereto may be reduced by at least about 1%, at least about 3%, at least about 5%, at least about 10%, at least about 20%, at least about 30%, at least about 40% or at least about 50%. In some embodiments the at least one tobacco-specific nitrosamine or precursor thereto may be reduced by between about 5% and about 50%, by between about 10% and about 50%, by between about 20% and about 50%, by between about 30% and about 50%, or by between about 40% and 50%.

In relation to processed (e.g. cured) tobacco leaf (e.g. cured or reconstituted), the at least one tobacco-specific nitrosamine or precursor thereto may be reduced by between about 5000 ng/g and about 50 ng/g, by between about 4000 ng/g and about 100 ng/g, by between about 3000 ng/g and 500 ng/g or by between 2000 ng/g and 1000 ng/g. In some embodiments the at least one tobacco-specific nitrosamine or precursor thereto may be reduced by at least about 5000 ng/g, at least about 4000 ng/g, at least about 3000 ng/g, at least about 2000 ng/g, at least about 1000 ng/g, at least about 500 ng/g, at least about 100 ng/g or at least about 50 ng/g.

Biomass production

In some instances, it may be desirable to produce plants or biomass with high alkaloid levels e.g. high levels of nicotine content so that nicotine may be purified to produce a pure nicotine product for example for use in devices which utilize liquid containing nicotine (e.g. e-cigarettes) or within tobacco heating devices. For example, the production of nicotine in this way could reduce costs of nicotine extraction for the production of e-liquids for e-cigarettes.

In one aspect, the present invention provides a method of producing a biomass comprising: growing a cell which has been engineered to modulate (e.g. increase) the activity or expression of a gene encoding a FAD synthetase under conditions to produce a biomass. Suitably, the activity or expression of a FAD synthetase may be increased in order to increase the concentration and/or total nicotine content.

In one embodiment, the present invention provides a method of producing a biomass having modified (e.g. increased) concentration and/or total content of nicotine, comprising growing a cell which has been engineered to increase the activity or expression of at least one FAD synthetase comprising an amino acid sequence as set out in SEQ ID No. 3, or a functional variant or functional fragment or orthologue thereof, or a sequence which has at least 80% identity to SEQ ID No. 3; or wherein the at least one FAD synthetase comprises a nucleotide sequence as set out in SEQ ID No. 1 or 2, or a functional variant or functional fragment or orthologue of SEQ ID No. 1 or 2, or a nucleic acid sequence which has at least 80% identity to SEQ ID No. 1 or 2.

The cell may be engineered by any method known in the art to modify the activity or expression of at least one FAD synthetase. Suitably, the cell may be engineered to express an exogenous gene encoding a FAD synthetase. Suitably, the cell may be engineered to overexpress a gene encoding a FAD synthetase.

Suitably, the biomass may contain a higher concentration and/or total content of nicotine compared with the biomass produced by a comparable cell which has not been modified in accordance with the present invention.

Suitably the cell for use in biomass production may be a plant cell, such as a tobacco cell.

Suitably the cell for use in biomass production may be a yeast cell.

In one embodiment the cell (e.g. yeast cell) may be further modified to comprise one or more sequences that increases nicotinic alkaloid biosynthesis. Suitably these one or more sequences may be incorporated into a nucleic acid construct that is suitable for cell (e.g. yeast cell) transformation. The one or more sequences may be overexpressed in the cell (e.g. yeast cell). The sequences may be selected from one or more of the following genes: MPO (or Methylputrescine Oxidase or MPO1 or MPO2); A622 (or Isoflavone reductase-like protein or

Isoflavone reductase homolog or Isoflavone reductase-like protein); BBL (or Berberine bridge enzyme or Berberine bridge enzyme-like or BBE or NBB1); PMT (or Putrescine N-Methyltransferase or putrescine methyltransferase or S-adenosyl-L-methionine:putrescine N-methyltransferase or PMT or PMT1 or PMT2 or PMT3 or PMT4) and QPT (or quinolinate phosphoribosyltransferase). In one embodiment the sequences may be selected from one or more of the following genes: BBL, A622, PMT and MPO (MPO1 or MPO2). Genes suitable for modification in this way may be taught in US2016032299 for example, which is incorporated herein by reference.

Commercially desirable traits

In one embodiment the plants of the present invention have modified (i.e. increased or decreased) total alkaloid content and/or modified (i.e. increased or decreased) content of one or more alkaloids, whilst the flavour characteristics and/or other commercially desirable traits are at least maintained. Suitably, the plants of the present invention may have decreased total alkaloid content and/or decreased content of one or more alkaloids, whilst the flavour characteristics and/or other commercially desirable traits are at least maintained.

In one embodiment the plants of the present invention produce leaves of a similar grade and/or quality to plants which have not been modified according to the invention.

In one embodiment the plants of the present invention have reduced nornicotine and/or PON and/or anabasine and/or anatabine content without a significant change in the flavour characteristics of the plant (e.g. compared with the same plant which has not been modified in accordance with the present invention).

In one embodiment the plants of the present invention have decreased TSNA precursor content without a significant change (e.g. decrease) in other commercially desirable traits of the plant (e.g. compared with the same plant which has not been modified in accordance with the present invention). In particular the yield of the modified plant is preferably not reduced compared with the same plant which has not been modified in accordance with the present invention.

Therefore in one embodiment the methods and uses of the present invention relate to decreasing TSNA precursor content whilst maintaining the flavour characteristics and/or other commercially desirable traits (e.g. yield).

The term “commercially desirable traits” as used herein will include traits such as yield, mature plant height, harvestable leaf number, average node length, cutter leaf length, cutter leaf width, quality (e.g. leaf quality, suitably cured leaf quality), abiotic (for instance drought) stress tolerance, herbicide tolerance and/or biotic (for instance insect, bacteria or fungus) stress tolerance.

Leaf quality may be measured based on colour, texture and aroma of the cured leaf, for example according to United States Department of Agriculture (USDA) grades and standards.

Tobacco grades are evaluated based on factors including, but not limited to, the leaf stalk position, leaf size, leaf colour, leaf uniformity and integrity, ripeness, texture, elasticity, sheen (related with the intensity and the depth of coloration of the leaf as well as the shine), hygroscopicity (the faculty of the tobacco leaves to absorb and to retain the ambient moisture), and green nuance or cast.

Leaf grade can be determined using standard methods known in the art, for example, using an Official Standard Grade published by the Agricultural Marketing Service of the US Department of Agriculture (7 U.S.C. §511). See, e.g., Official Standard Grades for Burley Tobacco (U.S. Type 31 and Foreign Type 93), effective November 5, 1990 (55 F.R. 40645); Official Standard Grades for Flue-Cured Tobacco (U.S. Types 11, 12, 13, 14 and Foreign Type 92), effective March 27, 1989 (54 F.R. 7925); Official Standard Grades for Pennsylvania Seedleaf Tobacco (U.S. Type 41), effective January 8, 1965 (29 F.R. 16854); Official Standard Grades for Ohio Cigar-Leaf Tobacco (U.S. Types 42, 43, and 44), effective December 8, 1963 (28 F.R. 11719 and 28 F.R. 11926); Official Standard Grades for Wisconsin Cigar-Binder Tobacco (U.S. Types 54 and 55), effective November 20, 1969 (34 F.R. 17061); Official Standard Grades for Wisconsin Cigar-Binder Tobacco (U.S. Types 54 and 55), effective November 20, 1969 (34 F.R. 17061); Official Standard Grades for Georgia and Florida ShadeGrown Cigar-Wrapper Tobacco (U.S. Type 62), Effective April 1971. A USDA grade index value can be determined according to an industry accepted grade index. See e.g. Bowman et al. (1988) Tobacco Science, 32:39-40; Legacy Tobacco Document Library (Bates Document #523267826-523267833, July 1, 1988, Memorandum on the Proposed Burley Tobacco Grade Index); and Miller et al. (1990) Tobacco Intern., 192:55-57 (all foregoing references are incorporated herein in their entirety).

In one aspect, a USDA grade index is a 0-100 numerical representation of federal grade received and is a weighted average of all stalk positions. A higher grade index indicates higher quality. Alternatively, leaf grade may be determined via hyper-spectral imaging. See e.g. WO 2011/027315 (which is incorporated herein by reference).

In one embodiment, a tobacco plant of the present invention provides tobacco of commercially acceptable grade.

Suitably, the tobacco plant of the present invention provides cured tobacco of commercially acceptable grade.

In one embodiment, a tobacco plant of the present invention is capable of producing leaves having a USDA grade index value of at least about 70% of the USDA grade index value of leaves of a comparable plant when grown in similar growth conditions. Suitably, tobacco plants

disclosed herein may be capable of producing leaves having a USDA grade index value of at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, or at least about 98% of the USDA grade index value of a control plant when grown in similar growth conditions. Suitably, tobacco plants disclosed herein may be capable of producing leaves having a USDA grade index value of between 65% and 130%, between 70% and 130%, between 75% and 130%, between 80% and 130%, between 85% and 130%, between 90% and 130%, between 95% and 130%, between 100% and 130%, between 105% and 130%, between 110% and 130%, between 115% and 130%, or between 120% and 130% of the USDA grade index value of a comparable plant.

In one aspect, the tobacco plant of the present invention is capable of producing leaves having a USDA grade index value of at least 50. Suitably, tobacco plants disclosed herein may be capable of producing leaves having a USDA grade index value of 55 or more, 60 or more, 65 or more, 70 or more, 75 or more, 80 or more, 85 or more, 90 or more, and 95 or more.

Unless specified otherwise, used herein, tobacco yield refers to cured leaf yield which is calculated based on the weight of cured tobacco leaves per acre under standard field conditions following standard agronomic and curing practice.

In one aspect, a plant (e.g. a tobacco plant) of the present invention has a yield between 50% and 150%, between 55% and 145%, between 60% and 140%, between 65% and 135%, between 70% and 130%, between 75% and 125%, between 80% and 120%, between 85% and 115%, between 90% and 110%, between 95% and 105%, between 50% and 100%, between 55% and 100%, between 60% and 100%, between 65% and 100%, between 70% and 100%, between 75% and 100%, between 80% and 100%, between 85% and 100%, between 90% and 100%, between 95% and 100%, between 100% and 150%, between 105% and 150%, between 110% and 150%, between 115% and 150%, between 120% and 150%, between 125% and 150%, between 130% and 150%, between 135% and 150%, between 140% and 150%, or between 145% and 150% of the yield of a comparable plant when grown in similar field conditions.

In another aspect, the plant (e.g. a tobacco plant) yield of the present invention is approximately 1.5, 1.6, 1.7, 1.8, 1.9, 2.0, 2.1, 2.2, 2.3, 2.4, 2.5, 2.6, 2.7, 2.8, 2.9, or 3.0 times of the yield of a comparable plant when grown in similar field conditions.

In another aspect, the yield of a tobacco plant of the present invention is comparable to the yield of the flue cured comparable plant when grown in similar field conditions.

In one aspect, a tobacco plant of the present invention provides a yield selected from the group consisting of about between 1200 and 3500, between 1300 and 3400, between 1400 and 3300, between 1500 and 3200, between 1600 and 3100, between 1700 and 3000, between 1800 and

2900, between 1900 and 2800, between 2000 and 2700, between 2100 and 2600, between 2200 and 2500, and between 2300 and 2400 lbs/acre.

In another aspect, a tobacco plant of the present invention provides a yield selected from the group consisting of about between 1200 and 3500, between 1300 and 3500, between 1400 and 3500, between 1500 and 3500, between 1600 and 3500, between 1700 and 3500, between 1800 and 3500, between 1900 and 3500, between 2000 and 3500, between 2100 and 3500, between 2200 and 3500, between 2300 and 3500, between 2400 and 3500, between 2500 and 3500, between 2600 and 3500, between 2700 and 3500, between 2800 and 3500, between 2900 and 3500, between 3000 and 3500, and between 3100 and 3500 lbs/acre.

In a further aspect, a tobacco plant of the present invention provides a yield selected from the group consisting of about between 1200 and 3500, between 1200 and 3400, between 1200 and 3300, between 1200 and 3200, between 1200 and 3100, between 1200 and 3000, between 1200 and 2900, between 1200 and 2800, between 1200 and 2700, between 1200 and 2600, between 1200 and 2500, between 1200 and 2400, between 1200 and 2300, between 1200 and 2200, between 1200 and 2100, between 1200 and 2000, between 1200 and 1900, between 1200 and 1800, between 1200 and 1700, between 1200 and 1600, between 1200 and 1500, and between 1200 and 1400 lbs/acre.

Plant breeding

In one embodiment the present invention provides a method of producing a plant having a modified alkaloid content and/or modified content of a TSNA precursor comprising:

a. crossing a donor plant having modified nicotine content and/or modified content of a TSNA precursor and wherein the activity or expression of at least one FAD synthetase according to the present invention has been modulated in the donor plant in accordance with the present invention with a recipient tobacco plant that does not have modified nicotine content or modified content of a TSNA precursor and possesses commercially desirable traits;
b. isolating genetic material from a progeny of said donor plant crossed with said recipient plant; and
c. performing molecular marker-assisted selection with a molecular marker comprising:
    i. identifying an introgressed region comprising a mutation in a polynucleotide sequence encoding a protein defined in a.

Suitably, the activity or expression of a protein comprising an amino acid sequence as set out in SEQ ID No. 3, or a functional variant or functional fragment or orthologue thereof, or a

sequence which has at least 80% identity to SEQ ID No. 3; or a protein encoded by a nucleotide sequence as set out in SEQ ID No. 1 or 2 or a functional variant or functional fragment or orthologue of SEQ ID No. 1 or 2 or a nucleic acid sequence which has at least 80% identity to SEQ ID No. 1 or 2, is modulated in the donor plant when compared to a comparable plant. Suitably, the alkaloid content and/or TSNA precursor content is decreased by said method. Suitably, the alkaloid content and/or TSNA precursor content is decreased and the activity or expression of said FAD synthetase is decreased or inhibited.

The molecular marker assisted selection may comprise performing PCR to identify an introgressed nucleic acid sequence comprising a mutation which modulates the activity or expression of a protein comprising the amino acid sequence shown as SEQ ID No. 3, or an amino acid sequence which has at least 80% identity thereto.

Tobacco plants

The present invention provides methods, uses directed to plants (e.g. tobacco plants) as well as a cell (e.g. a tobacco cell), a cell culture, a plant (e.g. a tobacco plant) and a plant propagation material.

The term "tobacco plant" as used herein refers to a plant in the genus *Nicotiana* that is used in the production of tobacco industry products. Non-limiting examples of suitable "tobacco" plants include *N. tabacum* and *N. rustica* (for example, *N. tabacum* L., LA B21, LN KY171, TI 1406, Basma, Galpao, Perique, Beinhart 1000-1, and Petico).

The tobacco material can be derived or obtained from varieties of *Nicotiana tabacum* types, commonly known as Burley varieties, flue or bright varieties and dark varieties. In some embodiments, the tobacco material is derived from a Burley, Virginia or a dark tobacco plant. The tobacco plant may be selected from Burley tobacco, rare tobacco, speciality tobacco, expanded tobacco or the like.

The use of tobacco cultivars and elite tobacco cultivars is also contemplated herein. The tobacco plant for use herein may therefore be a tobacco variety or elite tobacco cultivar. Particularly useful *Nicotiana tabacum* varieties include Flue-cured Virginia type, Burley type, and Oriental type.

In some embodiments, the tobacco plant may be, for example, selected from one or more of the following varieties: L. cultivar T.I. 1068, AA 37-1, B 13P, Xanthi (Mitchell-Mor), KT D#3 Hybrid 107, Bel-W3, 79-615, Samsun Holmes NN, F4 from cross BU21 x Hoja Parado, line 97, KTRDC#2 Hybrid 49, KTRDC#4 Hybrid 1 10, Burley 21, PM016, KTRDC#5 KY 160 SI, KTRDC#7 FCA, KTRDC#6 TN 86 SI, PM021, K 149, K 326, K 346, K 358, K 394, K 399, K 730, KY 10, KY 14, KY 160, KY 17, KY 8959, KY 9, KY 907, MD 609, McNair 373, NC 2000, PG 01,

PG 04, P01, P02, P03, RG 11, RG 17, RG 8, Speight G-28, TN 86, TN 90, VA 509, AS44, Banket A1, Basma Drama B84/31, Basma I Zichna ZP4/B, Basma Xanthi BX 2A, Batek, Besuki Jember, C104, Coker 319, Coker 347, Criollo Misionero, PM092, Delcrest, Djebel 81, DVH 405, Galpao Comum, HB04P, Hicks Broadleaf, Kabakulak Elassona, PM102, Kutsage E1 , KY 14 x L8, KY 171, LA BU 21, McNair 944, NC 2326, NC 71, NC 297, NC 3, PVH 03, PVH 09, PVH 19, PVH 21 10, Red Russian, Samsun, Saplak, Simmaba, Talgar 28, PM132, Wislica, Yayaldag, NC 4, TR Madole, Prilep HC-72, Prilep P23, Prilep PB 156/1, Prilep P12-2/1, Yaka JK-48, Yaka JB 125/3, TI-1068, KDH-960, TI-1070, TW136, PM204, PM205, Basma, TKF 4028, L8, TKF 2002, TN 90, GR141 , Basma xanthi, GR149, GR153, and Petit Havana.

Non-limiting examples of varieties or cultivars are: BD 64, CC 101 , CC 200, CC 27, CC 301 , CC 400, CC 500, CC 600, CC 700, CC 800, CC 900, Coker 176, Coker 319, Coker 371 Gold, Coker 48, CD 263, DF91 1 , DT 538 LC, Galpao tobacco, GL 26H, GL 350, GL 600, GL 737, GL 939, GL 973, HB 04P, HB 04P LC, HB3307PLC, Hybrid 403LC, Hybrid 404LC, Hybrid 501 LC, K 149, K 326, K 346, K 358, K394, K 399, K 730, KDH 959, KT 200, KT204LC, KY10, KY14, KY 160, KY 17, KY 171 , KY 907, KY907LC, KTY14xL8 LC, Little Crittenden, McNair 373, McNair 944, msKY 14xL8, Narrow Leaf Madole, Narrow Leaf Madole LC, NBH 98, N-126, N-777LC, N-7371 LC, NC 100, NC 102, NC 2000, NC 291 , NC 297, NC 299, NC 3, NC 4, NC 5, NC 6, NC7, NC 606, NC 71 , NC 72, NC 810, NC BH 129, NC 2002, Neal Smith Madole, OXFORD 207, PD 7302 LC, PD 7309 LC, PD 7312 LC 'Periq'e' tobacco, PVH03, PVH09, PVH19, PVH50, PVH51 , R 610, R 630, R 7-1 1 , R 7-12, RG 17, RG 81 , RG H51 , RGH 4, RGH 51 , RS 1410, Speight 168, Speight 172, Speight 179, Speight 210, Speight 220, Speight 225, Speight 227, Speight 234, Speight G-28, Speight G-70, Speight H-6, Speight H20, Speight NF3, TI 1406, TI 1269, TN 86, TN86LC, TN 90, TN 97, TN97LC, TN D94, TN D950, TR (Tom Rosson) Madole, VA 309, VA359, AA 37-1 , B 13P, Xanthi (Mitchell-Mor), Bel-W3, 79-615, Samsun Holmes NN, KTRDC number 2 Hybrid 49, Burley 21 , KY 8959, KY 9, MD 609, PG 01 , PG 04, P01 , P02, P03, RG 1 1 , RG 8, VA 509, AS44, Banket A1 , Basma Drama B84/31 , Basma I Zichna ZP4/B, Basma Xanthi BX 2A, Batek, Besuki Jember, C104, Coker 347, Criollo Misionero, Delcrest, Djebel 81 , DVH 405, Galpao Comum, HB04P, Hicks Broadleaf, Kabakulak Elassona, Kutsage E1 , LA BU 21, NC 2326, NC 297, PVH 21 10, Red Russian, Samsun, Saplak, Simmaba, Talgar 28, Wislica, Yayaldag, Prilep HC-72, Prilep P23, Prilep PB 156/1 , Prilep P12-2/1 , Yaka JK-48, Yaka JB 125/3, TI-1068, KDH-960, TI-1070, TW136, Basma, TKF 4028, L8, TKF 2002, GR141 , Basma xanthi, GR149, GR153, Petit Havana. Low converter subvarieties of the above, even if not specifically identified herein, are also contemplated.

The tobacco plant may be a Burley, Flue-cured Virginia, or Oriental.

In one embodiment the plant propagation material may be obtainable from a plant (e.g. a tobacco plant) of the invention.

A "plant propagation material" as used herein refers to any plant matter taken from a plant from which further plants may be produced. Suitably, a plant propagation material may be selected from a seed, plant calli and plant clumps. Suitably the plant propagation material may be a seed. Suitably, the plant propagation material may be plant calli. Suitably the plant propagation material may be plant clumps.

In one embodiment the cell (e.g. tobacco cell), cell culture, tobacco plant and/or plant propagation material may be obtainable (e.g. obtained) by a method according to the invention.

