

(12) STANDARD PATENT
(19) AUSTRALIAN PATENT OFFICE

(11) Application No. **AU 2012324024 B2**

(54) Title
Processes for producing lipids

(51) International Patent Classification(s)
C12N 15/54 (2006.01)

(21) Application No: **2012324024**

(22) Date of Filing: **2012.12.21**

(87) WIPO No: **WO13/096993**

(30) Priority Data

(31) Number
61/718,563
61/580,590

(32) Date
2012.10.25
2011.12.27

(33) Country
US
US

(43) Publication Date: **2013.07.11**

(43) Publication Journal Date: **2013.07.11**

(44) Accepted Journal Date: **2016.06.16**

(71) Applicant(s)
Commonwealth Scientific and Industrial Research Organisation

(72) Inventor(s)
Vanhercke, Thomas;Petrie, James Robertson;El Tahchy, Anna;Singh, Surinder Pal;Liu, Qing

(74) Agent / Attorney
FB Rice, L 14 90 Collins ST, Melbourne, VIC, 3000

(56) Related Art
US 2010/0184130 A1

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

(19) World Intellectual Property
Organization
International Bureau



(10) International Publication Number
WO 2013/096993 A1

(43) International Publication Date
4 July 2013 (04.07.2013)

(51) International Patent Classification:

C12N 15/54 (2006.01) *C11B 1/10* (2006.01)
C12N 15/29 (2006.01) *C12P 1/00* (2006.01)
A01H 5/00 (2006.01)

(21) International Application Number:

PCT/AU2012/001598

(22) International Filing Date:

21 December 2012 (21.12.2012)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

61/580,590 27 December 2011 (27.12.2011) US
61/718,563 25 October 2012 (25.10.2012) US

(71) Applicant: **COMMONWEALTH SCIENTIFIC AND INDUSTRIAL RESEARCH ORGANISATION** [AU/AU]; Limestone Avenue, Campbell, Australian Capital Territory 2612 (AU).

(72) Inventors: **VANHERCKE, Thomas**; 9/29 Moorhouse Street, O'Connor, Australian Capital Territory 2602 (AU). **PETRIE, James Robertson**; 16 Henry Street, Goulburn, New South Wales 2580 (AU). **EL TAHCHY, Anna**; 2/25 Harrison Street, Sculin, Australian Capital Territory 2614 (AU). **SINGH, Surinder Pal**; 10 Lucas Place, Downer, Australian Capital Territory 2602 (AU). **LIU, Qing**; 25 Mirrabucca Crescent, Giralang, Australian Capital Territory 2617 (AU).

(74) Agent: **FB RICE**; Level 23, 200 Queen Street, Melbourne, Victoria 3000 (AU).

(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, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, 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, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, 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: PROCESSES FOR PRODUCING LIPIDS

(57) Abstract: The present invention relates to methods of producing lipids. In particular, the present invention relates to methods of increasing the level of one or more non-polar lipids and/or the total non-polar lipid content in a vegetative plant part or a transgenic organism or part thereof.



WO 2013/096993 A1

PROCESSES FOR PRODUCING LIPIDS

FIELD OF THE INVENTION

The present invention relates to methods of producing lipids. In particular, the present invention relates to methods of increasing the level of one or more non-polar lipids and/or the total non-polar lipid content in a transgenic organism or part thereof. In one particular embodiment, the present invention relates to any combination of one or more monoacylglycerol acyltransferases (MGATs), diacylglycerol acyltransferases (DGATs), glycerol-3-phosphate acyltransferases (GPATs), oil body proteins and/or transcription factors regulating lipid biosynthesis while silencing key enzymatic steps in the starch biosynthesis and fatty acid desaturation pathways to increase the level of one or more non-polar lipids and/or the total non-polar lipid content and/or mono-unsaturated fatty acid content in plants or any part thereof including plant seed and/or leaves, algae and fungi.

BACKGROUND OF INVENTION

The majority of the world's energy, particularly for transportation, is supplied by petroleum derived fuels, which have a finite supply. Alternative sources which are renewable are needed, such as from biologically produced oils.