Suitably a tobacco plant according to the present invention may have modulated (e.g. decreased) nicotine content when compared to an unmodified tobacco plant, wherein the tobacco plant has been modified to modulate (e.g. decrease) the activity or expression of at least one FAD synthetase. Suitably a tobacco plant according to the present invention may have decreased nicotine content when compared to an unmodified tobacco plant, wherein the tobacco plant has been modified to decrease or inhibit the activity or expression of at least one FAD synthetase.

Suitably a tobacco plant according to the present invention may have modulated (e.g. reduced) content of a TSNA precursor when compared to an unmodified tobacco plant, wherein the tobacco plant has been modified to modulate (e.g. decrease) the activity or expression of at least one FAD synthetase.

Suitably a tobacco plant according to the present invention may have decreased TSNA precursor content when compared to an unmodified tobacco plant, wherein the tobacco plant has been modified to decrease or inhibit the activity or expression of at least one FAD synthetase.

In one embodiment the tobacco plant in accordance with the present invention comprises a tobacco cell of the invention.

In another embodiment the plant propagation material may be obtainable (e.g. obtained) from a tobacco plant of the invention.

In one embodiment there is provided the use of a tobacco plant as described herein to breed a tobacco plant.

The present invention also provides in another embodiment the use of a tobacco plant of the foregoing embodiments for the production of a tobacco industry product.

In another embodiment there is provided the use of a tobacco plant of the invention to grow a crop.

In one embodiment there is provided the use of a cell as provided for in the foregoing embodiments for production of a tobacco industry product.

In one embodiment the present invention provides a cell culture (e.g. in *in vitro* culture).

The tobacco cell culture may be a cell suspension culture. These cells cultured in vitro may be incorporated into a tobacco industry product, e.g. as a substitute for conventional tobacco particles, shreds, fine cut or long cut tobacco lamina, as an additive ingredient or as both a substitute and an additive. Suitably, the cell culture may produce nicotine.

In one embodiment there is provided the use of a cell culture, e.g. a harvested and/or processed cell culture according to the present invention for the production of a tobacco industry product.

The tobacco cells harvested from an *in vitro* culture may be dried, e.g. freeze-dried, for example to produce a powder.

In one embodiment, the cell culture is a tobacco cell culture. The skilled person will be aware of known methods for establishing *in vitro* cultures of tobacco cells. By way of example only, the following method may be used: collecting seeds form a tobacco plant of interest and sterilising their exterior to eliminate unwanted organisms, planting said seeds to grown a tobacco plant of interest, removing tissue from the tobacco plant (for example, from the tobacco stem) for use as an explant, establishing a callus culture form the tobacco explant, establishing a cell suspension culture from the callus culture, and harvesting culture material (e.g. including tobacco cells) to produce a tobacco cell culture.

The tobacco cells can be harvested by various methods, including filtration, e.g. vacuum filtration. The sample may be washed in the filter by adding water and the remaining liquid removed with the filtration, e.g. vacuum filtration.

The harvested tobacco cell culture may be further processed, e.g. dried, such as air-dried and/or freeze-dried. The harvested tobacco cell culture or dried harvested tobacco cell culture or an extract therefrom may be incorporated into tobacco industry products according to the present invention.

In one embodiment, the present invention provides a plant (e.g. tobacco plant) or part thereof for use in molecular farming. Suitably, a plant or part thereof modified in accordance with the present invention may be used in the manufacture of proteins such as therapeutics e.g. antibiotics, virus like particles, neutraceuticals or small molecules.

In one embodiment, the present invention provides a method for the production of proteins (e.g. therapeutic proteins), the method comprising modifying a plant or part thereof capable of producing said protein (e.g. therapeutic protein) by modulating (e.g. decreasing) the activity or expression of at least one FAD synthetase protein having an amino acid sequence as set out in SEQ ID No. 3, or a functional variant or functional fragment or orthologue thereof, or a

sequence which has at least 80% identity to SEQ ID No. 3; or a homologue of SEQ ID No. 3; or wherein the at least one FAD synthetase comprises a nucleotide sequence as set out in SEQ ID No. 1 or 2, or a functional variant or functional fragment or orthologue of SEQ ID No. 1 or 2, or a nucleic acid sequence which has at least 80% identity to SEQ ID No. 1 or 2; or a homologue of SEQ ID No. 1 or 2; and culturing the plant under conditions sufficient to allow the production of said protein (e.g. therapeutic protein).

Products

The present invention also provides for products obtainable or obtained from plants according to the present invention. Products are provided which are obtainable or obtained from a plant in which the activity or expression of a FAD synthetase has been modulated.

In one embodiment, the product may comprise a construct of the invention which modulates the activity or expression of at least one FAD synthetase as defined herein. In one embodiment, the product may comprise a construct of the invention which modifies the nucleic acid sequence of at least one FAD synthetase as defined herein.

The present invention also provides for products obtainable or obtained from tobacco according to the present invention.

In one embodiment there is provided the use of a tobacco plant of the invention to produce a tobacco leaf.

Suitably the tobacco leaf may be subjected to downstream applications such as processing.

Thus in one embodiment the use of the foregoing embodiment may provide a processed tobacco leaf. Suitably the tobacco leaf may be subjected to curing, fermenting, pasteurising or combinations thereof. In another embodiment the tobacco leaf may be cut. In some embodiments the tobacco leaf may be cut before or after being subjected to curing, fermenting, pasteurising or combinations thereof.

In one embodiment the present invention provides a harvested leaf of a tobacco plant of the invention.

In a further embodiment the harvested leaf may be obtainable (e.g. obtained) from a tobacco plant propagated from a propagation material of the present invention.

In another embodiment there is provided a harvest leaf obtainable from a method or use of the present invention.

Suitably the harvested leaf may be a cut harvested leaf.

In some embodiments the harvested leaf may comprise viable tobacco cells. In other embodiments the harvested leaf may be subjected to further processing.

There is also provided a processed tobacco leaf.

The processed tobacco leaf may be obtainable from a tobacco plant of the invention. Suitably the processed tobacco leaf may be obtainable from a tobacco plant obtained in accordance with any of the methods and/or uses of the present invention.

Suitably, the processed leaf may comprise reduced content of one or more TSNAs selected from NNN, NNK, NAT and NAB. Suitably, the content of NNN may be reduced. Suitably, the content of NNK may be reduced. Suitably, the content of NAT may be reduced. Suitably, the content of NAB may be reduced. Suitably, the reduction in TSNA content is in relation to a comparable product which has not been modified according to the present invention.

In another embodiment the processed tobacco leaf may be obtainable (e.g. obtained) from a tobacco plant propagated form a tobacco plant propagation material according to the present invention.

The processed tobacco leaf of the present invention may be obtainable (e.g. obtained) by processing a harvested leaf of the invention.

The term "processed tobacco leaf" as used herein refers to a tobacco leaf that has undergone one or more processing steps to which tobacco is subjected to in the art. A "processed tobacco leaf" comprises no or substantially no viable cells.

The term "viable cells" refers to cells which are able to grow and/or are metabolically active. Thus, if a cell is said to not be viable, also referred to as "non-viable" then a cell does not display the characteristics of a viable cell.

The term "substantially no viable cells" means that less than about 5% of the total cells are viable. Preferably, less than about 3%, more preferably less than about 1%, even more preferably less than about 0.1% of the total cells are viable.

In one embodiment the processed tobacco leaf may be processed by one or more of: curing, fermenting and/or pasteurising.

Suitably the processed tobacco leaf may be processed by curing.

Tobacco leaf may be cured by any method known in the art. In one embodiment tobacco leaf may be cured by one or more of the curing methods selected from the group consisting of: air curing, fire curing, flue curing and sun curing.

Suitably the tobacco leaf may be air cured.

Typically air curing is achieved by hanging tobacco leaf in well-ventilated barns and allowing to dry. This is usually carried out over a period of four to eight weeks. Air curing is especially suitable for burley tobacco.

Suitably the tobacco leaf may be fire cured. Fire curing is typically achieved by hanging tobacco leaf in large barns where fires of hardwoods are kept on continuous or intermittent low smoulder

and usually takes between three days and ten weeks, depending on the process and the tobacco.

In another embodiment the tobacco leaf may be flue cured. Flue curing may comprise stringing tobacco leaves onto tobacco sticks and hanging them from tier-poles in curing barns. The barns usually have a flue which runs from externally fed fire boxes. Typically this results in tobacco that has been heat-cured without being exposed to smoke. Usually the temperature will be raised slowly over the course of the curing with the whole process taking approximately 1 week.

Suitably the tobacco leaf may be sun cured. This method typically involves exposure of uncovered tobacco to the sun.

Suitably the processed tobacco leaf may be processed by fermenting.

Fermentation can be carried out in any manner known in the art. Typically during fermentation, the tobacco leaves are piled into stacks (a bulk) of cured tobacco covered in e.g. burlap to retain moisture. The combination of the remaining water inside the leaf and the weight of the tobacco generates a natural heat which ripens the tobacco. The temperature in the centre of the bulk is monitored daily. In some methods every week, the entire bulk is opened. The leaves are then removed to be shaken and moistened and the bulk is rotated so that the inside leaves go outside and the bottom leaves are placed on the top of the bulk. This ensures even fermentation throughout the bulk. The additional moisture on the leaves, plus the actual rotation of the leaves themselves, generates heat, releasing the tobacco's natural ammonia and reducing nicotine, while also deepening the colour and improving the tobacco's aroma. Typically the fermentation process continues for up to 6 months, depending on the variety of tobacco, stalk position on the leaf, thickness and intended use of leaf.

Suitably the processed tobacco leaf may be processed by pasteurising. Pasteurising may be particularly preferred when the tobacco leaf will be used to make a smokeless tobacco industry product, most preferably snus.

Tobacco leaf pasteurisation may be carried out by any method known in the art. For example pasteurisation may be carried out as detailed in J Foulds, L Ramstrom, M Burke, K Fagerstrom. Effect of smokeless tobacco (snus) on smoking and public health in Sweden Tobacco Control (2003) 12: 349–359, the teaching of which is incorporated herein by reference.

During the production of snus, pasteurisation is typically carried out by a process in which the tobacco is heat treated with steam for 24–36 hours (reaching temperatures of approximately 100°C). This results in an almost sterile product and without wishing to be bound by theory one of the consequences of this is believed to be a limitation of further TSNA formation.

In one embodiment the pasteurisation may be steam pasteurisation.

In some embodiments the processed tobacco leaf may be cut. The processed tobacco leaf may be cut before or after processing. Suitably, the processed tobacco leaf may be cut after processing.

In one embodiment, the use of the foregoing embodiment may provide reconstituted tobacco.

In one embodiment, there is provided reconstituted tobacco.

"Reconstituted" as used herein may also be referred to as recon, recycled or homogenized sheet tobacco and refers to tobacco material generated from remnants of tobacco leaf after processing. Reconstituted tobacco allows the production of a consistent, high quality blend and allows the adjustment of the ratio of individual components.

Reconstituted tobacco may be nano fibre recon (nanofibers can be extracted in solid or liquid form), paper making recon (which uses stems, scraps, and midribs, etc. as the raw material) or slurry type recon (which uses a mixture of fines and tobacco stems, ground to power, mixed with water and vegetable binding agent; the soluble residue is formed to sheets by extracting the water).

Any method known in the art may be used for making reconstituted tobacco, for example see CORESTA Congress, Sapporo, 2012, Smoke Science/Product Technology Groups, SSPT 12 (incorporated herein by reference).

In some embodiments the tobacco plant, harvested leaf of a tobacco plant and/or processed tobacco leaf may be used to extract nicotine. The extraction of nicotine can be achieved using any method known in the art. For example a method for extracting nicotine from tobacco is taught in US 2,162,738 which is incorporated herein by reference.

In one aspect, the present invention provides cured tobacco material made from a tobacco plant or part thereof according to the invention.

Suitably, the cured tobacco may comprise a reduced content of one or more TSNAs selected from NNK, NNN, NAT and NAB. Suitably, the content of NNN may be reduced. Suitably, the content of NNK may be reduced. Suitably, the content of NAT may be reduced. Suitably, the content of NAB may be reduced. Suitably, the reduction in TSNA content is in relation to a comparable product which has not been modified according to the present invention.

In another aspect, the present invention provides a tobacco blend comprising tobacco material made from a tobacco plant or part thereof according to the present invention, or from a tobacco cell culture according to the present invention. In one aspect, the present invention provides a tobacco blend comprising cured tobacco material according to the present invention.

Suitably, the tobacco blend according to the present invention may comprise approximately 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% or 90% tobacco from a tobacco plant or part thereof according to the present invention, or from a tobacco cell culture according to the

present invention. Suitably, the tobacco blend may comprise approximately 10% tobacco from a tobacco plant or part thereof according to the present invention, or from a tobacco cell culture according to the present invention. Suitably, the tobacco blend may comprise approximately 20% tobacco from a tobacco plant or part thereof according to the present invention, or from a tobacco cell culture according to the present invention. Suitably, the tobacco blend may comprise approximately 30% tobacco from a tobacco plant or part thereof according to the present invention, or from a tobacco cell culture according to the present invention. Suitably, the tobacco blend may comprise approximately 40% tobacco from a tobacco plant or part thereof according to the present invention, or from a tobacco cell culture according to the present invention. Suitably, the tobacco blend may comprise approximately 50% tobacco from a tobacco plant or part thereof according to the present invention, or from a tobacco cell culture according to the present invention. Suitably, the tobacco blend may comprise approximately 60% tobacco from a tobacco plant or part thereof according to the present invention, or from a tobacco cell culture according to the present invention. Suitably, the tobacco blend may comprise approximately 70% tobacco from a tobacco plant or part thereof according to the present invention, or from a tobacco cell culture according to the present invention. Suitably, the tobacco blend may comprise approximately 80% tobacco from a tobacco plant or part thereof according to the present invention, or from a tobacco cell culture according to the present invention. Suitably, the tobacco blend may comprise approximately 90% tobacco from a tobacco plant or part thereof according to the present invention, or from a tobacco cell culture according to the present invention.

In one aspect, a tobacco blend product of the present invention comprises at least about 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, or 95 percent by dry weight of tobacco cured from a tobacco plant or part thereof according to the present invention, or a tobacco cell culture according to the present invention.

Suitably, the cured tobacco material may be air cured. Suitably, the cured tobacco material may be flue cured. Suitably, the cured tobacco material may be sun cured. Suitably, the cured tobacco material may be fire cured.

A tobacco industry product or smoking article according to the present invention may comprise the tobacco material (e.g. cured tobacco material or reconstituted tobacco material) according to the present invention.

In another aspect the present invention provides a tobacco industry product.

In one embodiment the tobacco industry product according to the present invention may be a blended tobacco industry product. Suitably, the tobacco blend may comprise cured tobacco material according to the present invention.

In one embodiment the tobacco industry product may be prepared from a tobacco plant of the invention or a part thereof.

Suitably the tobacco plant or part thereof may be propagated from a tobacco plant propagation material according to the present invention.

The term "part thereof" as used herein in the context of a tobacco plant refers to a portion of the tobacco plant. Suitably, the "part thereof" may be a leaf, root or stem of a tobacco plant or the flowers. Suitably, the "part thereof" may be a leaf, root or stem of a tobacco plant.

Tobacco industry product

As used herein, the term "tobacco industry product" is intended to include combustible smoking articles such as cigarettes, cigarillos, cigars, tobacco for pipes or for roll-your-own cigarettes, (whether based on tobacco, tobacco derivatives, expanded tobacco, reconstituted tobacco, tobacco substitutes or other smokable material), non-combustible aerosol provision systems such as heating products that release compounds from substrate materials without burning such as electronic cigarettes, tobacco heating products, and hybrid systems to generate aerosol from a combination of substrate materials, for example hybrid systems containing a liquid or gel or solid substrate, as well as aerosolizable substrate materials used within these aerosol provision systems; and aerosol-free delivery articles such as lozenges, gums, patches, articles comprising breathable powders and smokeless tobacco industry products such as snus and snuff, which aerosol-free delivery articles may or may not deliver nicotine.

In one embodiment the tobacco industry product may be prepared from (e.g. may comprise) a tobacco plant of the invention or a part thereof.

Suitably the tobacco plant or part thereof may be propagated from a tobacco plant propagation material according to the present invention.

The term "part thereof" as used herein in the context of a tobacco plant refers to a portion of the tobacco plant. Preferably the "part thereof" is a leaf of a tobacco plant.

In another embodiment the tobacco industry product may be prepared from a harvested leaf of the invention.

In a further embodiment the tobacco industry product may be prepared from a processed tobacco leaf of the invention.

Suitably the tobacco industry product may be prepared from a tobacco leaf processed by one or more of: curing, fermenting and/or pasteurising.

Suitably the tobacco industry product may comprise a cut tobacco leaf, optionally processed as per the foregoing embodiment.

In another embodiment, the tobacco industry product may be prepared from a tobacco cell culture according to the present invention.

In another embodiment, the tobacco industry product may be prepared from (e.g. may comprise) a cured tobacco material according to the present invention.

In another embodiment, the tobacco industry product may be prepared from (e.g. may comprise) a tobacco blend according to the present invention.

In one embodiment the tobacco industry product may be a smoking article.

As used herein, the term "smoking article" can include smokeable products, such as rolling tobacco, cigarettes, cigars and cigarillos whether based on tobacco, tobacco derivatives, expanded tobacco, reconstituted tobacco or tobacco substitutes.

In another embodiment the tobacco industry product may be a smokeless tobacco industry product.

The term "smokeless tobacco industry product" as used herein refers to a tobacco industry product that is not intended to be smoked and/or subjected to combustion.

Smokeless tobacco industry products (including heat-not-burn materials) may contain tobacco in any form, including dried particles, shreds, granules, powders, or slurry, deposited on, mixed in, surrounded by, or combined with other ingredients in any format, such as flakes, films, tabs, foams, or beads.

In one embodiment a smokeless tobacco industry product may include snus, snuff, chewing tobacco or the like.

In one embodiment, the tobacco industry product is a combustible smoking article, selected from the group consisting of a cigarette, a cigarillo and a cigar.

In one embodiment, the tobacco industry product comprises one or more components of a combustible smoking article, such as a filter, a filter rod, a filter rod segments, tobacco, a tobacco rod, a tobacco rod segment, a spill, an additive release component such as a capsule, a thread, beads, a paper such as a plug wrap, a tipping paper or a cigarette paper.

In one embodiment, the tobacco industry product is a non-combustible aerosol provision system.

In one embodiment, the tobacco industry product comprises one or more components of a non-combustible aerosol provision system, such as a heater and an aerosolizable substrate.

In one embodiment, the aerosol provision system is an electronic cigarette also known as a vaping device.

In one embodiment the electronic cigarette comprises a heater, a power supply capable of supplying power to the heater, an aerosolizable substrate such as a liquid or gel, a housing and optionally a mouthpiece.

In one embodiment the aerosolizable substrate is contained in a substrate container. In one embodiment the substrate container is combined with or comprises the heater.

In one embodiment, the tobacco industry product is a heating product which releases one or more compounds by heating, but not burning, a substrate material. The substrate material is an aerosolizable material which may be for example tobacco or other non-tobacco products, which may or may not contain nicotine. In one embodiment, the heating product is a tobacco heating product.

In one embodiment, the heating product is an electronic device.

In one embodiment, the tobacco heating product comprises a heater, a power supply capable of supplying power to the heater, an aerosolizable substrate such as a solid or gel material.

In one embodiment the heating product is a non-electronic article.

In one embodiment the heating product comprises an aerosolizable substrate such as a solid or gel material and a heat source which is capable of supplying heat energy to the aerosolizable substrate without any electronic means, such as by burning a combustion material, such as charcoal.

In one embodiment the heating product also comprises a filter capable of filtering the aerosol generated by heating the aerosolizable substrate.

In some embodiments the aerosolizable substrate material may comprise a vapour or aerosol generating agent or a humectant, such as glycerol, propylene glycol, triacetin or diethylene glycol.

In one embodiment, the tobacco industry product is a hybrid system to generate aerosol by heating, but not burning, a combination of substrate materials. The substrate materials may comprise for example solid, liquid or gel which may or may not contain nicotine. In one embodiment, the hybrid system comprises a liquid or gel substrate and a solid substrate. The solid substrate may be for example tobacco or other non-tobacco products, which may or may not contain nicotine. In one embodiment, the hybrid system comprises a liquid or gel substrate and tobacco.

In a further embodiment the tobacco industry product may be a tobacco heating device or hybrid device or e-cigarette or the like.

Typically in tobacco heating devices or hybrid devices, an aerosol is generated by the transfer of heat from a heat source to a physically separate aerosol-forming substrate or material, which may be located within, around or downstream of the heat source. During smoking, volatile compounds are released from the aerosol-forming substrate by heat transfer from the heat source and entrained in air drawn through the smoking article. As the released compounds cool, they condense to form an aerosol that is inhaled by the user.

Aerosol-generating articles and devices for consuming or smoking tobacco heating devices are known in the art. They can include, for example, electrically heated aerosol-generating devices in which an aerosol is generated by the transfer of heat from one or more electrical heating elements of the aerosol-generating device to the aerosol-forming substrate of a tobacco heating device.