Triacylglycerol biosynthesis

Triacylglycerols (TAG) constitute the major form of lipids in seeds and consist of three acyl chains esterified to a glycerol backbone. The fatty acids are synthesized in the plastid as acyl-acyl carrier protein (ACP) intermediates where they can undergo a first desaturation catalyzed. This reaction is catalyzed by the stearyl-ACP desaturase and yields oleic acid (C18:1^{Δ9}). Subsequently, the acyl chains are transported to the cytosol and endoplasmic reticulum (ER) as acyl-Coenzyme (CoA) esters. Prior to entering the major TAG biosynthesis pathway, also known as the Kennedy or glycerol-3-phosphate (G3P) pathway, the acyl chains are typically integrated into phospholipids of the ER membrane where they can undergo further desaturation. Two key enzymes in the production of polyunsaturated fatty acids are the membrane-bound FAD2 and FAD3 desaturases which produce linoleic (C18:2^{Δ9,12}) and α-linolenic acid (C18:3^{Δ9,12,15}) respectively.

TAG biosynthesis via the Kennedy pathway consists of a series of subsequent acylations, each using acyl-CoA esters as the acyl-donor. The first acylation step typically occurs at the *sn*1-position of the G3P backbone and is catalyzed by the glycerol-3-phosphate acyltransferase (*sn*1-GPAT). The product, *sn*1-lysophosphatidic acid (*sn*1-LPA) serves as a substrate for the lysophosphatidic acid acyltransferase

(LPAAT) which couples a second acyl chain at the *sn2*-position to form phosphatidic acid. PA is further dephosphorylated to diacylglycerol (DAG) by the phosphatidic acid phosphatase (PAP) thereby providing the substrate for the final acylation step. Finally, a third acyl chain is esterified to the *sn3*-position of DAG in a reaction catalyzed by the diacylglycerol acyltransferase (DGAT) to form TAG which accumulates in oil bodies. A second enzymatic reaction, phosphatidyl glycerol acyltransferase (PDAT), also results in the conversion of DAG to TAG. This reaction is unrelated to DGAT and uses phospholipids as the acyl-donors.

To maximise yields for the commercial production of lipids, there is a need for further means to increase the levels of lipids, particularly non-polar lipids such as DAGs and TAGs, in transgenic organisms or parts thereof such as plants, seeds, leaves, algae and fungi. Attempts at increasing neutral lipid yields in plants have mainly focused on individual critical enzymatic steps involved in fatty acid biosynthesis or TAG assembly. These strategies, however, have resulted in modest increases in seed or leaf oil content. Recent metabolic engineering work in the oleaginous yeast *Yarrowia lipolytica* has demonstrated that a combined approach of increasing glycerol-3-phosphate production and preventing TAG breakdown via β -oxidation resulted in cumulative increases in the total lipid content (Dulermo et al., 2011).

Plant lipids such as seedoil triacylglycerols (TAGs) have many uses, for example, culinary uses (shortening, texture, flavor), industrial uses (in soaps, candles, perfumes, cosmetics, suitable as drying agents, insulators, lubricants) and provide nutritional value. There is also growing interest in using plant lipids for the production of biofuel.

To maximise yields for the commercial biological production of lipids, there is a need for further means to increase the levels of lipids, particularly non-polar lipids such as DAGs and TAGs, in transgenic organisms or parts thereof such as plants, seeds, leaves, algae and fungi.

SUMMARY OF THE INVENTION

The present inventors have demonstrated significant increases in the lipid content of organisms, particularly in the vegetative parts and seed of plants, by manipulation of both fatty acid biosynthesis and lipid assembly pathways. Various combinations of genes were used to achieve substantial increases in oil content, which is of great significance for production of biofuels and other industrial products derived from oil.

In a first aspect, the invention provides a process for producing an industrial product from a vegetative plant part or non-human organism or part thereof comprising high levels of non-polar lipid.

In an embodiment, the invention provides a process for producing an industrial product, the process comprising the steps of:

i) obtaining a vegetative plant part having a total non-polar lipid content of at least about 3%, preferably at least about 5% or at least about 7% (w/w dry weight),

ii) converting at least some of the lipid *in situ* in the vegetative plant part to the industrial product by heat, chemical, or enzymatic means, or any combination thereof, and

iii) recovering the industrial product, thereby producing the industrial product.

In another embodiment, the process for producing an industrial product comprises the steps of:

i) obtaining a vegetative plant part having a total non-polar lipid content of at least about 3%, preferably at least about 5% or at least about 7% (w/w dry weight),

ii) physically processing the vegetative plant part of step i),

iii) converting at least some of the lipid in the processed vegetative plant part to the industrial product by applying heat, chemical, or enzymatic means, or any combination thereof, to the lipid in the processed vegetative plant part, and

iv) recovering the industrial product, thereby producing the industrial product.

In another embodiment, the process for producing an industrial product comprises the steps of:

i) obtaining a non-human organism or a part thereof comprising one or more exogenous polynucleotide(s), wherein each of the one or more exogenous polynucleotide(s) is operably linked to a promoter which is capable of directing expression of the polynucleotide in a non-human organism or a part thereof, and wherein the non-human organism or part thereof has an increased level of one or more non-polar lipids relative to a corresponding non-human organism or a part thereof lacking the one or more exogenous polynucleotide(s), and

ii) converting at least some of the lipid *in situ* in the non-human organism or part thereof to the industrial product by heat, chemical, or enzymatic means, or any combination thereof, and

iii) recovering the industrial product, thereby producing the industrial product.

In a second aspect, the invention provides a process for producing an industrial product, the process comprising the steps of:

(i) obtaining a non-human organism or a part thereof as described herein, and
 (ii) optionally, physically processing the non-human organism or part thereof of step (i), and

(iii) converting at least some of the lipid in the non-human organism or part thereof of step (i), or in the processed non-human organism or part thereof obtained by step (ii), to the industrial product by applying heat, chemical, or enzymatic means, or any combination thereof, to the lipid, *in situ* in the non-human organism or part thereof of step (i), or in the processed non-human organism or part thereof obtained by step (ii), and

(iv) recovering the industrial product, thereby producing the industrial product.

In a further embodiment of the above aspects, the process for producing an industrial product comprises the steps of:

i) obtaining a non-human organism or a part thereof comprising one or more exogenous polynucleotides, wherein the non-human organism or part thereof has an increased level of one or more non-polar lipids relative to a corresponding non-human organism or a part thereof lacking the one or more exogenous polynucleotides,

ii) physically processing the non-human organism or part thereof of step i),

iii) converting at least some of the lipid in the processed non-human organism or part thereof to the industrial product by applying heat, chemical, or enzymatic means, or any combination thereof, to the lipid in the processed non-human organism or part thereof, and

iv) recovering the industrial product, thereby producing the industrial product.

In each of the above embodiments, it would be understood by a person skilled in the art that the converting step could be done simultaneously with or subsequent to the physical processing step.

In each of the above embodiments, the total non-polar lipid content of the vegetative plant part, or non-human organism or part thereof, preferably a plant leaf or part thereof, stem or tuber, is at least about 3%, more preferably at least about 5%, preferably at least about 7%, more preferably at least about 10%, more preferably at least about 11%, more preferably at least about 12%, more preferably at least about 13%, more preferably at least about 14%, or more preferably at least about 15% (w/w dry weight). In a further preferred embodiment, the total non-polar lipid content is between 5% and 25%, between 7% and 25%, between 10% and 25%, between 12% and 25%, between 15% and 25%, between 7% and 20%, between 10% and 20%, between 10% and 15%, between 15% and 20%, between 20% and 25%, about 10%, about 11%, about 12%, about 13%, about 14%, about 15%, about 16%, about 17%,

about 18%, about 20%, or about 22%, each as a percentage of dry weight. In a particularly preferred embodiment, the vegetative plant part is a leaf (or leaves) or a portion thereof. In a more preferred embodiment, the vegetative plant part is a leaf portion having a surface area of at least 1 cm².

Furthermore, in each of the above embodiments, the total TAG content of the vegetative plant part, or non-human organism or part thereof, preferably a plant leaf or part thereof, stem or tuber, is at least about 3%, more preferably at least about 5%, preferably at least about 7%, more preferably at least about 10%, more preferably at least about 11%, more preferably at least about 12%, more preferably at least about 13%, more preferably at least about 14%, more preferably at least about 15%, or more preferably at least about 17% (w/w dry weight). In a further preferred embodiment, the total TAG content is between 5% and 30%, between 7% and 30%, between 10% and 30%, between 12% and 30%, between 15% and 30%, between 7% and 30%, between 10% and 30%, between 20% and 28%, between 18% and 25%, between 22% and 30%, about 10%, about 11%, about 12%, about 13%, about 14%, about 15%, about 16%, about 17%, about 18%, about 20%, or about 22%, each as a percentage of dry weight. In a particularly preferred embodiment, the vegetative plant part is a leaf (or leaves) or a portion thereof. In a more preferred embodiment, the vegetative plant part is a leaf portion having a surface area of at least 1 cm².