Suitably the tobacco heating device may be an aerosol-generating device.

Preferably the tobacco heating device may be a heat-not-burn device. Heat-not-burn devices are known in the art and release compounds by heating, but not burning, tobacco.

An example of a suitable, heat-not-burn device may be one taught in WO2013/034459 or GB2515502 which are incorporated herein by reference.

In one embodiment the aerosol-forming substrate of a tobacco heating device may be a tobacco industry product in accordance with the present invention.

In one embodiment the tobacco heating device may be a hybrid device.

Polynucleotides/polypeptides/constructs

In certain embodiments of the present invention, constructs which modulate activity or expression of at least one FAD synthetase as described herein may be transformed into plant cells, suitably under the direction of a promoter.

In certain embodiments of the present invention, constructs which decrease (i.e. inhibit) the activity or expression of at least one FAD synthetase as described herein may be transformed into plant cells under the direction of a promoter. For example, the genetic construct may be a gene editing construct or may comprise an RNAi molecule, which may comprise a small interfering RNA (siRNA) molecule, or a short hairpin loop (shRNA) molecule.

In certain embodiments of the present invention, constructs which increase activity or expression of gene encoding a FAD synthetase as described herein may be transformed into plant cells, suitably under the direction of a promoter e.g. constructs which encode a gene encoding a FAD synthetase such as an endogenous FAD synthetase.

Constructs may be introduced into plants according to the present invention by means of suitable vector, e.g. plant transformation vectors. A plant transformation vector may comprise an expression cassette comprising 5'-3' in the direction of transcription, a promoter sequence, a construct sequence targeting gene encoding a FAD synthetase as described herein and, optionally a 3' untranslated, terminator sequence including a stop signal for RNA polymerase and a polyadenylation signal for polyadenylase. The promoter sequence may be present in one or more copies, and such copies may be identical or variants of a promoter sequence as described above. The terminator sequence may be obtained from plant, bacterial or viral genes.

Suitable terminator sequences are the pea *rbcS E9* terminator sequence, the *nos* terminator sequence derived from the *nopaline synthase* gene *of Agrobacterium tumefaciens* and the *35S terminator* sequence from cauliflower mosaic virus, for example. A person skilled in the art will be readily aware of other suitable terminator sequences.

The construct of the present invention may also comprise a gene expression enhancing mechanism to increase the strength of the promoter. An example of such an enhancer element is one derived from a portion of the promoter of the pea *plastocyanin* gene, and which is the subject of International Patent Application No. WO 97/20056 which is incorporated herein by reference. Suitable enhancer elements may be the *nos* enhancer element derived from the *nopaline synthase* gene *of Agrobacterium tumefaciens* and the *35S* enhancer element from cauliflower mosaic virus, for example.

These regulatory regions may be derived from the same gene as the promoter DNA sequence or may be derived from different genes, from *Nicotiana tabacum* or other organisms, for example from a plant of the family *Solanaceae*, or from the subfamily *Cestroideae*. All of the regulatory regions should be capable of operating in cells of the tissue to be transformed.

The promoter DNA sequence may be derived from the same gene as the gene of interest, e.g. the gene the promoter is going to direct, for instance a gene encoding a FAD synthetase according to the invention, a coding sequence used in the present invention or may be derived from a different gene, from *Nicotiana tabacum,* or another organism, for example from a plant of the family *Solanaceae*, or from the subfamily *Cestroideae*.

The expression cassette may be incorporated into a basic plant transformation vector, such as *pBIN 19 Plus*, *pBI 101*, pKYLX71:35S2, pCAMBIA2300 or other suitable plant transformation vectors known in the art. In addition to the expression cassette, the plant transformation vector will contain such sequences as are necessary for the transformation process. These may include the *Agrobacterium vir* genes, one or more T-DNA border sequences, and a selectable marker or other means of identifying transgenic plant cells.

The term "expression vector or plant transformation vector" means a construct capable of *in vivo* or *in vitro* expression. Preferably, the expression vector is incorporated in the genome of the organism. In one embodiment the vector of the present invention expresses a protein e.g. a FAD synthetase as described herein. The term "incorporated" preferably relates to stable incorporation into the genome.

Techniques for transforming plants are well known within the art and include *Agrobacterium-mediated* transformation, for example. The basic principle in the construction of genetically modified plants is to insert genetic information in the plant genome so as to obtain a stable maintenance of the inserted genetic material. A review of the general techniques may be found

in articles by Potrykus *(Annu Rev Plant Physiol Plant Mol Biol* [1991] 42:205-225) and Christon (AgroFood-Industry Hi-Tech March/April1994 17-27), which are incorporated herein by reference.

Typically, in *Agrobacterium-mediated* transformation a binary vector carrying a foreign DNA of interest, i.e. a construct according to the present invention, is transferred from an appropriate *Agrobacterium* strain to a target plant by the co-cultivation of the *Agrobacterium* with explants from the target plant. Transformed plant tissue is then regenerated on selection media, which selection media comprises a selectable marker and plant growth hormones. An alternative is the floral dip method (Clough & Bent, 1998 Plant J. 1998 Dec;16(6):735-43, which is incorporated herein by reference) whereby floral buds of an intact plant are brought into contact with a suspension of the *Agrobacterium* strain containing the chimeric gene, and following seed set, transformed individuals are germinated and identified by growth on selective media. Direct infection of plant tissues by *Agrobacterium* is a simple technique which has been widely employed and which is described in Butcher et al. (1980) *Tissue Culture Methods for Plant Pathologists,* eds.: D. S. Ingrams and J.P. Helgeson, 203-208 which is incorporated herein by reference.

Further suitable transformation methods include direct gene transfer into protoplasts using polyethylene glycol or electroporation techniques, particle bombardment, micro-injection and the use of silicon carbide fibres for example. Transforming plants using ballistic transformation and production of fertile transgenic maize plants by silicon carbide whisker-mediated transformation is taught in Frame et al. (1994) The Plant Journal 6(6): 941-948, which is incorporated herein by reference, and viral transformation techniques is taught in, for example, Meyer et al. (1992) Mol. Gen. Genet. 231(3): 345-352, which is incorporated herein by reference. The use of cassava mosaic virus as a vector system for plants is taught in Meyer et al. (1992) Gene 110: 213-217, which is incorporated herein by reference. Further teachings on plant transformation may be found in EP-A-0449375, incorporated herein by reference.

In a further aspect, the present invention relates to a vector system which carries a construct and introducing it into the genome of an organism, such as a plant, suitably a tobacco plant. The vector system may comprise one vector, but it may comprise two vectors. In the case of two vectors, the vector system is normally referred to as a binary vector system. Binary vector systems are described in further detail in Gynheung et al. (1980) Binary Vectors, Plant Molecular Biology Manual A3, 1-19, which is incorporated herein by reference.

One extensively employed system for transformation of plant cells uses the Ti plasmid from *Agrobacterium tumefaciens* or a Ri plasmid *from Agrobacterium rhizogenes* described by An et al. (1986) Plant Physiol. 81, 301-305 and Butcher et al. (1980) *Tissue Culture Methods for Plant*

Pathologists eds.: D. S. Ingrams and J.P. Helgeson, 203-208 which are incorporated herein by reference. After each introduction method of the desired exogenous gene according to the present invention in the plants, the presence and/or insertion of further DNA sequences may be necessary. The use of T-DNA for the transformation of plant cells has been intensively studied and is described in EP-A-120516; Hoekema (1985) The Binary Plant Vector System, Offset-drukkerij Kanters B. B., Amsterdam Chapter V; Fraley et al. Crit. Rev. Plant Sci. 4:1-46; and An et al. (1985) EMBO J 4: 277-284, all incorporated herein by reference.

Plant cells transformed with construct(s) which modulate the activity or expression of at least one FAD synthetase may be grown and maintained in accordance with well-known tissue culturing methods such as by culturing the cells in a suitable culture medium supplied with the necessary growth factors such as amino acids, plant hormones, vitamins, etc.

The term "transgenic plant" in relation to the present invention includes any plant that comprises a construct which modulates the activity or expression of at least one FAD synthetase according to the invention. Accordingly a transgenic plant is a plant which has been transformed with a construct according to the invention. Preferably the transgenic plant exhibits modulated (e.g. reduced) alkaloid content and/or modulated (e.g. reduced) TSNA precursor content according to the present invention. The term "transgenic plant" does not cover native nucleotide coding sequences in their natural environment when they are under the control of their native promoter which is also in its natural environment.

In one aspect, a gene encoding a FAD synthetase, a construct, a plant transformation vector or a plant cell according to the present invention is in an isolated form. The term "isolated" means that the sequence is at least substantially free from at least one other component with which the sequence is naturally associated in nature and as found in nature.

In one aspect, a gene encoding a FAD syntehtase, a construct, plant transformation vector or a plant cell according to the invention is in a purified form. The term "purified" means in a relatively pure state, e.g. at least about 90% pure, or at least about 95% pure or at least about 98% pure.

The term “nucleotide sequence” as used herein refers to an oligonucleotide sequence or polynucleotide sequence, and variant, homologues, fragments and derivatives thereof (such as portions thereof). The nucleotide sequence may be of genomic or synthetic or recombinant origin, which may be double-stranded or single-stranded whether representing the sense or anti-sense strand.

The term "nucleotide sequence" in relation to the present invention includes genomic DNA, cDNA, synthetic DNA, and RNA. Preferably it means DNA, more preferably cDNA sequence coding for the present invention.

In a preferred embodiment, the nucleotide sequence when relating to and when encompassed by the *per se* scope of the present invention, i.e. the gene encoding a FAD synthetase according to the present invention, includes the native nucleotide sequence when in its natural environment and when it is linked to its naturally associated sequence(s) that is/are also in its/their natural environment. For ease of reference, we shall call this preferred embodiment the "native nucleotide sequence". In this regard, the term "native nucleotide sequence" means an entire nucleotide sequence that is in its native environment and when operatively linked to an entire promoter with which it is naturally associated, which promoter is also in its native environment.

The nucleotide sequence for use in the present invention may be present in a vector in which the nucleotide sequence is operably linked to regulatory sequences capable of providing for the expression of the nucleotide sequence by a suitable host organism. The constructs for use in the present invention may be transformed into a suitable host cell as described herein to provide for expression of a polypeptide of the present invention. The choice of vector e.g. a plasmid, cosmid, or phage vector will often depend on the host cell into which it is to be introduced. Vectors may be used *in vitro*, for example for the production of RNA or used to transfect, transform, transduce or infect a host cell.

In some applications, the nucleotide sequence for use in the present invention is operably linked to a regulatory sequence which is capable of providing for the expression of the nucleotide sequence, such as by the chosen host cell. By way of example, the present invention covers a vector comprising the nucleotide sequence of gene encoding a FAD synthetase as described herein operably linked to such a regulatory sequence, i.e. the vector is an expression vector.

The term "operably linked" refers to a juxtaposition wherein the components described are in a relationship permitting them to function in their intended manner. A regulatory sequence "operably linked" to a coding sequence is ligated in such a way that expression of the coding sequence is achieved under conditions compatible with the control sequences.

The term "regulatory sequences" includes promoters and enhancers and other expression regulation signals. The term "promoter" is used in the normal sense of the art, e.g. an RNA polymerase binding site. The nucleotide sequence within a construct which encodes a FAD synthetase may be operably linked to at least a promoter.

The term "construct" - which is synonymous with terms such as "cassette" or "vector" - includes a nucleotide sequence for use according to the present invention directly or indirectly attached to a promoter.

An example of an indirect attachment is the provision of a suitable spacer group such as an intron sequence, such as the Sh1-intron or the ADH intron, intermediate the promoter and the

nucleotide sequence of the present invention. The same is true for the term "fused" in relation to the present invention which includes direct or indirect attachment. In some cases, the terms do not cover the natural combination of the nucleotide sequence coding for the protein ordinarily associated with the wild type gene promoter and when they are both in their natural environment. The construct may even contain or express a marker, which allows for the selection of the genetic construct.

In some embodiments, a promoter may be operably linked to nucleotide sequence in a construct or vector which is used to modulate the concentration and/or total content of nicotine in a cell or cell culture or tobacco plant or part thereof.

In some embodiments the promoter may be selected from the group consisting of: a constitutive promoter, a tissue-specific promoter, a developmentally-regulated promoter and an inducible promoter.

In one embodiment the promoter may be a constitutive promoter.

A constitutive promoter directs the expression of a gene throughout the various parts of a plant continuously during plant development, although the gene may not be expressed at the same level in all cell types. Examples of known constitutive promoters include those associated with the cauliflower mosaic virus 35S transcript (Odell JT, Nagy F, Chua NH. (1985). Identification of DNA sequences required for activity of the cauliflower mosaic virus 35S promoter. Nature. 313 810-2), the rice actin 1 gene (Zhang W, McElroy D, Wu R. (1991). Analysis of rice Act1 5' region activity in transgenic rice plants. Plant Cell 3 1155-65) and the maize ubiquitin 1 gene (Cornejo MJ, Luth D, Blankenship KM, Anderson OD, Blechl AE. (1993). Activity of a maize ubiquitin promoter in transgenic rice. Plant Molec. Biol. 23 567-81). Constitutive promoters such as the Carnation Etched Ring Virus (CERV) promoter (Hull R, Sadler J, LongstaffM (1986) (CaMV/35S), figwort mosaic virus 35S promoter. The sequence of carnation etched ring virus DNA: comparison with cauliflower mosaic virus and retroviruses. EMBO Journal, 5(2):3083-3090).

The constitutive promoter may be selected from a: a carnation etched ring virus (CERV) promoter, a cauliflower mosaic virus (CaMV 35S promoter), a promoter from the rice actin 1 gene or the maize ubiquitin 1 gene.

The promoter may be a tissue specific promoter. A tissue-specific promoter is one which directs the expression of a gene in one (or a few) parts of a plant, usually throughout the lifetime of those plant parts. The category of tissue-specific promoter commonly also includes promoters whose specificity is not absolute, i.e. they may also direct expression at a lower level in tissues other than the preferred tissue. Tissue specific promoters include the phaseolin-promoter,

legumin b4-promoter, usp-promoter, sbp-promoter, ST-LS1 promoter, B33 (patatin class I promoter).

In another embodiment the promoter may be a developmentally-regulated promoter.

A developmentally-regulated promoter directs a change in the expression of a gene in one or more parts of a plant at a specific time during plant development. The gene may be expressed in that plant part at other times at a different (usually lower) level, and may also be expressed in other plant parts.

In one embodiment the promoter may be an inducible promoter.

An inducible promoter is capable of directing the expression of a gene in response to an inducer. In the absence of the inducer the gene will not be expressed. The inducer may act directly upon the promoter sequence, or may act by counteracting the effect of a repressor molecule. The inducer may be a chemical agent such as a metabolite, a protein, a growth regulator (such as auxin and salicylic acid which activate the OCS promoter), or a toxic element, a physiological stress such as heat, light (such as the soybean SSU promoter), wounding (e.g. the nos, nopaline synthase promoter), or osmotic pressure, or an indirect consequence of the action of a pathogen or pest. A developmentally-regulated promoter might be described as a specific type of inducible promoter responding to an endogenous inducer produced by the plant or to an environmental stimulus at a particular point in the life cycle of the plant. Examples of known inducible promoters include those associated with wound response, such as described by Warner SA, Scott R, Draper J. ((1993) Plant J. 3 191-201), temperature response as disclosed by Benfey & Chua (1989) (Benfey, P.N., and Chua, N-H. ((1989) Science 244 174-181), and chemically induced, as described by Gatz ((1995) Methods in Cell Biol. 50 411-424).

A nucleotide sequence encoding either a protein which has the specific properties as gene encoding a FAD synthetase as defined herein or a protein which is suitable for modification may be identified and/or isolated and/or purified from any cell or organism producing said protein. Various methods are well known within the art for the identification and/or isolation and/or purification of nucleotide sequences. By way of example, PCR amplification techniques to prepare more of a sequence may be used once a suitable sequence has been identified and/or isolated and/or purified.

In a yet further alternative, the nucleotide sequence encoding the FAD synthetase may be prepared synthetically by established standard methods, e.g. the phosphoroamidite method described by Beucage et al. (1981) Tetrahedron Letters 22, 1859-1869 which is incorporated herein by reference, or the method described by Matthes et al. (1984) EMBO J. 3, 801-805 which is incorporated herein by reference. In the phosphoroamidite method, oligonucleotides

are synthesised, e.g. in an automatic DNA synthesiser, purified, annealed, ligated and cloned in appropriate vectors.

As used herein, the term "amino acid sequence" is synonymous with the term "polypeptide" and/or the term "protein". In some instances, the term "amino acid sequence" is synonymous with the term "peptide". In some instances, the term "amino acid sequence" is synonymous with the term "enzyme".

The present invention also encompasses the use of sequences having a degree of sequence identity or sequence homology with amino acid sequence(s) of a polypeptide having the specific properties defined herein or of any nucleotide sequence i.e. a gene encoding a FAD synthetase (hereinafter referred to as a "homologous sequence(s)"). Here, the term "homologue" means an entity having a certain homology with the subject amino acid sequences and the subject nucleotide sequences. Here, the term "homology" can be equated with "identity".

The homologous amino acid sequence and/or nucleotide sequence and/or fragments should provide and/or encode a polypeptide which retains the functional activity and/or enhances the activity of the FAD synthetase. Typically, the homologous sequences will comprise the same active sites etc. as the subject amino acid sequence for instance or will encode the same active sites. Although homology can also be considered in terms of similarity (i.e. amino acid residues having similar chemical properties/functions), in the context of the present invention it is preferred to express homology in terms of sequence identity. Homologous sequences typically retain functional domains or motifs.

In one embodiment, a homologous sequence is taken to include an amino acid sequence or nucleotide sequence which has one, two or several additions, deletions and/or substitutions compared with the subject sequence.

In one embodiment, homologues of FAD synthetase SEQ ID No. 3 are provided in Table 1. The present invention extends to the utilisation of homologues listed in Table 1 and to sequences having at least 80% sequence identity thereto.

Sequence identity

Sequence identity comparisons can be conducted by eye, or more usually, with the aid of readily available sequence comparison programs. These commercially available computer programs can calculate % homology between two or more sequences. % homology or % identity may be calculated over contiguous sequences, i.e. one sequence is aligned with the other sequence and each amino acid in one sequence is directly compared with the corresponding amino acid in the other sequence, one residue at a time. This is called an

"ungapped" alignment. Typically, such ungapped alignments are performed only over a relatively short number of residues.

Although this is a very simple and consistent method, it fails to take into consideration that, for example, in an otherwise identical pair of sequences, one insertion or deletion will cause the following amino acid residues to be put out of alignment, thus potentially resulting in a large reduction in % homology when a global alignment is performed. Consequently, most sequence comparison methods are designed to produce optimal alignments that take into consideration possible insertions and deletions without penalising unduly the overall homology score. This is achieved by inserting "gaps" in the sequence alignment to try to maximise local homology.

However, these more complex methods assign "gap penalties" to each gap that occurs in the alignment so that, for the same number of identical amino acids, a sequence alignment with as few gaps as possible - reflecting higher relatedness between the two compared sequences - will achieve a higher score than one with many gaps. "Affine gap costs" are typically used that charge a relatively high cost for the existence of a gap and a smaller penalty for each subsequent residue in the gap. This is the most commonly used gap scoring system. High gap penalties will of course produce optimised alignments with fewer gaps. Most alignment programs allow the gap penalties to be modified. However, it is preferred to use the default values when using such software for sequence comparisons.

Calculation of maximum % homology therefore firstly requires the production of an optimal alignment, taking into consideration gap penalties. A suitable computer program for carrying out such an alignment is the Vector NTI (Invitrogen Corp.). Examples of software that can perform sequence comparisons include, but are not limited to, the BLAST package (see Ausubel et al. (1999) Short Protocols in Molecular Biology, 4th Ed - Chapter 18), BLAST 2 (see FEMS Microbiol Lett 1999 174(2): 247-50; FEMS Microbiol Lett 1999 177(1): 187-8 and tatiana@ncbi.nlm.nih.gov), FASTA (Altschul et al. 1990 J. Mol. Biol. 403-410) and AlignX for example. At least BLAST, BLAST 2 and FASTA are available for offline and online searching (see Ausubel et al. 1999, pages 7-58 to 7-60).

Although the final % homology can be measured in terms of identity, the alignment process itself is typically not based on an all-or-nothing pair comparison. Instead, a scaled similarity score matrix is generally used that assigns scores to each pairwise comparison based on chemical similarity or evolutionary distance. An example of such a matrix commonly used is the BLOSUM62 matrix - the default matrix for the BLAST suite of programs. Vector NTI programs generally use either the public default values or a custom symbol comparison table if supplied (see user manual for further details). For some applications, it is preferred to use the default values for the Vector NTI package.

Alternatively, percentage homologies may be calculated using the multiple alignment feature in Vector NTI (Invitrogen Corp.), based on an algorithm, analogous to CLUSTAL (Higgins DG & Sharp PM (1988), Gene 73(1), 237-244). Once the software has produced an optimal alignment, it is possible to calculate % homology, preferably % sequence identity. The software typically does this as part of the sequence comparison and generates a numerical result.

Should gap penalties be used when determining sequence identity, then preferably the following parameters are used for pairwise alignment:

FOR BLAST	
GAP OPEN	0
GAP EXTENSION	0

FOR CLUSTAL	DNA	PROTEIN	
WORD SIZE	2	1	K triple
GAP PENALTY	15	10	
GAP EXTENSION	6.66	0.1	

In one embodiment, CLUSTAL may be used with the gap penalty and gap extension set as defined above. In some embodiments the gap penalties used for BLAST or CLUSTAL alignment may be different to those detailed above. The skilled person will appreciate that the standard parameters for performing BLAST and CLUSTAL alignments may change periodically and will be able to select appropriate parameters based on the standard parameters detailed for BLAST or CLUSTAL alignment algorithms at the time.

Suitably, the degree of identity with regard to a nucleotide sequence is determined over at least 50 contiguous nucleotides, preferably over at least 60 contiguous nucleotides, preferably over at least 70 contiguous nucleotides, preferably over at least 80 contiguous nucleotides, preferably over at least 90 contiguous nucleotides, preferably over at least 100 contiguous nucleotides, preferably over at least 150 contiguous nucleotides, preferably over at least 200 contiguous nucleotides, preferably over at least 250 contiguous nucleotides, preferably over at least 300 contiguous nucleotides, preferably over at least 350 contiguous nucleotides, preferably over at least 400 contiguous nucleotides, preferably over at least 450 contiguous nucleotides, preferably over at least 500 contiguous nucleotides, preferably over at least 550 contiguous nucleotides, preferably over at least 600 contiguous nucleotides, preferably over at least 650 contiguous nucleotides, or preferably over at least 700 contiguous nucleotides.

Suitably, the degree of identity with regard to a nucleotide, cDNA, cds or amino acid sequence may be determined over the whole sequence.

The sequences may also have deletions, insertions or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent substance. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues as long as the secondary binding activity of the substance is retained. For example, negatively charged amino acids include aspartic acid and glutamic acid; positively charged amino acids include lysine and arginine; and amino acids with uncharged polar head groups having similar hydrophilicity values include leucine, isoleucine, valine, glycine, alanine, asparagine, glutamine, serine, threonine, phenylalanine, and tyrosine.

Conservative substitutions may be made, for example according to the Table below. Amino acids in the same block in the second column and preferably in the same line in the third column may be substituted for each other:

ALIPHATIC	Non-polar	G A P
		I L V
	Polar – uncharged	C S T M
		N Q
	Polar – charged	D E
		K R
AROMATIC		H F W Y

The present invention also encompasses homologous substitution (substitution and replacement are both used herein to mean the interchange of an existing amino acid residue, with an alternative residue) that may occur i.e. like-for-like substitution such as basic for basic, acidic for acidic, polar for polar etc. Non-homologous substitution may also occur i.e. from one class of residue to another or alternatively involving the inclusion of unnatural amino acids such as ornithine (hereinafter referred to as Z), diaminobutyric acid ornithine (hereinafter referred to as B), norleucine ornithine (hereinafter referred to as O), pyriylalanine, thienylalanine, naphthylalanine and phenylglycine.

Replacements may also be made by unnatural amino acids include; alpha* and alpha-disubstituted* amino acids, N-alkyl amino acids*, lactic acid*, halide derivatives of natural amino acids such as trifluorotyrosine*, p-Cl-phenylalanine*, p-Br-phenylalanine*, p-I-phenylalanine*, L-allyl-glycine*, β-alanine*, L-α-amino butyric acid*, L-γ-amino butyric acid*, L-α-amino isobutyric acid*, L-ε-amino caproic acid#, 7-amino heptanoic acid*, L-methionine sulfone#*, L-norleucine*, L-norvaline*, p-nitro-L-phenylalanine*, L-hydroxyproline#, L-thioproline*, methyl derivatives of phenylalanine (Phe) such as 4-methyl-Phe*, pentamethyl-Phe*, L-Phe (4-amino)#, L-Tyr (methyl)*, L-Phe (4-isopropyl)*, L-Tic (1,2,3,4-tetrahydroisoquinoline-3-carboxyl acid)*, L-

diaminopropionic acid# and L-Phe (4-benzyl)*. The notation * has been utilised for the purpose of the discussion above (relating to homologous or non-homologous substitution), to indicate the hydrophobic nature of the derivative whereas # has been utilised to indicate the hydrophilic nature of the derivative, #* indicates amphipathic characteristics.

Variant amino acid sequences may include suitable spacer groups that may be inserted between any two amino acid residues of the sequence including alkyl groups such as methyl, ethyl or propyl groups in addition to amino acid spacers such as glycine or β-alanine residues. A further form of variation, involves the presence of one or more amino acid residues in peptoid form, which will be well understood by those skilled in the art. For the avoidance of doubt, "the peptoid form" is used to refer to variant amino acid residues wherein the α-carbon substituent group is on the residue's nitrogen atom rather than the α-carbon. Processes for preparing peptides in the peptoid form are known in the art, for example Simon et al. (1992) PNAS 89(20), 9367-9371 and Horwell (1995) Trends Biotechnol. 13(4), 132-134.

The nucleotide sequences for use in the present invention may include within them synthetic or modified nucleotides. A number of different types of modification to oligonucleotides are known in the art. These include methylphosphonate and phosphorothioate backbones and/or the addition of acridine or polylysine chains at the 3' and/or 5' ends of the molecule. For the purposes of the present invention, it is to be understood that the nucleotide sequences described herein may be modified by any method available in the art. Such modifications may be carried out in order to enhance the *in vivo* activity or life span of nucleotide sequences of the present invention.

The present invention also encompasses sequences that are complementary to the nucleic acid sequences of the present invention or sequences that are capable of hybridising either to the sequences of the present invention or to sequences that are complementary thereto. The term "hybridisation" as used herein shall include "the process by which a strand of nucleic acid joins with a complementary strand through base pairing" as well as the process of amplification as carried out in polymerase chain reaction (PCR) technologies.

The present invention also relates to nucleotide sequences that can hybridise to the nucleotide sequences of the present invention (including complementary sequences of those presented herein). Preferably, hybridisation is determined under stringency conditions (e.g. 50°C and 0.2xSSC {1xSSC = 0.15 M NaCl, 0.015 M Na_3citrate pH 7.0}). More preferably, hybridisation is determined under high stringency conditions (e.g. 65°C and 0.1xSSC {1xSSC = 0.15 M NaCl, 0.015 M Na_3citrate pH 7.0}).

A review of the general techniques used for transforming plants may be found in articles such as Potrykus et al. (1991) Annu Rev Plant Physiol. Plant Mol. Biol. 42:205-225 and Christou et al.

(1994) Agro-Food-Industry Hi-Tech March/April 17-27, which are incorporated herein by reference. Further teachings on plant transformation may be found in EP-A-0449375, incorporated herein by reference.

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this disclosure belongs. Singleton, et al., DICTIONARY OF MICROBIOLOGY AND MOLECULAR BIOLOGY, 20 ED., John Wiley and Sons, New York (1994), and Hale & Marham, THE HARPER COLLINS DICTIONARY OF BIOLOGY, Harper Perennial, NY (1991) provide one of skill with a general dictionary of many of the terms used in this disclosure.

This disclosure is not limited by the exemplary methods and materials disclosed herein, and any methods and materials similar or equivalent to those described herein can be used in the practice or testing of embodiments of this disclosure. Numeric ranges are inclusive of the numbers defining the range. Unless otherwise indicated, any nucleic acid sequences are written left to right in 5' to 3' orientation; amino acid sequences are written left to right in amino to carboxy orientation, respectively.

The headings provided herein are not limitations of the various aspects or embodiments of this disclosure which can be had by reference to the specification as a whole. Accordingly, the terms defined immediately below are more fully defined by reference to the specification as a whole.

Amino acids are referred to herein using the name of the amino acid, the three letter abbreviation or the single letter abbreviation. The term “protein", as used herein, includes proteins, polypeptides, and peptides. As used herein, the term “amino acid sequence” is synonymous with the term “polypeptide” and/or the term “protein”. In some instances, the term “amino acid sequence” is synonymous with the term “peptide”. In some instances, the term “amino acid sequence” is synonymous with the term “enzyme”.

In the present disclosure and claims, the conventional one-letter and three-letter codes for amino acid residues may be used. The 3-letter code for amino acids as defined in conformity with the IUPACIUB Joint Commission on Biochemical Nomenclature (JCBN). It is also understood that a polypeptide may be coded for by more than one nucleotide sequence due to the degeneracy of the genetic code.

Other definitions of terms may appear throughout the specification. Before the exemplary embodiments are described in more detail, it is to understand that this disclosure is not limited to particular embodiments described, as such may, of course, vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only,

and is not intended to be limiting, since the scope of the present disclosure will be limited only by the appended claims.

Where a range of values is provided, it is understood that each intervening value, to the tenth of the unit of the lower limit unless the context clearly dictates otherwise, between the upper and lower limits of that range is also specifically disclosed. Each smaller range between any stated value or intervening value in a stated range and any other stated or intervening value in that stated range is encompassed within this disclosure. The upper and lower limits of these smaller ranges may independently be included or excluded in the range, and each range where either, neither or both limits are included in the smaller ranges is also encompassed within this disclosure, subject to any specifically excluded limit in the stated range. Where the stated range includes one or both of the limits, ranges excluding either or both of those included limits are also included in this disclosure.

It must be noted that as used herein and in the appended claims, the singular forms "a", "an", and "the" include plural referents unless the context clearly dictates otherwise. Thus, for example, reference to "an enzyme" or “a nitrate reductase” includes a plurality of such candidate agents and equivalents thereof known to those skilled in the art, and so forth.

Advantages

It has been surprisingly found that by modulating the activity or expression of FAD synthetase as taught herein which acts as a positive regulator of alkaloid content in tobacco, the alkaloid and/or TSNA precursor content of plants can be modulated. Thereby tobacco industry products with modulated (e.g. reduced) alkaloid content) and/or reduced TSNA precursor content and commercially desirable traits sought after by consumers of tobacco industry products can be produced.

The present inventors have surprisingly determined a method for modulating the alkaloid content and/or TSNA precursor content of a plant (e.g. tobacco plant) by modulating the activity or expression of FAD synthetase. Alkaloid or TSNA precursor content of a plant (e.g. tobacco plant) may be decreased by decreasing or inhibiting the activity or expression of FAD synthetase as described herein. Prior to the present invention it had not been known that modulation of the activity or expression of FAD synthetase as described herein could be used to modulate alkaloid (and/or TSNA precursor content of a plant (e.g. a tobacco plant).

The publications discussed herein are provided solely for their disclosure prior to the filing date of the present application. Nothing herein is to be construed as an admission that such publications constitute prior art to the claims appended hereto.

EXAMPLES

Example 1 – Virus-induced gene silencing (VIGS) of Nitab4.5_0005997g0050.2 decreases alkaloid content in leaves

Virus-induced gene silencing (VIGS)

For virus induced gene silencing, a 300-nucleotide cDNA fragment of Nitab4.5_0005997g0050.2:

GGGATAGTAGCTCTGGGGAAGTTTGATGCACTCCATATTGGTCATCGTGAGCTTGCAATCCAAGCAGCTAAGAGAGG
AATTCCATTTCTTCTTTCATTTGTTGGAATGGCGGAAGTACTTGGATGGGAACCGAGGGCTCCCATTGTTGCTGACT
GTGATCGCAAAAGGATTCTTTCATCTTGGGCTCCTTATTGCGGTAATGTAATCCCAAGGGAATTCCAGATAGAATTT
TCCAAAGTTCGATCTCTTACGCCTCGTCAGTTTGTGGAGAAGCTGTCCAAAGAGCTGGGAGTGCGAGGA (SEQ ID No. 43)

was synthesized and cloned with In-Fusion cloning kit into pTV00 (between EcoRI and XhoI sites).The plasmid was then transformed into *A. tumefaciens* GV3101.

The TRV vector comprising both (TRV RNA1 SEQ ID No. 34) and (TRV RNA2) comprising the targeted nucleotide sequence were separately propagated in *A. tumefaciens*. These cultures were mixed (1:1) and syringe-infiltration into 2-week-old TN90 plants. The silencing effect was assessed five weeks post-virus infection by assessing the expression level of the target gene.

Silencing

VIGS assays were performed as previously described (Ratcliff et al., 2001) Ratcliff, F et al., (2001), The Plant Journal, 25: 237-245 (incorporated herein by reference). Briefly, independent cultures of *A. tumefaciens* GV3101 carrying TRV2 and TRV1 plasmids were propagated overnight in LB medium supplemented with appropriate antibiotics. Cultures were resuspended in VIGS buffer (10 mM morpholineethanesulfonic acid pH 5.6, 10 mM $MgCL_2$, and 100 µM acetocyringone) adjusting optical density to OD_{600}=1, and incubated overnight at room temperature in the dark. These cultures were mixed (1:1) and syringe-infiltrated into 2-week-old TN90 plants. The silencing effect was assessed two weeks post-virus infection by assessing the expression level of the target gene. TRV-Luciferase was used as a negative control and TRV-PDS (reduced chlorophyll content of the silenced leaves) was used as a phenotypic silencing control.

Alkaloid measurement

Relative content of pyridine alkaloids was determined by reversed phase high performance liquid chromatography with tandem mass spectrometry (LC-MS/MS). Chromatographic

separation was achieved using a Gemini-NX column (100 mm × 3.0 mm, particle size 3 µm, Phenomenex) and gradient chromatographic separation using 6.5 mM ammonium acetate buffer (aq) (pH10) and Methanol.

Mass Spectrometer operates in electrospray (ESI) positive mode using scheduled MRM data acquisition. Two MRM transitions were monitored for each analyte and one for the isotope labelled internal standard.

Analyte	Precursor Ion	Daughter Ion (quant/confirm)
Nicotine	163.1	130/106
Nicotine d4	167.1	134.1
Anabasine	163.1	80/120
Anatabine	161.1	144/80
Nornicotine	149.1	80/130
Nornicotine d4	153.1	84.1
PON	176.1	106.0/148
PON d4	183.1	110.0

Results

Alkaloid content of 5-week-old tobacco leaves silenced for Nitab4.5_0005997g0050.2 is shown in Figure 1. Content is represented relative to control and comprises three biological replicates analysed by one-way ANOVA and Tukey's multiple-comparison post-test. Values are shown as means ± SEM. Asterisks indicate statistical significance of P value ≤ 0.001. *Pyridine alkaloids: nicotine, nornicotine, anabasine (ANAB), anatabine (ANAT), pseudooxynicotine (PON).*

VIGS of Nitab4.5_0005997g0050.2 leads to a decrease in alkaloid content in leaves, in particular a decrease in nicotine, nornicotine, anabasine, PON and anatabine content.

Conclusions

Nitab4.5_0005997g0050.2 is a positive regulator of alkaloid content, in particular alkaloid content in leaves and is a regulator of pyridine alkaloids in tobacco.

Example 2 – Nitab4.5_0005997g0050.2 has FAD synthetase functional activity

The present inventors characterised Nitab4.5_0005997g0050.2 as a gene encoding for FAD synthetase. This enzyme catalyses the adenylation of flavin mononucleotide (FMN) to form flavin adenine dinucleotide (FAD) coenzyme (ATP + FMN → diphosphate + FAD).

To confirm the categorisation a functional assay was performed to determine the activity of Nitab4.5_0005997g0050.2. The gene was PCR-amplified from N. tabacum and cloned into pMAL-CX5 to allow expression in *E. coli* as N-terminal maltose binding protein (MBP) fusion. Native, untagged MBP was also expressed and purified as a control for the assay.The Flavin Adenine Dinucleotide (FAD) Assay Kit from Abcam (ab204710), was used to confirm the enzymatic activity. The assay measures the formation of FAD by reaction with an OxiRed probe that results in an increase in absorbance at 570 nm.

Following lysis and affinity purification, 0.5 µg of Nitab4.5_0005997g0050.2 enzyme was added to 10 µM ATP and 10 µM FMN. The production of FAD over 60 minutes was measured by endpoint absorbance values at 570 nm.

Results

Figure 2 shows measured endpoint absorbance values (570 nm) for the production of FAD over a 60 minute reaction by 0.5 µg of Nitab4.5_0005997g0050.2 enzyme in the presence of 10 µM ATP and 10 µM FMN. The results for the control reactions (MBP and –VE, no enzyme) are also shown on the same graph and indicate that the reaction is FMN-dependant and that the reaction has reduced activity in the absence of added ATP. Error bars are standard deviation from four replicates.

Conclusion

Nitab4.5_0005997g0050.2 has FAD synthetase functional activity, suggesting pyridine alkaloid synthesis is correlated with FAD synthetase levels.

Example 3 – Homologue testing

The effects of the homologues of SEQ ID No. 3, namely those listed in Table 1, are tested in assays as described in the above examples.

CLAIMS

1. A method of modulating (e.g. decreasing) the alkaloid content of a tobacco plant or a part thereof or tobacco plant cell, the method comprising modifying said plant or plant cell by modulating (e.g. decreasing) the activity or expression of a FAD synthetase.

2. A method of modulating (e.g. decreasing) the content of a tobacco specific nitrosamine (TSNA) or a precursor of a TSNA in a tobacco plant or plant part thereof or tobacco plant cell, the method comprising modifying said plant or plant cell by modulating (e.g. decreasing) the activity or expression of a FAD synthetase.

3. A method for producing a plant or part thereof, a cell or cell culture, a plant propagation material, a leaf, a cut harvested leaf, a processed leaf or a cut and processed leaf which has modulated (e.g. decreased) alkaloid content, the method comprising modifying said plant or cell culture to modulate the activity or expression of a FAD synthetase.

4. Use of at least one gene encoding a FAD synthetase for modulating alkaloid content of a tobacco cell or tobacco plant or part thereof.

5. A method or use according to any preceding claim, wherein the FAD synthetase:

a) comprises an amino acid sequence as set out in SEQ ID No. 3; or a functional variant or functional fragment or orthologue of SEQ ID No. 3; or a sequence which has at least 80% identity to SEQ ID No. 3, or a homologue of SEQ ID No. 3; or

b) is encoded by a nucleotide sequence as set out in SEQ ID No. 1 or 2; or a functional variant or functional fragment or orthologue of SEQ ID No. 1 or 2; or a nucleic acid sequence which has at least 80% identity to SEQ ID No. 1 or 2, or a homologue of SEQ ID No. 1 or 2.

6. A method or use according to any one of the preceding claims, wherein the alkaloid content is modulated (e.g. decreased) in comparison to a plant or cell culture which has not been modified to modulate the activity or expression of a FAD synthetase.

7. A tobacco plant or part thereof or a tobacco cell or cell culture which has been modified to modulate (e.g. decrease) the activity or expression of a FAD synthetase and the tobacco plant or part thereof or tobacco cell or cell culture has decreased alkaloid and or TSNA precursor content in comparison to an unmodified plant or unmodified cell or cell culture.

8. A plant propagation material obtainable (e.g. obtained) from a plant according to claim 7 or from a plant or cell or cell culture produced by the method or use of any one of claims 1 to 6.

9. A method or use according to any one of claims 1-6, or a plant or part thereof or cell or cell culture according to claim 7, or a plant propagation material according to claim 8, wherein the alkaloid content of the plant is decreased in comparison to a plant or cell culture which has not been modified to modulate the activity or expression of a FAD synthetase.

10. A method or use according to any one of claims 1-6 or 9, a tobacco plant or part thereof or tobacco cell or cell culture according to claim 7 or 9, or a plant propagation material according to claim 8, wherein the content of one or more alkaloids selected from nicotine, nornicotine, PON, anabasine, myosmine and anatabine is modulated (e.g. decreased), preferably the content of nicotine, nornicotine and/or PON is modulated (e.g. decreased).

11. A method or use according to claim 10, a tobacco plant or part thereof or tobacco cell or tobacco cell culture according to claim 10, or a plant propagation material according to claim 10 wherein the nicotine content is decreased.

12. Use of a tobacco plant or part thereof or tobacco cell or cell culture according to any one of claims 7 or 9-11, or of a plant produced by the method of any one of claims 1 to 6, or 9 to 11 to breed a plant.

13. Use of a tobacco plant or part thereof or a tobacco cell or cell culture according to any one of claims 7 or 9-11, or of a plant produced by the method of any one of claims 1 to 6, or 9 to 11 for production of a product.

14. Use of a tobacco plant or part thereof according to any one of claims 7 or 9-11, or of a plant produced by the method of any one of claims 1 to 6, or 9 to 11 to grow a crop.

15. Use of a tobacco plant or part thereof according to any one of claims 7 or 9-11, or of a plant produced by the method of any one of claims 1 to 6, or 9 to 11 to produce a leaf.

16. A harvested leaf of a plant according to any one of claims 7 or 9-11, or obtainable from a plant propagated from a propagation material according to any one of claims 8-11, or obtainable

from a plant obtained by a use according to any one of claims 4-6 or 9-11, or obtainable from a plant produced by the method of any one of claims 1-6, or 9-11.