Furthermore, in each of the above embodiments, the total lipid content of the vegetative plant part, or non-human organism or part thereof, preferably a plant leaf or part thereof, stem or tuber, is at least about 3%, more preferably at least about 5%, preferably at least about 7%, more preferably at least about 10%, more preferably at least about 11%, more preferably at least about 12%, more preferably at least about 13%, more preferably at least about 14%, more preferably at least about 15%, more preferably at least about 17% (w/w dry weight), more preferably at least about 20%, more preferably at least about 25%. In a further preferred embodiment, the total lipid content is between 5% and 35%, between 7% and 35%, between 10% and 35%, between 12% and 35%, between 15% and 35%, between 7% and 35%, between 10% and 20%, between 18% and 28%, between 20% and 28%, between 22% and 28%, about 10%, about 11%, about 12%, about 13%, about 14%, about 15%, about 16%, about 17%, about 18%, about 20%, about 22%, or about 25%, each as a percentage of dry weight. Typically, the total lipid content of the vegetative plant part, or non-human organism or part thereof is about 2-3% higher than the non-polar lipid content. In a particularly preferred embodiment, the vegetative plant part is a leaf (or leaves) or a portion thereof. In a more preferred embodiment, the vegetative plant part is a leaf portion having a surface area of at least 1 cm².

The industrial product may be a hydrocarbon product such as fatty acid esters, preferably fatty acid methyl esters and/or a fatty acid ethyl esters, an alkane such as methane, ethane or a longer-chain alkane, a mixture of longer chain alkanes, an alkene, a biofuel, carbon monoxide and/or hydrogen gas, a bioalcohol such as ethanol, propanol, or butanol, biochar, or a combination of carbon monoxide, hydrogen and biochar. The industrial product may be a mixture of any of these components, such as a mixture of alkanes, or alkanes and alkenes, preferably a mixture which is predominantly (>50%) C4-C8 alkanes, or predominantly C6 to C10 alkanes, or predominantly C6 to C8 alkanes. The industrial product is not carbon dioxide and not water, although these molecules may be produced in combination with the industrial product. The industrial product may be a gas at atmospheric pressure/room temperature, or preferably, a liquid, or a solid such as biochar, or the process may produce a combination of a gas component, a liquid component and a solid component such as carbon monoxide, hydrogen gas, alkanes and biochar, which may subsequently be separated. In an embodiment of the above aspects, the hydrocarbon product is predominantly fatty acid methyl esters. In an alternative embodiment of the above aspects, the hydrocarbon product is a product other than fatty acid methyl esters.

The industrial product may be an intermediate product, for example, a product comprising fatty acids, which can subsequently be converted to, for example, biofuel by, for example, trans-esterification to fatty acid esters.

Heat may be applied in the process, such as by pyrolysis, combustion, gasification, or together with enzymatic digestion (including anaerobic digestion, composting, fermentation). Lower temperature gasification takes place at, for example, between about 700°C to about 1000°C. Higher temperature gasification takes place at, for example, between about 1200°C to about 1600°C. Lower temperature pyrolysis (slower pyrolysis), takes place at, for example, about 400°C, whereas higher temperature pyrolysis takes place at, for example, about 500°C. Mesophilic digestion takes place between, for example, about 20°C and about 40°C. Thermophilic digestion takes place from, for example, about 50°C to about 65°C.

Chemical means include, but are not limited to, catalytic cracking, anaerobic digestion, fermentation, composting and transesterification. In an embodiment of the above aspects, a chemical means uses a catalyst or mixture of catalysts, which may be applied together with heat. The process may use a homogeneous catalyst, a heterogeneous catalyst and/or an enzymatic catalyst. In an embodiment of the above aspects, the catalyst is a transition metal catalyst, a molecular sieve type catalyst, an activated alumina catalyst or sodium carbonate. Catalysts include acid catalysts such as sulphuric acid, or alkali catalysts such as potassium or sodium hydroxide or other

hydroxides. The chemical means may comprise transesterification of fatty acids in the lipid, which process may use a homogeneous catalyst, a heterogeneous catalyst and/or an enzymatic catalyst. The conversion may comprise pyrolysis, which applies heat and may apply chemical means, and may use a transition metal catalyst, a molecular sieve type catalyst, an activated alumina catalyst and/or sodium carbonate.