17. A harvested leaf of a plant according to claim 16, wherein the harvested leaf of a plant is a cut harvested leaf.

18. A processed leaf, preferably a processed tobacco leaf, preferably a non-viable processed tobacco leaf:

obtainable (e.g. obtained) from a plant obtainable from a use according to any one of claims 4-6 or 9-11;

obtainable (e.g. obtained) by processing a plant according to any one of claims 7 or 9-11;

obtainable (e.g. obtained) from a plant propagated from a plant propagation material according to any one of claims 8-11; or

obtainable (e.g. obtained) by processing a harvested leaf of a plant according to claim 16 or 17; or

obtainable (e.g. obtained) from a plant produced by the method of any one of claims 1 to 6, or 9 to 11.

19. A processed leaf according to claim 18, wherein the leaf is processed by curing, fermenting, pasteurising or a combination thereof.

20. A processed leaf according to claim 18 or 19, wherein the processed leaf is a cut processed leaf.

21. Cured tobacco material made from a plant or a part thereof according to any one of claims 7 or 9-11, or a harvested leaf according to claim 16 or 17, or a processed leaf according to any of claims 18-20.

22. A tobacco blend comprising said cured tobacco material of claim 21.

23. A tobacco industry product prepared from:

a tobacco plant or part thereof or tobacco cell or cell culture according to any one of claims 7 or 9-11;

a tobacco plant or part thereof propagated from a tobacco plant propagation material according to claim 8;
a harvested leaf of a plant according to claim 16 or 17;
a processed leaf according to any one of claims 18-20.

24. A tobacco industry product according to claim 23, wherein the tobacco product is a combustible smoking article.

25. A tobacco industry product according to claim 23, wherein the tobacco product is a smokeless tobacco product.

26. A tobacco product according to claim 23, wherein the tobacco product is a non-combustible aerosol provision system such as a tobacco heating device or an aerosol-generating device.

27. A combustible smoking article, non-combustible aerosol provisioning system, smokeless tobacco product or tobacco heating device comprising a plant or a part thereof according to any one of claims 7 or 9-11 or an extract (e.g. a tobacco extract) thereof or a tobacco cell culture according to any one of claims 7 or 9-11; or a cured tobacco material according to claim 21; or a tobacco blend according to claim 22.

28. Use of a nucleotide sequence encoding a FAD synthetase protein which:
a) encodes an amino acid sequence as set out in SEQ ID No. 3; or a functional variant or functional fragment or orthologue of SEQ ID No. 3; or a sequence which has at least 80% identity to SEQ ID No. 3; or
b) comprises a sequence as set out in SEQ ID No. 1 or 2; or a functional variant or functional fragment or orthologue of SEQ ID No. 1 or 2; or a nucleic acid sequence which has at least 80% identity to SEQ ID No. 1 or 2;
to select a plant having modulated (e.g. reduced) alkaloid content and/or modulated (e.g. reduced) content of TSNA or a precursor of a TSNA.

29. A mutant of a plant carrying a heritable mutation in a nucleotide sequence which:
a) encodes an amino acid sequence as set out in SEQ ID No. 3; or a functional variant or functional fragment or orthologue of SEQ ID No. 3; or a sequence which has at least 80% identity to SEQ ID No. 3; or

b) comprises a sequence as set out in SEQ ID No. 1 or 2; or a functional variant or functional fragment or orthologue of SEQ ID No. 1 or 2; or a nucleic acid sequence which has at least 80% identity to SEQ ID No. 1 or 2;

wherein said heritable mutation modulates (e.g. decreases) the activity or expression of FAD synthetase and wherein the mutant plant has modulated (e.g. decreased) alkaloid content and/or modulated content of a TSNA or a precursor of a TSNA relative to a comparable plant which does not carry said heritable mutation.

30. Progeny or seed of a mutant plant which carries the heritable mutation according to claim 29.

31. A harvested leaf, a processed leaf or cured tobacco material produced from a plant comprising a modification in a nucleotide sequence which:

a) encodes an amino acid sequence as set out in SEQ ID No. 3; or a functional variant or functional fragment or orthologue of SEQ ID No. 3; or a sequence which has at least 80% identity to SEQ ID No. 3; or

b) comprises a sequence as set out in SEQ ID No. 1 or 2; or a functional variant or functional fragment or orthologue of SEQ ID No. 1 or 2; or a nucleic acid sequence which has at least 80% identity to SEQ ID No. 1 or 2;

wherein said modification modulates (e.g. decreases) the activity or expression of FAD synthetase and wherein said plant has modulated (e.g. decreased) alkaloid content and/or modulated content of a TSNA or a precursor of a TSNA relative to a comparable plant which does not carry said modification in said FAD synthetase.

Figure 1

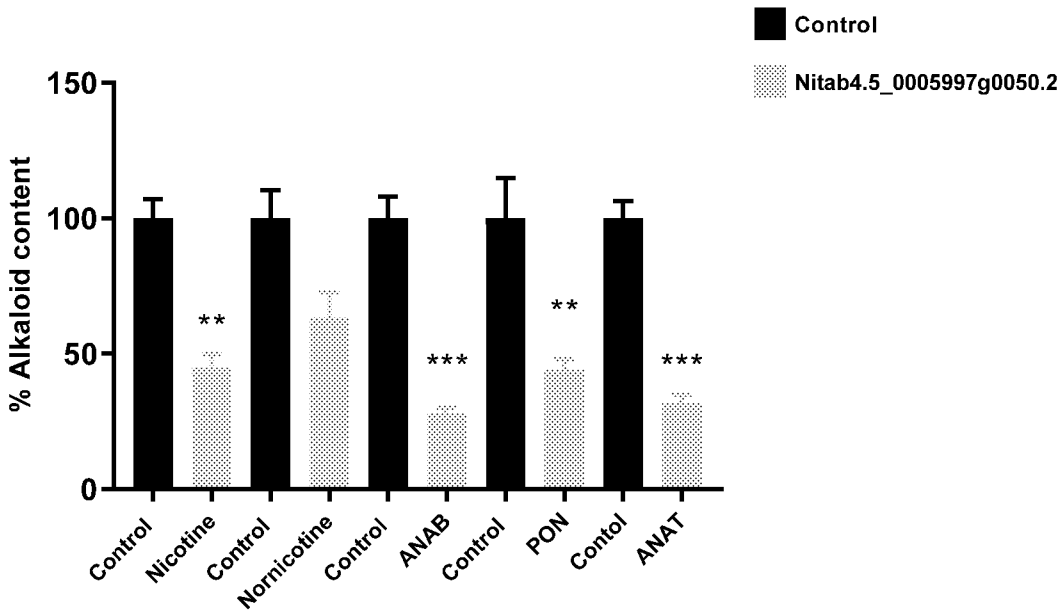

Figure 2

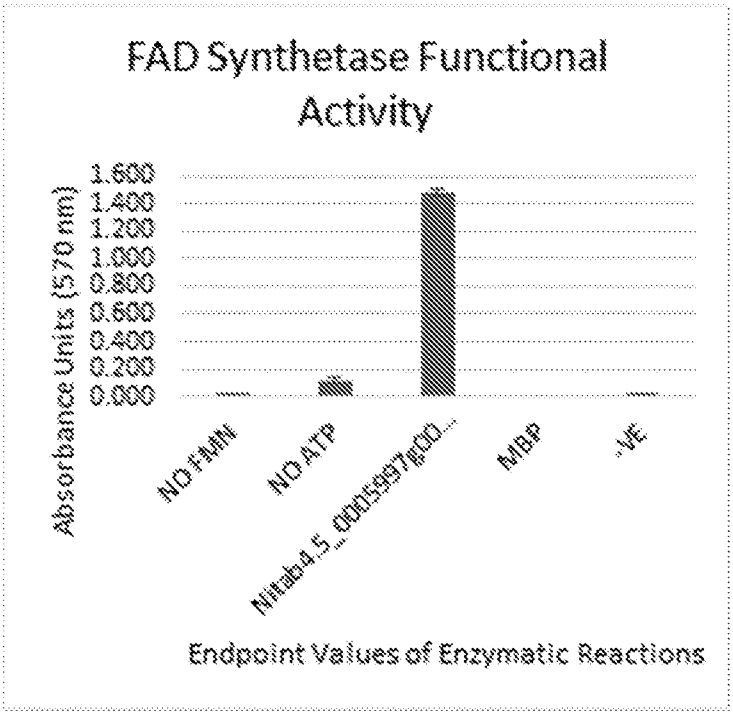

Figure 3

Nitab4.5_0005997g0050.2 Genomic sequence (SEQ ID NO: 1):

TTAGTAAAACATATATTATTAAATTCCTCAGCCACCCTTCTCTCTCCCCAACAAAAGAAAAATAAACA
AAAAGGAGAATCAAAACTATCGATACAAATAAAACCAAGAAAAAAAATTGCTACAGCAATAGGCCGGC
CAACACCCATTTCCATATTCTTCTTCTGATGAAAAATAAGAATGTTGGGTGGGTCTCGCTGTATATCC
CAGCACTTAAGAGACTCTAATTTTCACCAATTCTGCACTTCTACAACTCCCATTTTCACTGCTGCTTC
TCCACCAATCTCAACTTCTTCATTCAACTACTTCAGATTTAGATTAAGAAATAAAGTTAGTGATTTTG
GGGTCAATTGTGTTCAGCCCAATTCTTCAAATACCCAAAGCTCTTTGCTTTCTCAAACTTTGAGGTAA
ATCTTATAGTAACATCTTTTACTATTTCAGCAGCAATTTTAGTGATTTTTTATTTGAGCTGTAAATTT
AGTCTTTTTGGGTACTAGCCAAATTTTCAGAATATTCAAAGCTCTTTACTTTCTGAAAGTTTGAGGTA
ATTTGTACATCTTTTACTCTTTCAGCAGTAAATTTAGTGCTTTTTGGGTGTACATTTGAGCTGTAAAT
TTGGTTTTTTTTGGGTACTTTCTATGAATTGTGTACTAGCCAAATCTTCAGAATATCCAAAGCTCTTT
GCTTTCTGAATGTTTGAGGTAAATTTGTAACATCTTTATTTTTTAGCAGTAAAATTAGTGATTTTTCG
GGTAATTTGAGCAGTAAATTTAGTTCTTTTGGGGTACTTTTGTGAATTGTGTATAATCCAAATCTTCA
GAATATCTAAAGCTCTTTATTTTCTGAAAGTTTGAGTCAATTTGTCATAATTTTTCTGCATTGAGCAG
TAAATTTAGTTTATTTTTTGTGGTAGTTTGAGGTATTTATCAGAGATGTTCTTATTTAAGCAGTGATT
TCAGTTATTTGGGGGTGGTTTTTTGTCAATTTTGTACGACCCAGATATTCAGAATACAGGAAGCTCTT
AACTTTCTGGAAGTTTGAGGTAATTTAACATAGTTCTTGTTTGAGCAATAAATCGAGTGATTTTGGCC
CCAGGCTTTCGTCTAGTGGTAAGAGCGCAGCACGGGATTCGAACCCTGCAGTGGATGTTGGTCTTTAA
GTGAGGAAGGGCAGAGGAGCAGGCTCGTTATGCATTGAGTTTCGAACTGTGCACCACTGACGTTCAGG
AATTTCTCGATATTATAAGAATAAAAATTGAGTGAGTTTGGGGTCTGTTTTATCAACTGTGTGCAACC
AGAATACAAAGCTCCTTGCTTTCTAATAGTTTGAGATAAATTTTATCTCTATTACACTGGTATTTTAA
GACCTTTTGTCATTTAAAAGGAAAGTGTAGAAAGGAAGTTTTTTAAGAAAACTAGAATTATAAGCTCA
TTATAAAAAATAGCTCATTATTCTTGCGAATAGAGTAAGCGAAAGAAGGGCTGGAAGGAGATTTTCTT
TTATAACTGTAATAATATTGAAGTGGAAAATTTGTTGCAACATTAAGCTGGGATTACATTTTTGACGA
AAATAATATGATAAGTTTATTGGAATATCATATAAGTAGGTGGGTATGCGTTTGTGCATACCAAACTG
TGGTGTAAATATAATTATTTCTCTCAGACAGTTTAAATTTTAGATGAGAGGTTTCACACGCTTCAACC
AGATAACAGAGCAGACACATGCCTGTGTTTGAATCTCATCGCGTGTCATCTATAAAAATATTCGGACG
TGCTTGGCCCACTAAAAAGATTCAGGCCTTGCACATGAGAGAGCATGTTGTAGACCTAAAACAAATAA
ATAAGACTGTCTTCCTTCTAATTGTTATAAGTTTTTAGATGAGAGGGTTATACATTTAAATAGTTTCA
CCTGGACATAGATTTTGCCGGTAGTTTCCATTTTATGTTTAAATGAAGCTTACCGAGGTTGGGTTTCC
ATTAAATCTAAGTCAGAGTTCTTTATTCCATTGCATTGACATTGTAATGTCATTTTGTATTCTACAGT
ATGATTTACTTTAGAATCCCAGTGCTCATATGTGTTTTCTTCACAATTTTTATGTAAAGCCAATCTGG
AAAAGATGAGGAGCCTCCATCAGAAAGATTGCCAATGCTTGCAGGTGTGTTCTAACCCATATACTACT
TGTAATAATATTATTAAATGGTATCATATGTATGCTGAGTATATATCAATACCCATTTTGTTTTATCA
TTTATTATAAATCATTAAGTTTCTTGAGTGATTAAGGGCTACACAGATACCTTGTTCTACAATTGTTC
TCAGCCATTCTTATTTCATAATTCTTGGTTACGTTGCCATTGGCTCTTGTGCTCATTCTTTCCTTTTG
TAATTTTTTTCTTTTCCTATTCCTGGTTCTCTAAAGAAAATGGGGAAAGAGGATAACAAGATTTCCTG
CAATAATGGAAGGTTCTTCTACTTCAGTTCTAACATAAGCTGGAGATCAACTATCTCGGCAAAGCATG
AAATTTCACACAATATATGACCAACCATCAGATTTCCTGCAATCACTTCAGGAAATTACAAAATATGT
TCCACTGCGAGGACATCTTGTCTTTATGTTCTTGGATAACTTCCATTTCAAGGTTAGGAAAATTTCTC
TCACCAAAGCACACCTAGAGGTATTGGTAACATGAAGGGTATGTATGACCCTGTGAAAGCTTCAGTTT
AATGCCTTCCTCTGACAGAAATTTGACCACAAGATCGTAGAAAAGAAAAAGTGAAAGTCATTGATACA
TGCTCCAACATGGGTCATTATAAAGGGAAATATTGTCCATTCTTTCCTCAATTTGTTTTGCTAACACT
TTAGACCTGTTCTTTCCATACCGTGCACTTATTTTTGTATAAGAAGAAATAGGGATAATCTAGCTGTG
AACAGTTGAATGGTCAATCTGCAAATTTTTTTCTAAATTTACTTTGGACCTACAATTACATTTTTATA
TTAGTGAGTTTCATTTCTACTTGTTTCTTATCATCAATAGAAAAAAAAAGAGTTTCAATTCCACTTGC
TACTTTGGGTTGTAAATGAGAAGAAATAGAGGATTCTTGGATATCATTTATGTGTAAATTCTCGATCA
TTGGTATCATTTCAATAGGCGGGATAGTAGCTCTGGGGAAGTTTGATGCACTCCATATTGGTCATCGT
GAGCTTGCAATCCAAGCAGCTAAGAGAGGAATTCCATTTCTTCTTTCATTTGTTGGAATGGCGGAAGT
ACTTGGATGGGAACCGAGGTATAACTTCTGTTCTCTCTTCTACTTTACTCTACTCTAATAGAATGTTA
ATAGAGAAATATTCTCTCCCTGGAAGACCATAAAAGTAGAGAATGGTGAACTTCGACTCAACTCTCGC
TGCTACACCCAAACCTGTAATAGTAGCTAGGTTCAACCAGGGAGTTCTATCTCGCATTCTCTAAATGT
GTGTGTGTGTGTGAGTGAGTGTGTGTGTGAGAGAGAGAGAGGGAGGGAGGGATAAATGTGTGTGTG

TGAGAGAGAGAGAGAGAAAGAGCGAGCTAGAGAGCTGTCTCTTTGCAACCTCTTAAGTGGTAGCAATG
CTGCACCTTCATGCCTCTATTTAAAGCATAATTCTGTGGACCTACAGATGCTCTTTCTCCGAAGCACT
AGATTTGTTGACTTGCACTCTTTAGTTTAGGAATTTCTGTTAATAGGATGATCAATTCAGAGTATGTG
CAAAGAGCACGAAAATATAACCATATGTTTTTAATTCAAGATTTAACAAAAAGGATAAATTCTTCATA
AAAAAAAGGATAAATTGAAGAAATAAATTACTGGCTGAACGCATAGATGGCCCTTTAAACTTGCCCTA
TTTTTCCATTAAGACACCTGAACTAATGCTTGTACCTATTGAGCACTTCACCTATATTCAAGCTGTAT
CAATTGAACACATATTTTATAATCTAAACTAAGTTCAACACATGAGTACCTCTCTCTGTATGTAGCAA
ATATGACCCATAAGAAAGAGCCACGTTAGATAAGGAAAAAGTCAATATCTGGAGGAAAAAGGAAACAG
TAACCAGAAACAGAAGAATCCGGTGAAAAAGGGGACAATACCAGCCGGAATTACGTTTTCTTTGTTAT
CAAATTATCATCACCCCGACACTGGCAAAAACAAGTTCATCCCAGATTGACCCTGAATCCCCGTCAAC
TGCCTGGGCGACTCCATTTTCACCGGATAAAGTTGCCTTGTCACCAGTATCAATGAATCATCATTTTT
GTCATCTCCAAGCTCGATTGAGTCATCTTTGGAGAAAGGTTATTTTGGGGGGTAAGTCTCATTTTGAG
TTCACAAATTGCTGAAACATTCAGATCCCCCTTTTATCCCTTTTTTTTTCTTCTTCCATTCGCAAATA
CTCACCATCCTGGTTTGCCAATGTTCCACCACTGAGGAATCTCACCCCCATGACATCAACCGATAGGG
ACCAGATTACATCCTTCCTCCACAAATCACCATTACATACACAAAGACCAATCAACTAACCTCTCTCC
CCTAATTGATGGATTTGTAGATAATACAACAACAACAACGATTCAGTGACATTCCACTAGTGGGGTCC
GGGGAGGGTAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNGAAGACAAAAAGACAAAAGGGGACAATATCAGTAACAAC
ACAGAAATAATAGAAAAATAGGAACAACATGCAATCAAGAAGAATGATACACGATTTGTAGATAATAC
GTGAATGTATATATTACTCTTGTATTGTTCATTTAGAATTTTGTATAGTAGCTTTCTTGTTTAAAAGA
TAAGATTTTGAAAATTTGCTAAAAAGGTACGACTATTATTATTCTTGAGGGAGAATGGGTCAGTTGAC
GTATTTTTCCTTGGCTATATCATGGACCGTTTTCTGATGAAGGTGGCGCATATTATTTAGGCGGGTCA
GAAGAGAGAATGGTGTCTGAATAGAATGATGGAGAGGAGTGGTGTTTGATAGGCTTTCTATAGGATAA
ATAAGAAGAAGATGAAATACTTTATTGGATTTGAATCTGGTATGGAGTCCAAAAAAGAAGAGGAAAAT
GAGTCGGAGAAGATGACGGAAAATAGCAGTTATCAGAAATAGCCAAAAAATGAATTTACCCATGATAA
CTTCATATTTGTTTTTCAGTAAAATGCGCATAGTTTTGGGCGAAAGACACGCTATGTGCCATGTCAGC
GAGTTGTGTCTCATAGGCACAGTTTGGCTCAAGTTCAAGTGTTCAATAGGCACAATAGTTAATTGAGG
TGTCTAGATGATAAAATGGGACAAGTTTAAAGGGCTGCATATGTATTTGGCTTAAATCAATTAATTCC
TGCAACAGTGTATATTCCCACAGGAAAACTTAAGTGGATGTCCTTTTTTCTTTCTTTCTGTTTTTAGT
GACTAGTAAGAGTTCCTCCATCTTAAATTCAAATGATGTCCTTAACCTCATGTTACAGAACCTTGATT
TGCTCCTTGAAAATAAAATATTTATCGAACTTTATTCTGCTCAAGTAGTATATAGGAAAAGTCATGGG
TTATAAAATTTGATCTCTGAAGCTGTAAGGTTTTTATAATCTGCTGAACATGTCCTCCTAACCTTACT
CAAGCGTGTTGAGTAGTTCCTTCTTTAGTTAGTGGTCTTTATGAAATGCTACATCATGAAATGAGAAA
GAATCATTGCATCATTTTGTGTATATTTTCTACACAGGGCTCCCATTGTTGCTGACTGTGATCGCAAA
AGGATTCTTTCATCTTGGGCTCCTTATTGCGGTAATGTAATCCCAAGGGAATTCCAGATAGAATTTTC
CAAAGTTCGATCTCTTACGCCTCGTCAGTTTGTGGAGAAGCTGTCCAAAGAGCTGGGAGTGCGAGGAG
TTGTTGCTGGTATGTTGTTACATCCTAACTTGACATAGAGATTCCAATATAAACTAAGAAAAGAAAAG
AATATATGAGGATGTCTTGATGATCTCAGGCGAAAACTATCGTTTTGGATATAGAGCTTCTGGTGATG
CATCGGACCTTGTGAAGCTCTGTGAAGAGTATGGATTAGAGGCTTATATAATCAATTCTGTCATGGAC
AAGAATCAAATTTCTCGAGCCTTAAACTCTCACGGAAAAAAGGAGAGAGGGCAAGTATCATCTACTCG
TGTTAGGTATGCCCTTGACAAGGGAGACATGAAATATGTGTCAGAGCTATTAGGTCGCAACCATCGTC
TTGTTTTAACGATGGAGGACCAAGAAAGGATTATTAGTGAGAGAAATAGGCTGTCAGCTCCCAAGTCT
TGTTTGTTGAATCTTGCACCCAAGGAAGGTCTTTATGAGAATTGTTCAGTTTCGATTGACGAGAATGT
TGTACCCTGCAGAGTCATAGTTGATACAACTGGTATTCATTTGGAATGCTATGAGGTAGCTAATTTTA
CTTGTATTACTTCTCGAGACGTAAAGATACTAGGTATTGACTTTGGTAGTCAATAGTTTGATGAGGTT
CAAATTATGTAATTTCATTCCCCATAATTCTTTAGTGCGACCATTTCTGAATGTCCTTCAGTCTTTAG
GTATGTGTACGATTCTTTATGGTACATTGTTGCACAATTCACATACTTTTGGTGGGAACATAGTTGAC
TTTACTGTTATTTTCTGAATGCCAAAAATATGGTAGTATGAACAGAAACGAGCTGATAAACGTGCAAG
AAAATAGCACATTTCATGAAGAAGTGACCTTTAATCATTACCATACATAATGGATGTTGTTCTATCAC
TAAGAGTGGACTGAGTTAGAAACTTGTTTACTGTTTTTCTTTGTTCGAGTGTTTTTGGATTTGGTCTA
ATTTGTGACAATCCCTTGATTATATTCCTGTTGTACTTGATTGTCTGTTATCCTTTTTGCCAATGAAA
TGATCATAGTGCTTGAGTTGTCCATTT

Figure 4

Nitab4.5_0005997g0050.2 Transcript: (SEQ ID NO: 2)

ATGTTGGGTGGGTCTCGCTGTATATCCCAGCACTTAAGAGACTCTAATTTTCACCAATTCTGCACTTC
TACAACTCCCATTTTCACTGCTGCTTCTCCACCAATCTCAACTTCTTCATTCAACTACTTCAGATTTA
GATTAAGAAATAAAGTTAGTGATTTTGGGGTCAATTGTGTTCAGCCCAATTCTTCAAATACCCAAAGC
TCTTTGCTTTCTCAAACTTTGAGCCAATCTGGAAAAGATGAGGAGCCTCCATCAGAAAGATTGCCAAT
GCTTGCAGGCGGGATAGTAGCTCTGGGGAAGTTTGATGCACTCCATATTGGTCATCGTGAGCTTGCAA
TCCAAGCAGCTAAGAGAGGAATTCCATTTCTTCTTTCATTTGTTGGAATGGCGGAAGTACTTGGATGG
GAACCGAGGGCTCCCATTGTTGCTGACTGTGATCGCAAAAGGATTCTTTCATCTTGGGCTCCTTATTG
CGGTAATGTAATCCCAAGGGAATTCCAGATAGAATTTTCCAAAGTTCGATCTCTTACGCCTCGTCAGT
TTGTGGAGAAGCTGTCCAAAGAGCTGGGAGTGCGAGGAGTTGTTGCTGGCGAAAACTATCGTTTTGGA
TATAGAGCTTCTGGTGATGCATCGGACCTTGTGAAGCTCTGTGAAGAGTATGGATTAGAGGCTTATAT
AATCAATTCTGTCATGGACAAGAATCAAATTTCTCGAGCCTTAAACTCTCACGGAAAAAAGGAGAGAG
GGCAAGTATCATCTACTCGTGTTAGGTATGCCCTTGACAAGGGAGACATGAAATATGTGTCAGAGCTA
TTAGGTCGCAACCATCGTCTTGTTTTAACGATGGAGGACCAAGAAAGGATTATTAGTGAGAGAAATAG
GCTGTCAGCTCCCAAGTCTTGTTTGTTGAATCTTGCACCCAAGGAAGGTCTTTATGAGAATTGTTCAG
TTTCGATTGACGAGAATGTTGTACCCTGCAGAGTCATAGTTGATACAACTGGTATTCATTTGGAATGC
TATGAGGTAGCTAATTTTACTTGTATTACTTCTCGAGACGTAAAGATACTAGGTATTGACTTTGGTAG
TCAATAG

Figure 5

Nitab4.5_0005997g0050.2 Amino acid sequence (SEQ ID NO: 3)

MLGGSRCISQHLRDSNFHQFCTSTTPIFTAASPPISTSSFNYFRFRLRNKVSDFGVNCVQPNSSNTQS
SLLSQTLSQSGKDEEPPSERLPMLAGGIVALGKFDALHIGHRELAIQAAKRGIPFLLSFVGMAEVLGW
EPRAPIVADCDRKRILSSWAPYCGNVIPREFQIEFSKVRSLTPRQFVEKLSKELGVRGVVAGENYRFG
YRASGDASDLVKLCEEYGLEAYIINSVMDKNQISRALNSHGKKERGQVSSTRVRYALDKGDMKYVSEL
LGRNHRLVLTMEDQERIISERNRLSAPKSCLLNLAPKEGLYENCSVSIDENVVPCRVIVDTTGIHLEC
YEVANFTCITSRDVKILGIDFGSQ

Figure 6
SEQ ID NO: 34 (TRV1 RNA1)

GATGGTAAGAAGAAAGGCGCGCAGTATGCGATAGCTCTTCACAGCCTGTATGACTTCAAGTTGAAAGA
CTTGATGGCTACTATGGTTGAGAAGAAAACTAAAGTGGTTCATGCTGCTATGCTTTTTGCTCCTGAAA
GTATGTTAGTGGACGAAGGTCCATTACCTTCTGTTGACGGTTACTACATGAAGAAGAACGGGAAGATC
TATTTCGGTTTTGAGAAAGATCCTTCCTTTTCTTACATTCATGACTGGGAAGAGTACAAGAAGTATCT
ACTGGGGAAGCCAGTGAGTTACCAAGGGAATGTGTTCTACTTCGAACCGTGGCAGGTGAGAGGAGACA
CAATGCTTTTTTCGATCTACAGGATAGCTGGAGTTCCGAGGAGGTCTCTATCATCGCAAGAGTACTAC
CGAAGAATATATATCAGTAGATGGGAAAGCATGGTTGTTGTCCCAATTTTCGATCTGGTCGAATCAAC
GCGAGAGTTGGTCAAGAAAGACCTGTTTGTAGAGAAACAATTCATGGACAAGTGTTTGGATTACATAG
CTAGGTTATCTGACCAGCAGCTGACCATAAGCAATGTTAAATCATACTTGAGTTCAAATAATTGGGTC
TTATTCATAAACGGGGCGGCCGTGAAGAACAAGCAAAGTGTAGATTCTCGAGATTTACAGTTGTTGGC
TCAAACTTTGCTAGTGAAGGAACAAGTGGCGAGACCTGTCATGAGGGAGTTGCGTGAAGCAATTCTGA
CTGAGACGAAACCTATCACGTCATTGACTGATGTGCTGGGTTTAATATCAAGAAAACTGTGGAAGCAG
TTTGCTAACAAGATCGCAGTCGGCGGATTCGTTGGCATGGTTGGTACTCTAATTGGATTCTATCCAAA
GAAGGTACTAACCTGGGCGAAGGACACACCAAATGGTCCAGAACTATGTTACGAGAACTCGCACAAAA
CCAAGGTGATAGTATTTCTGAGTGTTGTGTATGCCATTGGAGGAATCACGCTTATGCGTCGAGACATC
CGAGATGGACTGGTGAAAAAACTATGTGATATGTTTGATATCAAACGGGGGGCCCATGTCTTAGACGT
TGAGAATCCGTGCCGCTATTATGAAATCAACGATTTCTTTAGCAGTCTGTATTCGGCATCTGAGTCCG
GTGAGACCGTTTTACCAGATTTATCCGAGGTAAAAGCCAAGTCTGATAAGCTATTGCAGCAGAAGAAA
GAAATCGCTGACGAGTTTCTAAGTGCAAAATTCTCTAACTATTCTGGCAGTTCGGTGAGAACTTCTCC
ACCATCGGTGGTCGGTTCATCTCGAAGCGGACTGGGTCTGTTGTTGGAAGACAGTAACGTGCTGACCC

AAGCTAGAGTTGGAGTTTCAAGAAAGGTAGACGATGAGGAGATCATGGAGCAGTTTCTGAGTGGTCTT
ATTGACACTGAAGCAGAAATTGACGAGGTTGTTTCAGCCTTTTCAGCTGAATGTGAAAGAGGGGAAAC
AAGCGGTACAAAGGTGTTGTGTAAACCTTTAACGCCACCAGGATTTGAGAACGTGTTGCCAGCTGTCA
AACCTTTGGTCAGCAAAGGAAAAACGGTCAAACGTGTCGATTACTTCCAAGTGATGGGAGGTGAGAGA
TTACCAAAAAGGCCGGTTGTCAGTGGAGACGATTCTGTGGACGCTAGAAGAGAGTTTCTGTACTACTT
AGATGCGGAGAGAGTCGCTCAAAATGATGAAATTATGTCTCTGTATCGTGACTATTCGAGAGGAGTTA
TTCGAACTGGAGGTCAGAATTACCCGCACGGACTGGGAGTGTGGGATGTGGAGATGAAGAACTGGTGC
ATACGTCCAGTGGTCACTGAACATGCTTATGTGTTCCAACCAGACAAACGTATGGATGATTGGTCGGG
ATACTTAGAAGTGGCTGTTTGGGAACGAGGTATGTTGGTCAACGACTTCGCGGTCGAAAGGATGAGTG
ATTATGTCATAGTTTGCGATCAGACGTATCTTTGCAATAACAGGTAATAATCCTCTCTCTTGATATTT
TTAAATTATAGAATTAATTAGTTTACTTTATTCTTTACTATATGATTTAAATAGTTTAATCTTGTTTT
TGAGTAAACTATTCGATTTTGATATTTGTATTCGTCCTACAAAGTTGGAAATACTGATGATATTTTCT
TTTGAACGTGATACCTACCAATACTAATCTTACGGAATCTTTTAATAGAGCACTAATCAACATGGAAC
TAAAGACCAATTCTTAAGTGTCTCTGTTGTACAGTTCATTTTAGTAGTGCGTTTAAGTATTATTATCT
CCCTTCATGCGGGGCAATTATGTAGATTAAAATCGAAATTATATAAAATTTACATAAGTCTAAGTCTA
GGGTCTCCAGCTAATTGTTATTTTTTTAACGATGTTGACTAAAGCAATAACGACGTTGACTTGTGTTA
AACAGGTTGATCTTGGACAATTTAAGTGCCCTGGATCTAGGACCAGTTAACTGTTCTTTTGAATTAGT
TGACGGTGTACCTGGTTGTGGTAAGTCGACAATGATTGTCAACTCAGCTAATCCTTGTGTCGATGTGG
TTCTCTCTACTGGGAGAGCAGCAACCGACGACTTGATCGAGAGATTCGCGAGCAAAGGTTTTCCATGC
AAATTGAAAAGGAGAGTGAAGACGGTTGATTCTTTTTTGATGCATTGTGTCGATGGTTCTTTAACCGG
AGACGTGTTGCATTTCGACGAAGCTCTCATGGCCCATGCTGGTATGGTGTACTTTTGCGCTCAGATAG
CTGGTGCTAAACGATGTATCTGTCAAGGAGATCAGAATCAAATTTCTTTCAAGCCTAGGGTATCTCAA
GTTGATTTGAGGTTTTCTAGTCTGGTCGGAAAGTTTGACATTGTTACAGAAAAAAGAGAAACTTACAG
AAGTCCAGCAGATGTGGCTGCCGTATTGAACAAGTACTATACTGGAGATGTCAGAACACATAACGCGA
CTGCTAATTCGATGACGGTGAGGAAGATTGTGTCTAAAGAACAGGTTTCTTTGAAGCCTGGTGCTCAG
TACATAACTTTCCTTCAGTCTGAGAAGAAGGAGTTGGTAAATTTGTTGGCATTGAGGAAAGTGGCAGC
TAAAGTGAGTACAGTACACGAGTCGCAAGGAGAGACATTCAAAGATGTAGTCCTAGTCAGGACGAAAC
CTACGGATGACTCAATCGCTAGAGGTCGGGAGTACTTAATCGTGGCATTGTCGCGTCACACACAATCA
CTTGTGTATGAAACTGTGAAAGAGGACGATGTAAGCAAAGAGATCAGGGAAAGTGCCGCGCTTACGAA
GGCGGCTTTGGCAAGATTTTTTGTTACTGAGACCGTCTTATGACGGTTTCGGTCTAGGTTTGATGTCT
TTAGACATCATGAAGGGCCTTGCGCCGTTCCAGATTCAGGTACGATTACGGACTTGGAGATGTGGTAC
GACGCTTTGTTTCCGGGAAATTCGTTAAGAGACTCAAGCCTAGACGGGTATTTGGTGGCAACGACTGA
TTGCAATTTGCGATTAGACAATGTTACGATCAAAAGTGGAAACTGGAAAGACAAGTTTGCTGAAAAAG
AAACGTTTCTGAAACCGGTTATTCGTACTGCTATGCCTGACAAAAGGAAGACTACTCAGTTGGAGAGT
TTGTTAGCATTGCAGAAAAGGAACCAAGCGGCACCCGATCTACAAGAAAATGTGCACGCGACAGTTCT
AATCGAAGAGACGATGAAGAAGCTGAAATCTGTTGTCTACGATGTGGGAAAAATTCGGGCTGATCCTA
TTGTCAATAGAGCTCAAATGGAGAGATGGTGGAGAAATCAAAGCACAGCGGTACAGGCTAAGGTAGTA
GCAGATGTGAGAGAGTTACATGAAATAGACTATTCGTCTTACATGTATATGATCAAATCTGACGTGAA
ACCTAAGACTGATTTAACACCGCAATTTGAATACTCAGCTCTACAGACTGTTGTGTATCACGAGAAGT
TGATCAACTCGTTGTTCGGTCCAATTTTCAAAGAAATTAATGAACGCAAGTTGGATGCTATGCAACCA
CATTTTGTGTTCAACACGAGAATGACATCGAGTGATTTAAACGATCGAGTGAAGTTCTTAAATACGGA
AGCGGCTTACGACTTTGTTGAGATAGACATGTCTAAATTCGACAAGTCGGCAAATCGCTTCCATTTAC
AACTGCAGCTGGAGATTTACAGGTTATTTGGGCTGGATGAGTGGGCGGCCTTCCTTTGGGAGGTGTCG
CACACTCAAACTACTGTGAGAGATATTCAAAATGGTATGATGGCGCATATTTGGTACCAACAAAAGAG
TGGAGATGCTGATACTTATAATGCAAATTCAGATAGAACACTGTGTGCACTCTTGTCTGAATTACCAT
TGGAGAAAGCAGTCATGGTTACATATGGAGGAGATGACTCACTGATTGCGTTTCCTAGAGGAACGCAG
TTTGTTGATCCGTGTCCAAAGTTGGCTACTAAGTGGAATTTCGAGTGCAAGATTTTTAAGTACGATGT
CCCAATGTTTTGTGGGAAGTTCTTGCTTAAGACGTCATCGTGTTACGAGTTCGTGCCAGATCCGGTAA
AAGTTCTGACGAAGTTGGGGAAAAAGAGTATAAAGGATGTGCAACATTTAGCCGAGATCTACATCTCG
CTGAATGATTCCAATAGAGCTCTTGGGAACTACATGGTGGTATCCAAACTGTCCGAGTCTGTTTCAGA
CCGGTATTTGTACAAAGGTGATTCTGTTCATGCGCTTTGTGCGCTATGGAAGCATATTAAGAGTTTTA
CAGCTCTGTGTACATTATTCCGAGACGAAAACGATAAGGAATTGAACCCGGCTAAGGTTGATTGGAAG
AAGGCACAGAGAGCTGTGTCAAACTTTTACGACTGGTAATATGGAAGACAAGTCATTGGTCACCTTGA
AGAAGAAGACTTTCGAAGTCTCAAAATTCTCAAATCTAGGGGCCATTGAATTGTTTGTGGACGGTAGG
AGGAAGAGACCGAAGTATTTTCACAGAAGAAGAGAAACTGTCCTAAATCATGTTGGTGGGAAGAAGAG
TGAACACAAGTTAGACGTTTTTGACCAAAGGGATTACAAAATGATTAAATCTTACGCGTTTCTAAAGG