Enzymatic means include, but are not limited to, digestion by microorganisms in, for example, anaerobic digestion, fermentation or composting, or by recombinant enzymatic proteins.

The lipid that is converted to an industrial product in this aspect of the invention may be some, or all, of the non-polar lipid in the vegetative plant part or non-human organism or part thereof, or preferably the conversion is of at least some of the non-polar lipid and at least some of the polar lipid, and more preferably essentially all of the lipid (both polar and non-polar) in the vegetative plant part or non-human organism or part thereof is converted to the industrial product(s).

In an embodiment of the above aspects, the conversion of the lipid to the industrial product occurs *in situ* without physical disruption of the vegetative plant part or non-human organism or part thereof. In this embodiment of the above aspects, the vegetative plant part or non-human organism or part thereof may first be dried, for example by the application of heat, or the vegetative plant part or non-human organism or part thereof may be used essentially as harvested, without drying. In an alternative embodiment of the above aspects, the process comprises a step of physically processing the vegetative plant part, or the non-human organism or part thereof. The physical processing may comprise one or more of rolling, pressing such as flaking, crushing or grinding the vegetative plant part, non human organism or part thereof, which may be combined with drying of the vegetative plant part, or the non-human organism or part thereof. For example, the vegetative plant part, or non-human organism or part thereof may first be substantially dried and then ground to a finer material, for ease of subsequent processing.

In an embodiment of the above aspects, the weight of the vegetative plant part, or the non-human organism or part thereof used in the process is at least 1 kg or preferably at least 1 tonne (dry weight) of pooled vegetative plant parts, or the non-human organisms or parts thereof. The processes may further comprise a first step of harvesting vegetative plant parts, for example from at least 100 or 1000 plants grown in a field, to provide a collection of at least 1000 such vegetative plant parts, i.e., which are essentially identical. Preferably, the vegetative plant parts are harvested at a time when the yield of non-polar lipids are at their highest. In one embodiment of the above aspects, the vegetative plant parts are harvested about at the time of flowering. In another embodiment of the above aspects, the vegetative plant parts are

harvested from about at the time of flowering to about the beginning of senescence. In another embodiment of the above aspects, the vegetative plant parts are harvested when the plants are at least about 1 month of age.

The process may or may not further comprise extracting some of the non-polar lipid content of the vegetative plant part, or the non-human organism or part thereof prior to the conversion step. In an embodiment of the above aspects, the process further comprises steps of:

(a) extracting at least some of the non-polar lipid content of the vegetative plant part or the non-human organism or part thereof as non-polar lipid, and

(b) recovering the extracted non-polar lipid, wherein steps (a) and (b) are performed prior to the step of converting at least some of the lipid in the vegetative plant part, or the non-human organism or part thereof to the industrial product. The proportion of non-polar lipid that is first extracted may be less than 50%, or more than 50%, or preferably at least 75% of the total non-polar lipid in the vegetative plant part, or non-human organism or part thereof. In this embodiment, the extracted non-polar lipid comprises triacylglycerols, wherein the triacylglycerols comprise at least 90%, preferably at least 95% of the extracted lipid. The extracted lipid may itself be converted to an industrial product other than the lipid itself, for example by trans-esterification to fatty acid esters.

In a third aspect, the invention provides a process for producing extracted lipid from a non-human organism or a part thereof.

In an embodiment, the invention provides a process for producing extracted lipid, the process comprising the steps of:

i) obtaining a non-human organism or a part thereof comprising one or more exogenous polynucleotide(s) and an increased level of one or more non-polar lipid(s) relative to a corresponding non-human organism or a part thereof, respectively, lacking the one or more exogenous polynucleotide(s),

ii) extracting lipid from the non-human organism or part thereof, and

iii) recovering the extracted lipid,

thereby producing the extracted lipid, wherein each of the one or more exogenous polynucleotides is operably linked to a promoter which is capable of directing expression of the polynucleotide in a non-human organism or part thereof, and wherein one or more or all of the following features apply:

(a) the one or more exogenous polynucleotide(s) comprise a first exogenous polynucleotide which encodes an RNA or transcription factor polypeptide that increases the expression of one or more glycolytic or fatty acid biosynthetic genes in a non-human organism or a part thereof, and a second exogenous polynucleotide which

encodes an RNA or polypeptide involved in biosynthesis of one or more non-polar lipids,

5 (b) if the non-human organism is a plant, a vegetative part of the plant has a total non-polar lipid content of at least about 3%, more preferably at least about 5%, preferably at least about 7%, more preferably at least about 10%, more preferably at least about 11%, more preferably at least about 12%, more preferably at least about 13%, more preferably at least about 14%, or more preferably at least about 15% (w/w dry weight),

10 (c) the non-human organism is an alga selected from the group consisting of diatoms (bacillariophytes), green algae (chlorophytes), blue-green algae (cyanophytes), golden-brown algae (chrysophytes), haptophytes, brown algae and heterokont algae,

15 (d) the one or more non-polar lipid(s) comprise a fatty acid which comprises a hydroxyl group, an epoxy group, a cyclopropane group, a double carbon-carbon bond, a triple carbon-carbon bond, conjugated double bonds, a branched chain such as a methylated or hydroxylated branched chain, or a combination of two or more thereof, or any of two, three, four, five or six of the aforementioned groups, bonds or branched chains,

20 (e) the total fatty acid content in the non-polar lipid(s) comprises at least 2% more oleic acid and/or at least 2% less palmitic acid than the non-polar lipid(s) in the corresponding non-human organism or part thereof lacking the one or more exogenous polynucleotides,

25 (f) the non-polar lipid(s) comprise a modified level of total sterols, preferably free (non-esterified) sterols, steroyl esters, steroyl glycosides, relative to the non-polar lipid(s) in the corresponding non-human organism or part thereof lacking the one or more exogenous polynucleotides,

(g) the non-polar lipid(s) comprise waxes and/or wax esters,

30 (h) the non-human organism or part thereof is one member of a pooled population or collection of at least 1000 such non-human organisms or parts thereof, respectively, from which the lipid is extracted.

In an embodiment of (b) above, the total non-polar lipid content is between 5% and 25%, between 7% and 25%, between 10% and 25%, between 12% and 25%, between 15% and 25%, between 7% and 20%, between 10% and 20%, about 10%, about 11%, about 12%, about 13%, about 14%, about 15%, about 16%, about 17%, 35 about 18%, about 20%, or about 22%, each as a percentage of dry weight.

In an embodiment, the non-human organism is an alga, or an organism suitable for fermentation such as a fungus, or preferably a plant. The part of the non-human organism may be a seed, fruit, or a vegetative part of a plant. In a preferred

embodiment, the plant part is a leaf portion having a surface area of at least 1 cm². In another preferred embodiment, the non-human organism is a plant, the part is a plant seed and the extracted lipid is seedoil. In a more preferred embodiment, the plant is from an oilseed species, which is used commercially or could be used commercially for oil production. The species may be selected from a group consisting of a *Acrocomia aculeata* (macauba palm), *Arabidopsis thaliana*, *Aracinis hypogaea* (peanut), *Astrocaryum murumuru* (murumuru), *Astrocaryum vulgare* (tucumã), *Attalea geraensis* (Indaiá-rateiro), *Attalea humilis* (American oil palm), *Attalea oleifera* (andaiá), *Attalea phalerata* (uricuri), *Attalea speciosa* (babassu), *Avena sativa* (oats), *Beta vulgaris* (sugar beet), *Brassica* sp. such as *Brassica carinata*, *Brassica juncea*, *Brassica napobrassica*, *Brassica napus* (canola), *Camelina sativa* (false flax), *Cannabis sativa* (hemp), *Carthamus tinctorius* (safflower), *Caryocar brasiliense* (pequi), *Cocos nucifera* (Coconut), *Crambe abyssinica* (Abyssinian kale), *Cucumis melo* (melon), *Elaeis guineensis* (African palm), *Glycine max* (soybean), *Gossypium hirsutum* (cotton), *Helianthus* sp. such as *Helianthus annuus* (sunflower), *Hordeum vulgare* (barley), *Jatropha curcas* (physic nut), *Joannesia princeps* (arara nut-tree), *Lemna* sp. (duckweed) such as *Lemna aequinoctialis*, *Lemna disperma*, *Lemna ecuadoriensis*, *Lemna gibba* (swollen duckweed), *Lemna japonica*, *Lemna minor*, *Lemna minuta*, *Lemna obscura*, *Lemna paucicostata*, *Lemna perpusilla*, *Lemna tenera*, *Lemna trisulca*, *Lemna turionifera*, *Lemna valdiviana*, *Lemna yungensis*, *Licania rigida* (oiticica), *Linum usitatissimum* (flax), *Lupinus angustifolius* (lupin), *Mauritia flexuosa* (buriti palm), *Maximiliana maripa* (inaja palm), *Miscanthus* sp. such as *Miscanthus x giganteus* and *Miscanthus sinensis*, *Nicotiana* sp. (tobacco) such as *Nicotiana tabacum* or *Nicotiana benthamiana*, *Oenocarpus bacaba* (bacaba-do-azeite), *Oenocarpus bataua* (patauã), *Oenocarpus distichus* (bacaba-de-leque), *Oryza* sp. (rice) such as *Oryza sativa* and *Oryza glaberrima*, *Panicum virgatum* (switchgrass), *Paraqueiba paraensis* (mari), *Persea amencana* (avocado), *Pongamia pinnata* (Indian beech), *Populus trichocarpa*, *Ricinus communis* (castor), *Saccharum* sp. (sugarcane), *Sesamum indicum* (sesame), *Solanum tuberosum* (potato), *Sorghum* sp. such as *Sorghum bicolor*, *Sorghum vulgare*, *Theobroma grandiflorum* (cupuassu), *Trifolium* sp., *Trithrinax brasiliensis* (Brazilian needle palm), *Triticum* sp. (wheat) such as *Triticum aestivum* and *Zea mays* (corn). In an embodiment, the *Brassica napus* plant is of the variety Westar. In an alternative embodiment, if the plant is *Brassica napus*, it is of a variety or cultivar other than Westar. In an embodiment, the plant is of a species other than *Arabidopsis thaliana*. In another embodiment, the plant is of a species other than *Nicotiana tabacum*. In another embodiment, the plant is of a species other than *Nicotiana benthamiana*. In one embodiment, the plant is a perennial, for example, a switchgrass. Each of the features described for the plant of