TAGTAGGTGTACAACTAGTTGTAACATCACATCTACCTGCAGATACGCCTGGGTTCATTCAAATCGAT
CTGTTGGATTCGAGACTTACTGAGAAAAGAAAGAGAGGAAAGACTATTCAGAGATTCAAAGCTCGAGC
TTGCGATAACTGTTCAGTTGCGCAGTACAAGGTTGAATACAGTATTTCCACACAGGAGAACGTACTTG
ATGTCTGGAAGGTGGGTTGTATTTCTGAGGGCGTTCCGGTCTGTGACGGTACATACCCTTTCAGTATC
GAAGTGTCGCTAATATGGGTTGCTACTGATTCGACTAGGCGCCTCAATGTGGAAGAACTGAACAGTTC
GGATTACATTGAAGGCGATTTTACCGATCAAGAGGTTTTCGGTGAGTTCATGTCTTTGAAACAAGTGG
AGATGAAGACGATTGAGGCGAAGTACGATGGTCCTTACAGACCAGCTACTACTAGACCTAAGTCATTA
TTGTCAAGTGAAGATGTTAAGAGAGCGTCTAATAAGAAAAACTCGTCTTAATGCATAAAGAAATTTAT
TGTCAATATGACGTGTGTACTCAAGGGTTGTGTGAATGAAGTCACTGTTCTTGGTCACGAGACGTGTA
GTATCGGTCATGCTAACAAATTGCGAAAGCAAGTTGCTGACATGGTTGGTGTCACACGTAGGTGTGCG
GAAAATAATTGTGGATGGTTTGTCTGTGTTGTTATCAATGATTTTACTTTTGATGTGTATAATTGTTG
TGGCCGTAGTCACCTTGAAAAGTGTCGTAAACGTGTTGAAACAAGAAATCGAGAAATTTGGAAACAAA
TTCGACGAAATCAAGCTGAAAACATGTCTGCGACAGCTAAAAAGTCTCATAATTCGAAGACCTCTAAG
AAGAAATTCAAAGAGGACAGAGAATTTGGGACACCAAAAAGATTTTTAAGAGATGATGTTCCTTTCGG
GATTGATCGTTTGTTTGCTTTTTGATTTTATTTTATATTGTTATCTGTTTCTGTGTATAGACTGTTTG
AGATTGGCGCTTGGCCGACTCATTGTCTTACCATAGGGGAACGGACTTTGTTTGTGTTGTTATTTTAT
TTGTATTTTATTAAAATTCTCAATGATCTGAAAAGGCCTCGAGGCTAAGAGATTATTGGGGGGTGAGT
AAGTACTTTTAAAGTGATGATGGTTACAAAGGCAAAAGGGGTAAAACCCCTCGCCTACGTAAGCGTTA
TTACGCCCGGATCCCCCGGGGAGCTCGAATTCGCTGAAATCACCAGTCTCTCTCTACAAATCTATCTC
TCTCTATTTTTTCCATAAATAATGTGTGAGTAGTTTCCCGATAAGGGAAATTAGGGTTCTTATAGGGT
TTCGCTCATGTGTTGAGCATATAAGAAACCCTTAGTATGTATTTGTATTTGTAAAATACTTCTATTAT
CAATAAAATTTCTAATTCCTAAAACCAAAATCCAGTACTAAAATCCAGATCTCCTAAAGTCCCTATAG
ATCTTTGTCGTGAATATAAACCAGACACGAGACGACTAAACCTGGAGCCCAGACGCCGTTCGAAGCTA
GAAGTACCGCTTAGGCAGGAGGCCGTTAGGGAAAAGATGCTAAGGCAGGGTTGGTTACGTTGACTCCC
CCGTAGGTTTGGTTTAAATATGATGAAGTGGACGGAAGGAAGGAGGAAGACAAGGAAGGATAAGGTTG
CAGGCCCTGTGCAAGGTAAGAAGATGGAAATTTGATAGAGGTACGCTACTATACTTATACTATACGCT
AAGGGAATGCTTGTATTTATACCCTATACCCCCTAATAACCCCTTATCAATTTAAGAAATAATCCGCA
TAAGCCCCCGCTTAAAAATTGGTATCAGAGCCATGAATAGGTCTATGACCAAAACTCAAGAGGATAAA
ACCTCACCAAAATACGAAAGAGTTCTTAACTCTAAAGATAAAAGATCTTTCAAGATCAAAACTAGTTC
CCTCACACCGGAGCATGCGATATCCTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGC
TGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGT
AAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCA
GTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTA
TTGGGCCAAAGACAAAAGGGCGACATTCAACCGATTGAGGGAGGGAAGGTAAATATTGACGGAAATTA
TTCATTAAAGGTGAATTATCACCGTCACCGACTTGAGCCATTTGGGAATTAGAGCCAGCAAAATCACC
AGTAGCACCATTACCATTAGCAAGGCCGGAAACGTCACCAATGAAACCATCGATAGCAGCACCGTAAT
CAGTAGCGACAGAATCAAGTTTGCCTTTAGCGTCAGACTGTAGCGCGTTTTCATCGGCATTTTCGGTC
ATAGCCCCCTTATTAGCGTTTGCCATCTTTTCATAATCAAAATCACCGGAACCAGAGCCACCACCGGA
ACCGCCTCCCTCAGAGCCGCCACCCTCAGAACCGCCACCCTCAGAGCCACCACCCTCAGAGCCGCCAC
CAGAACCACCACCAGAGCCGCCGCCAGCATTGACAGGAGGCCCGATCTAGTAACATAGATGACACCGC
GCGCGATAATTTATCCTAGTTTGCGCGCTATATTTTGTTTTCTATCGCGTATTAAATGTATAATTGCG
GGACTCTAATCATAAAAACCCATCTCATAAATAACGTCATGCATTACATGTTAATTATTACATGCTTA
ACGTAATTCAACAGAAATTATATGATAATCATCGCAAGACCGGCAACAGGATTCAATCTTAAGAAACT
TTATTGCCAAATGTTTGAACGATCGGGGATCATCCGGGTCTGTGGCGGGAACTCCACGAAAATATCCG
AACGCAGCAAGATATCGCGGTGCATCTCGGTCTTGCCTGGGCAGTCGCCGCCGACGCCGTTGATGTGG
ACGCCGGGCCCGATCATATTGTCGCTCAGGATCGTGGCGTTGTGCTTGTCGGCCGTTGCTGTCGTAAT
GATATCGGCACCTTCGACCGCCTGTTCCGCAGAGATCCCGTGGGCGAAGAACTCCAGCATGAGATCCC
CGCGCTGGAGGATCATCCAGCCGGCGTCCCGGAAAACGATTCCGAAGCCCAACCTTTCATAGAAGGCG
GCGGTGGAATCGAAATCTCGTGATGGCAGGTTGGGCGTCGCTTGGTCGGTCATTTCGAACCCCAGAGT
CCCGCTCAGAAGAACTCGTCAAGAAGGCGATAGAAGGCGATGCGCTGCGAATCGGGAGCGGCGATACC
GTAAAGCACGAGGAAGCGGTCAGCCCATTCGCCGCCAAGCTCTTCAGCAATATCACGGGTAGCCAACG
CTATGTCCTGATAGCGGTCCGCCACACCCAGCCGGCCACAGTCGATGAATCCAGAAAAGCGGCCATTT
TCCACCATGATATTCGGCAAGCAGGCATCGCCATGGGTCACGACGAGATCATCGCCGTCGGGCATGCG
CGCCTTGAGCCTGGCGAACAGTTCGGCTGGCGCGAGCCCCTGATGCTCTTCGTCCAGATCATCCTGAT
CGACAAGACCGGCTTCCATCCGAGTACGTGCTCGCTCGATGCGATGTTTCGCTTGGTGGTCGAATGGG
CAGGTAGCCGGATCAAGCGTATGCAGCCGCCGCATTGCATCAGCCATGATGGATACTTTCTCGGCAGG

AGCAAGGTGAGATGACAGGAGATCCTGCCCCGGCACTTCGCCCAATAGCAGCCAGTCCCTTCCCGCTT
CAGTGACAACGTCGAGCACAGCTGCGCAAGGAACGCCCGTCGTGGCCAGCCACGATAGCCGCGCTGCC
TCGTCCTGCAGTTCATTCAGGGCACCGGACAGGTCGGTCTTGACAAAAAGAACCGGGCGCCCCTGCGC
TGACAGCCGGAACACGGCGGCATCAGAGCAGCCGATTGTCTGTTGTGCCCAGTCATAGCCGAATAGCC
TCTCCACCCAAGCGGCCGGAGAACCTGCGTGCAATCCATCTTGTTCAATCATGCGAAACGATCCAGAT
CCGGTGCAGATTATTTGGATTGAGAGTGAATATGAGACTCTAATTGGATACCGAGGGGAATTTATGGA
ACGTCAGTGGAGCATTTTTGACAAGAAATATTTGCTAGCTGATAGTGACCTTAGGCGACTTTTGAACG
CGCAATAATGGTTTCTGACGTATGTGCTTAGCTCATTAAACTCCAGAAACCCGCGGCTGAGTGGCTCC
TTCAACGTTGCGGTTCTGTCAGTTCCAAACGTAAAACGGCTTGTCCCGCGTCATCGGCGGGGGTCATA
ACGTGACTCCCTTAATTCTCCGCTCATGATCAGATTGTCGTTTCCCGCCTTCAGTTTAAACTATCAGT
GTTTGACAGGATATATTGGCGGGTAAACCTAAGAGAAAAGAGCGTTTATTAGAATAATCGGATATTTA
AAAGGGCGTGAAAAGGTTTATCCGTTCGTCCATTTGTATGTGCATGCCAACCACAGGGTTCCCCAGAT
CTGGCGCCGGCCAGCGAGACGAGCAAGATTGGCCGCCGCCCGAAACGATCCGACAGCGCGCCCAGCAC
AGGTGCGCAGGCAAATTGCACCAACGCATACAGCGCCAGCAGAATGCCATAGTGGGCGGTGACGTCGT
TCGAGTGAACCAGATCGCGCAGGAGGCCCGGCAGCACCGGCATAATCAGGCCGATGCCGACAGCGTCG
AGCGCGACAGTGCTCAGAATTACGATCAGGGGTATGTTGGGTTTCACGTCTGGCCTCCGGACCAGCCT
CCGCTGGTCCGATTGAACGCGCGGATTCTTTATCACTGATAAGTTGGTGGACATATTATGTTTATCAG
TGATAAAGTGTCAAGCATGACAAAGTTGCAGCCGAATACAGTGATCCGTGCCGCCCTGGACCTGTTGA
ACGAGGTCGGCGTAGACGGTCTGACGACACGCAAACTGGCGGAACGGTTGGGGGTTCAGCAGCCGGCG
CTTTACTGGCACTTCAGGAACAAGCGGGCGCTGCTCGACGCACTGGCCGAAGCCATGCTGGCGGAGAA
TCATACGCATTCGGTGCCGAGAGCCGACGACGACTGGCGCTCATTTCTGATCGGGAATGCCCGCAGCT
TCAGGCAGGCGCTGCTCGCCTACCGCGATGGCGCGCGCATCCATGCCGGCACGCGACCGGGCGCACCG
CAGATGGAAACGGCCGACGCGCAGCTTCGCTTCCTCTGCGAGGCGGGTTTTTCGGCCGGGGACGCCGT
CAATGCGCTGATGACAATCAGCTACTTCACTGTTGGGGCCGTGCTTGAGGAGCAGGCCGGCGACAGCG
ATGCCGGCGAGCGCGGCGGCACCGTTGAACAGGCTCCGCTCTCGCCGCTGTTGCGGGCCGCGATAGAC
GCCTTCGACGAAGCCGGTCCGGACGCAGCGTTCGAGCAGGGACTCGCGGTGATTGTCGATGGATTGGC
GAAAAGGAGGCTCGTTGTCAGGAACGTTGAAGGACCGAGAAAGGGTGACGATTGATCAGGACCGCTGC
CGGAGCGCAACCCACTCACTACAGCAGAGCCATGTAGACAACATCCCCTCCCCCTTTCCACCGCGTCA
GACGCCCGTAGCAGCCCGCTACGGGCTTTTTCATGCCCTGCCCTAGCGTCCAAGCCTCACGGCCGCGC
TCGGCCTCTCTGGCGGCCTTCTGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGT
TCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATA
ACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTG
GCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCG
AAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTC
CGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTTCCGCTGC
ATAACCCTGCTTCGGGGTCATTATAGCGATTTTTTCGGTATATCCATCCTTTTTCGCACGATATACAG
GATTTTGCCAAAGGGTTCGTGTAGACTTTCCTTGGTGTATCCAACGGCGTCAGCCGGGCAGGATAGGT
GAAGTAGGCCCACCCGCGAGCGGGTGTTCCTTCTTCACTGTCCCTTATTCGCACCTGGCGGTGCTCAA
CGGGAATCCTGCTCTGCGAGGCTGGCCGGCTACCGCCGGCGTAACAGATGAGGGCAAGCGGATGGCTG
ATGAAACCAAGCCAACCAGGAAGGGCAGCCCACCTATCAAGGTGTACTGCCTTCCAGACGAACGAAGA
GCGATTGAGGAAAAGGCGGCGGCGGCCGGCATGAGCCTGTCGGCCTACCTGCTGGCCGTCGGCCAGGG
CTACAAAATCACGGGCGTCGTGGACTATGAGCACGTCCGCGAGCTGGCCCGCATCAATGGCGACCTGG
GCCGCCTGGGCGGCCTGCTGAAACTCTGGCTCACCGACGACCCGCGCACGGCGCGGTTCGGTGATGCC
ACGATCCTCGCCCTGCTGGCGAAGATCGAAGAGAAGCAGGACGAGCTTGGCAAGGTCATGATGGGCGT
GGTCCGCCCGAGGGCAGAGCCATGACTTTTTTAGCCGCTAAAACGGCCGGGGGGTGCGCGTGATTGCC
AAGCACGTCCCCATGCGCTCCATCAAGAAGAGCGACTTCGCGGAGCTGGTGAAGTACATCACCGACGA
GCAAGGCAAGACCGAGCGCCTTTGCGACGCTCACCGGGCTGGTTGCCCTCGCCGCTGGGCTGGCGGCC
GTCTATGGCCCTGCAAACGCGCCAGAAACGCCGTCGAAGCCGTGTGCGAGACACCGCGGCCGCCGGCG
TTGTGGATACCTCGCGGAAAACTTGGCCCTCACTGACAGATGAGGGGCGGACGTTGACACTTGAGGGG
CCGACTCACCCGGCGCGGCGTTGACAGATGAGGGGCAGGCTCGATTTCGGCCGGCGACGTGGAGCTGG
CCAGCCTCGCAAATCGGCGAAAACGCCTGATTTTACGCGAGTTTCCCACAGATGATGTGGACAAGCCT
GGGGATAAGTGCCCTGCGGTATTGACACTTGAGGGGCGCGACTACTGACAGATGAGGGGCGCGATCCT
TGACACTTGAGGGGCAGAGTGCTGACAGATGAGGGGCGCACCTATTGACATTTGAGGGGCTGTCCACA
GGCAGAAAATCCAGCATTTGCAAGGGTTTCCGCCCGTTTTTCGGCCACCGCTAACCTGTCTTTTAACC
TGCTTTTAAACCAATATTTATAAACCTTGTTTTTAACCAGGGCTGCGCCCTGTGCGCGTGACCGCGCA
CGCCGAAGGGGGGTGCCCCCCCTTCTCGAACCCTCCCGGCCCGCTAACGCGGGCCTCCCATCCCCCCA

GGGGCTGCGCCCCTCGGCCGCGAACGGCCTCACCCCAAAAATGGCAGCGCTGGCAGTCCTTGCCATTG
CCGGGATCGGGGCAGTAACGGGATGGGCGATCAGCCCGAGCGCGACGCCCGGAAGCATTGACGTGCCG
CAGGTGCTGGCATCGACATTCAGCGACCAGGTGCCGGGCAGTGAGGGCGGCGGCCTGGGTGGCGGCCT
GCCCTTCACTTCGGCCGTCGGGGCATTCACGGACTTCATGGCGGGGCCGGCAATTTTTACCTTGGGCA
TTCTTGGCATAGTGGTCGCGGGTGCCGTGCTCGTGTTCGGGGGTGCGATAAACCCAGCGAACCATTTG
AGGTGATAGGTAAGATTATACCGAGGTATGAAAACGAGAATTGGACCTTTACAGAATTACTCTATGAA
GCGCCATATTTAAAAAGCTACCAAGACGAAGAGGATGAAGAGGATGAGGAGGCAGATTGCCTTGAATA
TATTGACAATACTGATAAGATAATATATCTTTTATATAGAAGATATCGCCGTATGTAAGGATTTCAGG
GGGCAAGGCATAGGCAGCGCGCTTATCAATATATCTATAGAATGGGCAAAGCATAAAAACTTGCATGG
ACTAATGCTTGAAACCCAGGACAATAACCTTATAGCTTGTAAATTCTATCATAATTGGGTAATGACTC
CAACTTATTGATAGTGTTTTATGTTCAGATAATGCCCGATGACTTTGTCATGCAGCTCCACCGATTTT
GAGAACGACAGCGACTTCCGTCCCAGCCGTGCCAGGTGCTGCCTCAGATTCAGGTTATGCCGCTCAAT
TCGCTGCGTATATCGCTTGCTGATTACGTGCAGCTTTCCCTTCAGGCGGGATTCATACAGCGGCCAGC
CATCCGTCATCCATATCACCACGTCAAAGGGTGACAGCAGGCTCATAAGACGCCCCAGCGTCGCCATA
GTGCGTTCACCGAATACGTGCGCAACAACCGTCTTCCGGAGACTGTCATACGCGTAAAACAGCCAGCG
CTGGCGCGATTTAGCCCCGACATAGCCCCACTGTTCGTCCATTTCCGCGCAGACGATGACGTCACTGC
CCGGCTGTATGCGCGAGGTTACCGACTGCGGCCTGAGTTTTTTAAGTGACGTAAAATCGTGTTGAGGC
CAACGCCCATAATGCGGGCTGTTGCCCGGCATCCAACGCCATTCATGGCCATATCAATGATTTTCTGG
TGCGTACCGGGTTGAGAAGCGGTGTAAGTGAACTGCAGTTGCCATGTTTTACGGCAGTGAGAGCAGAG
ATAGCGCTGATGTCCGGCGGTGCTTTTGCCGTTACGCACCACCCCGTCAGTAGCTGAACAGGAGGGAC
AGCTGATAGACACAGAAGCCACTGGAGCACCTCAAAAACACCATCATACACTAAATCAGTAAGTTGGC
AGCATCACCCATAATTGTGGTTTCAAAATCGGCTCCGTCGATACTATGTTATACGCCAACTTTGAAAA
CAACTTTGAAAAAGCTGTTTTCTGGTATTTAAGGTTTTAGAATGCAAGGAACAGTGAATTGGAGTTCG
TCTTGTTATAATTAGCTTCTTGGGGTATCTTTAAATACTGTAGAAAAGAGGAAGGAAATAATAAATGG
CTAAAATGAGAATATCACCGGAATTGAAAAAACTGATCGAAAAATACCGCTGCGTAAAAGATACGGAA
GGAATGTCTCCTGCTAAGGTATATAAGCTGGTGGGAGAAAATGAAAACCTATATTTAAAAATGACGGA
CAGCCGGTATAAAGGGACCACCTATGATGTGGAACGGGAAAAGGACATGATGCTATGGCTGGAAGGAA
AGCTGCCTGTTCCAAAGGTCCTGCACTTTGAACGGCATGATGGCTGGAGCAATCTGCTCATGAGTGAG
GCCGATGGCGTCCTTTGCTCGGAAGAGTATGAAGATGAACAAAGCCCTGAAAAGATTATCGAGCTGTA
TGCGGAGTGCATCAGGCTCTTTCACTCCATCGACATATCGGATTGTCCCTATACGAATAGCTTAGACA
GCCGCTTAGCCGAATTGGATTACTTACTGAATAACGATCTGGCCGATGTGGATTGCGAAAACTGGGAA
GAAGACACTCCATTTAAAGATCCGCGCGAGCTGTATGATTTTTTAAAGACGGAAAAGCCCGAAGAGGA
ACTTGTCTTTTCCCACGGCGACCTGGGAGACAGCAACATCTTTGTGAAAGATGGCAAAGTAAGTGGCT
TTATTGATCTTGGGAGAAGCGGCAGGGCGGACAAGTGGTATGACATTGCCTTCTGCGTCCGGTCGATC
AGGGAGGATATCGGGGAAGAACAGTATGTCGAGCTATTTTTTGACTTACTGGGGATCAAGCCTGATTG
GGAGAAAATAAAATATTATATTTTACTGGATGAATTGTTTTAGTACCTAGATGTGGCGCAACGATGCC
GGCGACAAGCAGGAGCGCACCGACTTCTTCCGCATCAAGTGTTTTGGCTCTCAGGCCGAGGCCCACGG
CAAGTATTTGGGCAAGGGGTCGCTGGTATTCGTGCAGGGCAAGATTCGGAATACCAAGTACGAGAAGG
ACGGCCAGACGGTCTACGGGACCGACTTCATTGCCGATAAGGTGGATTATCTGGACACCAAGGCACCA
GGCGGGTCAAATCAGGAATAAGGGCACATTGCCCCGGCGTGAGTCGGGGCAATCCCGCAAGGAGGGTG
AATGAATCGGACGTTTGACCGGAAGGCATACAGGCAAGAACTGATCGACGCGGGGTTTTCCGCCGAGG
ATGCCGAAACCATCGCAAGCCGCACCGTCATGCGTGCGCCCCGCGAAACCTTCCAGTCCGTCGGCTCG
ATGGTCCAGCAAGCTACGGCCAAGATCGAGCGCGACAGCGTGCAACTGGCTCCCCCTGCCCTGCCCGC
GCCATCGGCCGCCGTGGAGCGTTCGCGTCGTCTCGAACAGGAGGCGGCAGGTTTGGCGAAGTCGATGA
CCATCGACACGCGAGGAACTATGACGACCAAGAAGCGAAAAACCGCCGGCGAGGACCTGGCAAAACAG
GTCAGCGAGGCCAAGCAGGCCGCGTTGCTGAAACACACGAAGCAGCAGATCAAGGAAATGCAGCTTTC
CTTGTTCGATATTGCGCCGTGGCCGGACACGATGCGAGCGATGCCAAACGACACGGCCCGCTCTGCCC
TGTTCACCACGCGCAACAAGAAAATCCCGCGCGAGGCGCTGCAAAACAAGGTCATTTTCCACGTCAAC
AAGGACGTGAAGATCACCTACACCGGCGTCGAGCTGCGGGCCGACGATGACGAACTGGTGTGGCAGCA
GGTGTTGGAGTACGCGAAGCGCACCCCTATCGGCGAGCCGATCACCTTCACGTTCTACGAGCTTTGCC
AGGACCTGGGCTGGTCGATCAATGGCCGGTATTACACGAAGGCCGAGGAATGCCTGTCGCGCCTACAG
GCGACGGCGATGGGCTTCACGTCCGACCGCGTTGGGCACCTGGAATCGGTGTCGCTGCTGCACCGCTT
CCGCGTCCTGGACCGTGGCAAGAAAACGTCCCGTTGCCAGGTCCTGATCGACGAGGAAATCGTCGTGC
TGTTTGCTGGCGACCACTACACGAAATTCATATGGGAGAAGTACCGCAAGCTGTCGCCGACGGCCCGA
CGGATGTTCGACTATTTCAGCTCGCACCGGGAGCCGTACCCGCTCAAGCTGGAAACCTTCCGCCTCAT
GTGCGGATCGGATTCCACCCGCGTGAAGAAGTGGCGCGAGCAGGTCGGCGAAGCCTGCGAAGAGTTGC

GAGGCAGCGGCCTGGTGGAACACGCCTGGGTCAATGATGACCTGGTGCATTGCAAACGCTAGGGCCTT
GTGGGGTCAGTTCCGGCTGGGGGTTCAGCAGCCAGCGCTTTACTGGCATTTCAGGAACAAGCGGGCAC
TGCTCGACGCACTTGCTTCGCTCAGTATCGCTCGGGACGCACGGCGCGCTCTACGAACTGCCGATAAA
CAGAGGATTAAAATTGACAATTGTGATTAAGGCTCAGATTCGACGGCTTGGAGCGGCCGACGTGCAGG
ATTTCCGCGAGATCCGATTGTCGGCCCTGAAGAAAGCTCCAGAGATGTTCGGGTCCGTTTACGAGCAC
GAGGAGAAAAAGCCCATGGAGGCGTTCGCTGAACGGTTGCGAGATGCCGTGGCATTCGGCGCCTACAT
CGACGGCGAGATCATTGGGCTGTCGGTCTTCAAACAGGAGGACGGCCCCAAGGACGCTCACAAGGCGC
ATCTGTCCGGCGTTTTCGTGGAGCCCGAACAGCGAGGCCGAGGGGTCGCCGGTATGCTGCTGCGGGCG
TTGCCGGCGGGTTTATTGCTCGTGATGATCGTCCGACAGATTCCAACGGGAATCTGGTGGATGCGCAT
CTTCATCCTCGGCGCACTTAATATTTCGCTATTCTGGAGCTTGTTGTTTATTTCGGTCTACCGCCTGC
CGGGCGGGGTCGCGGCGACGGTAGGCGCTGTGCAGCCGCTGATGGTCGTGTTCATCTCTGCCGCTCTG
CTAGGTAGCCCGATACGATTGATGGCGGTCCTGGGGGCTATTTGCGGAACTGCGGGCGTGGCGCTGTT
GGTGTTGACACCAAACGCAGCGCTAGATCCTGTCGGCGTCGCAGCGGGCCTGGCGGGGGCGGTTTCCA
TGGCGTTCGGAACCGTGCTGACCCGCAAGTGGCAACCTCCCGTGCCTCTGCTCACCTTTACCGCCTGG
CAACTGGCGGCCGGAGGACTTCTGCTCGTTCCAGTAGCTTTAGTGTTTGATCCGCCAATCCCGATGCC
TACAGGAACCAATGTTCTCGGCCTGGCGTGGCTCGGCCTGATCGGAGCGGGTTTAACCTACTTCCTTT
GGTTCCGGGGGATCTCGCGACTCGAACCTACAGTTGTTTCCTTACTGGGCTTTCTCAGCCCCAGATCT
GGGGTCGATCAGCCGGGGATGCATCAGGCCGACAGTCGGAACTTCGGGTCCCCGACCTGTACCATTCG
GTGAGCAATGGATAGGGGAGTTGATATCGTCAACGTTCACTTCTAAAGAAATAGCGCCACTCAGCTTC
CTCAGCGGCTTTATCCAGCGATTTCCTATTATGTCGGCATAGTTCTCAAGATCGACAGCCTGTCACGG
TTAAGCGAGAAATGAATAAGAAGGCTGATAATTCGGATCTCTGCGAGGGAGATGATATTTGATCACAG
GCAGCAACGCTCTGTCATCGTTACAATCAACATGCTACCCTCCGCGAGATCATCCGTGTTTCAAACCC
GGCAGCTTAGTTGCCGTTCTTCCGAATAGCATCGGTAACATGAGCAAAGTCTGCCGCCTTACAACGGC
TCTCCCGCTGACGCCGTCCCGGACTGATGGGCTGCCTGTATCGAGTGGTGATTTTGTGCCGAGCTGCC
GGTCGGGGAGCTGTTGGCTGGCTGGTGGCAGGATATATTGTGGTGTAAACAAATTGACGCTTAGACAA
CTTAATAACACATTGCGGACGTTTTTAATGTACTGGGGTGGTTTTTCTTTTCACCAGTGAGACGGGCA
ACAGCTGATTGCCCTTCACCGCCTGGCCCTGAGAGAGTTGCAGCAAGCGGTCCACGCTGGTTTGCCCC
AGCAGGCGAAAATCCTGTTTGATGGTGGTTCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAG
CCCGAGATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAA
CGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCAAATCAAGTT
TTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCCCGATTTAGAGCTTGA
CGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCGCCATTCAGGC
TGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGA
TGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGC
CAGTGAATTCGAGCTCGGTACCCCCCTACTCCAAAAATGTCAAAGATACAGTCTCAGAAGACCAAAGG
GCTATTGAGACTTTTCAACAAAGGGTAATTTCGGGAAACCTCCTCGGATTCCATTGCCCAGCTATCTG
TCACTTCATCGAAAGGACAGTAGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAA
AGGCTATCATTCAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATC
GTGGAAAAAGAAGACGTCCCAACCACGTCTTCAAAGCAAGTGGATTGATGTGACATCTCCACTGACGT
AAGGGATGACGCACAATCCCACTATCCTTCGCAAGACCCTTCTTCTATATAAGGAAGTTCATTTCATT
TGGAGAGGACAGCCCAAGCTTTCTAGAGGATCCATAAAACATTTCAATCCTTTGAACGCGGTAGAACG
TGCTAATTGGATTTTGGTGAGAACGCGGTAGAACGTACTTATCACCTACAGTTTTATTTTGTTTTTCT
TTTTGGTTTAATCTATCCAGCTTAGTACCGAGTGGGGGAAAGTGACTGGTGTGCCTAAAACCTTTTCT
TTGATACTTTGTAAAAATACATACAGATACAATGGCGAACGGTAACTTCAAGTTGTCTCAATTGCTCA
ATGTGGACGAGATGTCTGCTGAGCAGAGGAGTCATTTCTTTGACTTGATGCTGACTAAACCTGATTGT
GAGATCGGGCAAATGATGCAAAGAGTTGTTGTTGATAAAGTCGATGACATGATTAGAGAAAGAAAGAC
TAAAGATCCAGTGATTGTTCATGAAGTTCTTTCTCAGAAGGAACAGAACAAGTTGATGGAAATTTATC
CTGAATTCAATATCGTGTTTAAAGACGACAAAAACATGGTTCATGGGTTTGCGGCTGCTGAGCGAAAA
CTACAAGCTTTATTGCTTTTAGATAGAGTTCCTGCTCTGCAAGAGGTGGATGACATCGGTGGTCAATG
GTCGTTTTGGGTAACTAGAGGTGAGAAAAGGATTCATTCCTGTTGTCCAAATCTAGATATTCGGGATG
ATCAGAGAGAAATTTCTCGACAGATATTTCTTACTGCTATTGGTGATCAAGCTAGAAGTGGTAAGAGA
CAGATGTCGGAGAATGAGCTGTGGATGTATGACCAATTTCGTGAAAATATTGCTGCGCCTAACGCGGT
TAGGTGCAATAATACATATCAGGGTTGTACATGTAGGGGTTTT

Figure 7

Identifier	Nic 1 ERF - Gene
Nitab4.5_0003090g0020.1	*ERF17L3ΔN*
Nitab4.5_0003090g0030.1	*ERF199*
Nitab4.5_0003665g0040.1	*JRE5L2*
Nitab4.5_0004620g0010.1	*ERF210*
Nitab4.5_0004620g0030.1	*ERF91*
Nitab4.5_0004620g0080.1	*ERF29*
Nitab4.5_0004620g0090.3	*ERF130*
Nitab4.5_0004620g0095.1	*ERF16*
Nitab4.5_0006382g0040.1	*ERF110*

Identifier	Nic 2 ERF - Gene
Nitab4.5_0002924g0010.1	ERF17LI
Nitab4.5_0002924g0020.2	ERF179
Nitab4.5_0002924g0040.2	ERF17
Nitab4.5_0002924g0045.1	ERF168
Nitab4.5_0002924g0050.2	ERF115
Nitab4.5_0006499g0010.1	ERF104
Nitab4.5_0006499g0020.2	ERF221
Nitab4.5_0012667g0020.2	ERF91L1
Nitab4.5_0015055g0010.2	ERF189

Figure 8

Nitab4.5_0003090g0030.1 (ERF199)

SEQ ID No. 35.>Nitab4.5_0003090g0030.1 720 residues [genomic]

ATGGCAATGGAAATGAATCCAGCTGACGAAACCTTGTTTTTCTCCGACTCTCATCTCCTTGAATCGATAAAGCAA
CATCTTCTTGACGATTCAGATTTTTCTGAAATTTTTTCGTCGATGAATTCTAGCAACGAAATATTGCCTAACAGT
CCTAGCTCAAGTTTTAGCAGCTTCGACTTCGACTGCAGCTTCCTTAATTGGGATGAAAACTCTGAGGAAACATTA
ATACCAACTGATCAGAATCCTTCACATGAATCCCATGAAAAGTACTCCGAGTCCGAGGAGAAAACCCAGGGCCCT
GGGGTGGCGCGTGAGAAAAACGCGCCGCGAGATTGGACGCGGTACATAGGAGTGAAACGGCGACCGTGGGGGACG
TTTTCGGCGGAGACAAGAGACCCAAGTAGGAAAGGTGAAGGTGCAAGGCTGTGGTTAGGAACTTACGAGACCGCA
GAGGATGCAGCGTTAGCTTACGATCAAGCCGCTTTCAAAATCCGCGGCTCGAGAGCTCGGCTCAATTTTCCTCAC
TTAATCGGCTCAAACATGCCTAAGCCGGCTAGAGTTACAGCGAGGCGTAGTCGTACGCGCTCACCCGAGCCATCG
TCTTCTTCATCCACCTCATCATCAGAAAATGTACCAAGGAAAAGGAATATAGATGTGATAAATTCCATAGCCAAA
GCCAAATTCCTTTGTCATAGCTTGAATTTACAGAGATTAGCTTAA

Figure 9

SEQ ID No. 36. >Nitab4.5_0003090g0030.1 720 residues [cdna]

ATGGCAATGGAAATGAATCCAGCTGACGAAACCTTGTTTTTCTCCGACTCTCATCTCCTTGAATCGATAAAGCAA
CATCTTCTTGACGATTCAGATTTTTCTGAAATTTTTTCGTCGATGAATTCTAGCAACGAAATATTGCCTAACAGT
CCTAGCTCAAGTTTTAGCAGCTTCGACTTCGACTGCAGCTTCCTTAATTGGGATGAAAACTCTGAGGAAACATTA
ATACCAACTGATCAGAATCCTTCACATGAATCCCATGAAAAGTACTCCGAGTCCGAGGAGAAAACCCAGGGCCCT

GGGGTGGCGCGTGAGAAAAACGCGCCGCGAGATTGGACGCGGTACATAGGAGTGAAACGGCGACCGTGGGGGACG
TTTTCGGCGGAGACAAGAGACCCAAGTAGGAAAGGTGAAGGTGCAAGGCTGTGGTTAGGAACTTACGAGACCGCA
GAGGATGCAGCGTTAGCTTACGATCAAGCCGCTTTCAAAATCCGCGGCTCGAGAGCTCGGCTCAATTTTCCTCAC
TTAATCGGCTCAAACATGCCTAAGCCGGCTAGAGTTACAGCGAGGCGTAGTCGTACGCGCTCACCCGAGCCATCG
TCTTCTTCATCCACCTCATCATCAGAAAATGTACCAAGGAAAAGGAATATAGATGTGATAAATTCCATAGCCAAA
GCCAAATTCCTTTGTCATAGCTTGAATTTACAGAGATTAGCTTAA

Figure 10

SEQ ID No. 37. >Nitab4.5_0003090g0030.1 720 residues [cds]

ATGGCAATGGAAATGAATCCAGCTGACGAAACCTTGTTTTTCTCCGACTCTCATCTCCTTGAATCGATAAAGCAA
CATCTTCTTGACGATTCAGATTTTTCTGAAATTTTTTCGTCGATGAATTCTAGCAACGAAATATTGCCTAACAGT
CCTAGCTCAAGTTTTAGCAGCTTCGACTTCGACTGCAGCTTCCTTAATTGGGATGAAAACTCTGAGGAAACATTA
ATACCAACTGATCAGAATCCTTCACATGAATCCCATGAAAAGTACTCCGAGTCCGAGGAGAAAACCCAGGGCCCT
GGGGTGGCGCGTGAGAAAAACGCGCCGCGAGATTGGACGCGGTACATAGGAGTGAAACGGCGACCGTGGGGGACG
TTTTCGGCGGAGACAAGAGACCCAAGTAGGAAAGGTGAAGGTGCAAGGCTGTGGTTAGGAACTTACGAGACCGCA
GAGGATGCAGCGTTAGCTTACGATCAAGCCGCTTTCAAAATCCGCGGCTCGAGAGCTCGGCTCAATTTTCCTCAC
TTAATCGGCTCAAACATGCCTAAGCCGGCTAGAGTTACAGCGAGGCGTAGTCGTACGCGCTCACCCGAGCCATCG
TCTTCTTCATCCACCTCATCATCAGAAAATGTACCAAGGAAAAGGAATATAGATGTGATAAATTCCATAGCCAAA
GCCAAATTCCTTTGTCATAGCTTGAATTTACAGAGATTAGCTTAA

Figure 11

SEQ ID No. 38. >Nitab4.5_0003090g0030.1 239 residues [peptide]

MAMEMNPADETLFFSDSHLLESIKQHLLDDSDFSEIFSSMNSSNEILPNSPSSSFSSFDFDCSFLNWDENSEETL
IPTDQNPSHESHEKYSESEEKTQGPGVAREKNAPRDWTRYIGVKRRPWGTFSAETRDPSRKGEGARLWLGTYETA
EDAALAYDQAAFKIRGSRARLNFPHLIGSNMPKPARVTARRSRTRSPEPSSSSSTSSSENVPRKRNIDVINSIAK
AKFLCHSLNLQRLA

Figure 12

Nitab4.5_0015055g0010.2 (ERF189)

SEQ ID No. 39. >Nitab4.5_0015055g0010.2 699 residues [genomic]

ATGGAAATGAATCTAGCTGACGAAACCTTGTTTTTCTCTGAGTCTCATCTCCTTGAATCGATAAAGCAACATCTT
CTTGATGATTCAGATTTTTCTGAAATTTTTTCGCCGATGAGTTCAAGCAACGAAATATTGCCTAACAGTCCTAGC
TCAAGTTTTAGCAGCTTCGACTGCAGCTTCCTCAATTGGGATGAAAACTTTGAGGAAACATTAATACCAACTGAT
CAAAATCCTTCACATGAGAAGTGCTCCGAGTCCGAGGAGCAAACCCAGGGCCCAGCGGTGGTGCGTGAGAAAAAC
GCGCCGCGAGATTGGACGCGGTATATAGGAGTGAAACGGCGGCCGTGGGGGACGTTTTCGGCGGAGACAAGAGAC
CCAAGTAGGAAAGGTGAAGGTGCAAGGCTGTGGTTAGGAACTTACGAGACCGCAGAGGATGCAGCGTTGGCTTAC
GATCAAGCCGCTTTCAAAATCCGCGGCTCGAGAGCTCGGCTCAATTTTCCCCACTTAATTGGCTCAAACATGCCT
AAGCCGGCTAGAGTAACAGCGAGGCGTAGTCGTACGCGCTCACCCGAGCCATCCTCTTCTTCATCCACTTCATCA
TCAGAAAATGTGCCAAGAAAAAGGAATATAGATGTGATAAATTCCATAGCCAAAGCCAAATTCCTTTGTCATAGC
TTAAATTTACAGAGATTAGCTTAA

Figure 13

SEQ ID No. 40. >Nitab4.5_0015055g0010.2 699 residues [cdna]

ATGGAAATGAATCTAGCTGACGAAACCTTGTTTTTCTCTGAGTCTCATCTCCTTGAATCGATAAAGCAACATCTT
CTTGATGATTCAGATTTTTCTGAAATTTTTTCGCCGATGAGTTCAAGCAACGAAATATTGCCTAACAGTCCTAGC
TCAAGTTTTAGCAGCTTCGACTGCAGCTTCCTCAATTGGGATGAAAACTTTGAGGAAACATTAATACCAACTGAT
CAAAATCCTTCACATGAGAAGTGCTCCGAGTCCGAGGAGCAAACCCAGGGCCCAGCGGTGGTGCGTGAGAAAAAC
GCGCCGCGAGATTGGACGCGGTATATAGGAGTGAAACGGCGGCCGTGGGGGACGTTTTCGGCGGAGACAAGAGAC
CCAAGTAGGAAAGGTGAAGGTGCAAGGCTGTGGTTAGGAACTTACGAGACCGCAGAGGATGCAGCGTTGGCTTAC
GATCAAGCCGCTTTCAAAATCCGCGGCTCGAGAGCTCGGCTCAATTTTCCCCACTTAATTGGCTCAAACATGCCT
AAGCCGGCTAGAGTAACAGCGAGGCGTAGTCGTACGCGCTCACCCGAGCCATCCTCTTCTTCATCCACTTCATCA
TCAGAAAATGTGCCAAGAAAAAGGAATATAGATGTGATAAATTCCATAGCCAAAGCCAAATTCCTTTGTCATAGC
TTAAATTTACAGAGATTAGCTTAA

Figure 14

SEQ ID No. 41. >Nitab4.5_0015055g0010.2 699 residues [cds]

ATGGAAATGAATCTAGCTGACGAAACCTTGTTTTTCTCTGAGTCTCATCTCCTTGAATCGATAAAGCAACATCTT
CTTGATGATTCAGATTTTTCTGAAATTTTTTCGCCGATGAGTTCAAGCAACGAAATATTGCCTAACAGTCCTAGC
TCAAGTTTTAGCAGCTTCGACTGCAGCTTCCTCAATTGGGATGAAAACTTTGAGGAAACATTAATACCAACTGAT
CAAAATCCTTCACATGAGAAGTGCTCCGAGTCCGAGGAGCAAACCCAGGGCCCAGCGGTGGTGCGTGAGAAAAAC

GCGCCGCGAGATTGGACGCGGTATATAGGAGTGAAACGGCGGCCGTGGGGGACGTTTTCGGCGGAGACAAGAGAC
CCAAGTAGGAAAGGTGAAGGTGCAAGGCTGTGGTTAGGAACTTACGAGACCGCAGAGGATGCAGCGTTGGCTTAC
GATCAAGCCGCTTTCAAAATCCGCGGCTCGAGAGCTCGGCTCAATTTTCCCCACTTAATTGGCTCAAACATGCCT
AAGCCGGCTAGAGTAACAGCGAGGCGTAGTCGTACGCGCTCACCCGAGCCATCCTCTTCTTCATCCACTTCATCA
TCAGAAAATGTGCCAAGAAAAAGGAATATAGATGTGATAAATTCCATAGCCAAAGCCAAATTCCTTTGTCATAGC
TTAAATTTACAGAGATTAGCTTAA

Figure 15

SEQ ID No. 42.

>Nitab4.5_0015055g0010.2 232 residues [peptide]
MEMNLADETLFFSESHLLESIKQHLLDDSDFSEIFSPMSSSNEILPNSPSSSFSSFDCSFLNWDENFEETLIPTD
QNPSHEKCSESEEQTQGPAVVREKNAPRDWTRYIGVKRRPWGTFSAETRDPSRKGEGARLWLGTYETAEDAALAY
DQAAFKIRGSRARLNFPHLIGSNMPKPARVTARRSRTRSPEPSSSSSTSSSENVPRKRNIDVINSIAKAKFLCHS
LNLQRL

INTERNATIONAL SEARCH REPORT

International application No
PCT/GB2023/050211

A. CLASSIFICATION OF SUBJECT MATTER
INV. A01H5/12 A01H6/82 A24B15/00 C12N15/82 A24B15/10
C07K14/415 C12N9/12
ADD.
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED
Minimum documentation searched (classification system followed by classification symbols)
A01H A24B C12N C07K

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

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EPO-Internal, Sequence Search, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2020/099875 A1 (BRITISH AMERICAN TOBACCO INVESTMENTS LTD [GB]) 22 May 2020 (2020-05-22) claims 1-33; pages 9, 32-34, 36, 37	23-27
X	DATABASE UniProt [Online] 12 April 2017 (2017-04-12), "RecName: Full=FAD synthase {ECO:0000256|ARBA:ARBA00012393}; EC=2.7.7.2 {ECO:0000256|ARBA:ARBA00012393};", XP93040953, retrieved from EBI accession no. UNIPROT:Unreviewed Database accession no. Unreviewed sequence	7-27, 29-31

-/--

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

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance;; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance;; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
24 April 2023	11/05/2023

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Kurz, Birgit

INTERNATIONAL SEARCH REPORT

International application No
PCT/GB2023/050211

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	YRUELA INMACULADA ET AL: "Evolutionary divergence of chloroplast FAD synthetase proteins", BMC EVOLUTIONARY BIOLOGY, BIOMED CENTRAL LTD., LONDON, GB, vol. 10, no. 1, 18 October 2010 (2010-10-18), page 311, XP021072353, ISSN: 1471-2148, DOI: 10.1186/1471-2148-10-311 abstract; pages 1, 2; Figures 1-3 -----	1-31

INTERNATIONAL SEARCH REPORT

International application No.
PCT/GB2023/050211

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:

   a. [X] forming part of the international application as filed.

   b. [ ] furnished subsequent to the international filing date for the purposes of international search (Rule 13*ter*.1(a)).

      [ ] accompanied by a statement to the effect that the sequence listing does not go beyond the disclosure in the international application as filed.

2. [ ] With regard to any nucleotide and/or amino acid sequence disclosed in the international application, this report has been established to the extent that a meaningful search could be carried out without a WIPO Standard ST.26 compliant sequence listing.

3. Additional comments:

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/GB2023/050211

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2020099875 A1	22-05-2020	AR 117078 A1	07-07-2021
		BR 112021009579 A2	17-08-2021
		CN 113272434 A	17-08-2021
		EP 3880825 A1	22-09-2021
		US 2022033837 A1	03-02-2022
		WO 2020099875 A1	22-05-2020