the third aspect can be applied *mutatis mutandis* to the vegetative plant part of the first and/or second aspects.

In an embodiment, the non-human organism is an oleaginous fungus such as an oleaginous yeast.

In a preferred embodiment, the lipid is extracted without drying the non-human organism or part thereof prior to the extraction. The extracted lipid may subsequently be dried or fractionated to reduce its moisture content.

In further embodiments of this aspect, the invention provides a process for producing extracted lipid from specific oilseed plants. In an embodiment, the invention provides a process for producing extracted canola oil, the process comprising the steps of:

- i) obtaining canola seed comprising at least 45% seedoil on a weight basis,
- ii) extracting oil from the canola seed, and
- iii) recovering the oil, wherein the recovered oil comprises at least 90% (w/w)

triacylglycerols (TAG),

thereby producing the canola oil. In a preferred embodiment, the canola seed has an oil content on a weight basis of at least 46%, at least 47%, at least 48%, at least 49%, at least 50%, at least 51%, at least 52%, at least 53%, at least 54%, at least 55% or at least 56%. The oil content is determinable by measuring the amount of oil that is extracted from the seed, which is threshed seed as commonly harvested, and calculated as a percentage of the seed weight, i.e., % (w/w). Moisture content of the canola seed is between 5% and 15%, and is preferably about 8.5%. In an embodiment, the oleic acid content is between about 58% and 62% of the total fatty acid in the canola oil, preferably at least 63%, and the palmitic acid content is about 4% to about 6% of the total fatty acids in the canola oil. Preferred canola oil has an iodine value of 110-120 and a chlorophyll level of less than 30ppm.

In another embodiment, the invention provides a process for producing extracted cornseed oil, the process comprising the steps of:

- i) obtaining corn seed comprising at least 5% seedoil on a weight basis,
- ii) extracting oil from the corn seed, and
- iii) recovering the oil, wherein the recovered oil comprises at least 80%, preferably at least 85% or at least 90% (w/w) triacylglycerols (TAG),

thereby producing the cornseed oil. In a preferred embodiment, the corn seed has an oil content on a seed weight basis (w/w) of at least 6%, at least 7%, at least 8%, at least 9%, at least 10%, at least 11%, at least 12% or at least 13%. The moisture content of the cornseed is about 13% to about 17%, preferably about 15%. Preferred corn oil comprises about 0.1% tocopherols.

In another embodiment, the invention provides a process for producing extracted soybean oil, the process comprising the steps of:

- i) obtaining soybean seed comprising at least 20% seedoil on a weight basis,
- ii) extracting oil from the soybean seed, and
- 5 iii) recovering the oil, wherein the recovered oil comprises at least 90% (w/w) triacylglycerols (TAG),

thereby producing the soybean oil. In a preferred embodiment, the soybean seed has an oil content on a seed weight basis (w/w) of at least 21%, at least 22%, at least 23%, at least 24%, at least 25%, at least 26%, at least 27%, at least 28%, at least 29%, at
10 least 30%, or at least 31%. In an embodiment, the oleic acid content is between about 20% and about 25% of the total fatty acid in the soybean oil, preferably at least 30%, the linoleic acid content is between about 45% and about 57%, preferably less than 45%, and the palmitic acid content is about 10% to about 15% of the total fatty acids in the soybean oil, preferably less than 10%. Preferably the soybean seed has a
15 protein content of about 40% on a dry weight basis, and the moisture content of the soybean seed is about 10% to about 16%, preferably about 13%.

In another embodiment, the invention provides a process for producing extracted lupinseed oil, the process comprising the steps of:

- i) obtaining lupin seed comprising at least 10% seedoil on a weight basis,
- 20 ii) extracting oil from the lupin seed, and
- iii) recovering the oil, wherein the recovered oil comprises at least 90% (w/w) triacylglycerols (TAG),

thereby producing the lupinseed oil. In a preferred embodiment, the lupin seed has an oil content on a seed weight basis (w/w) of at least 11%, at least 12%, at least 13%, at
25 least 14%, at least 15%, or at least 16%.

In another embodiment, the invention provides a process for producing extracted peanut oil, the process comprising the steps of:

- i) obtaining peanuts comprising at least 50% seedoil on a weight basis,
- ii) extracting oil from the peanuts, and
- 30 iii) recovering the oil, wherein the recovered oil comprises at least 90% (w/w) triacylglycerols (TAG),

thereby producing the peanut oil. In a preferred embodiment, the peanut seed (peanuts) have an oil content on a seed weight basis (w/w) of at least 51%, at least 52%, at least 53%, at least 54%, at least 55% or at least 56%. In an embodiment, the
35 oleic acid content is between about 38% and 59% of the total fatty acid in the peanut oil, preferably at least 60%, and the palmitic acid content is about 9% to about 13% of the total fatty acids in the peanut oil, preferably less than 9%.

In another embodiment, the invention provides a process for producing extracted sunflower oil, the process comprising the steps of:

- i) obtaining sunflower seed comprising at least 50% seedoil on a weight basis,
- ii) extracting oil from the sunflower seed, and
- 5 iii) recovering the oil, wherein the recovered oil comprises at least 90% (w/w) triacylglycerols (TAG),

thereby producing the sunflower oil. In a preferred embodiment, the sunflower seed have an oil content on a seed weight basis (w/w) of at least 51%, at least 52%, at least 53%, at least 54%, or at least 55%.

10 In another embodiment, the invention provides a process for producing extracted cottonseed oil, the process comprising the steps of:

- i) obtaining cottonseed comprising at least 41% seedoil on a weight basis,
- ii) extracting oil from the cottonseed, and
- 15 iii) recovering the oil, wherein the recovered oil comprises at least 90% (w/w) triacylglycerols (TAG),

thereby producing the cottonseed oil. In a preferred embodiment, the cotton seed have an oil content on a seed weight basis (w/w) of at least 42%, at least 43%, at least 44%, at least 45%, at least 46%, at least 47%, at least 48%, at least 49%, or at least 50%. In an embodiment, the oleic acid content is between about 15% and 22% of the total fatty acid in the cotton oil, preferably at least 22%, the linoleic acid content is between about 45% and about 57%, preferably less than 45%, and the palmitic acid content is about 20% to about 26% of the total fatty acids in the cottonseed oil, preferably less than 18%. In an embodiment, the cottonseed oil also contains cyclopropanated fatty acids such as sterculic and malvalic acids, and may contain small amounts of gossypol.

25 In another embodiment, the invention provides a process for producing extracted safflower oil, the process comprising the steps of:

- i) obtaining safflower seed comprising at least 35% seedoil on a weight basis,
- ii) extracting oil from the safflower seed, and
- 30 iii) recovering the oil, wherein the recovered oil comprises at least 90% (w/w) triacylglycerols (TAG),

thereby producing the safflower oil. In a preferred embodiment, the safflower seed have an oil content on a seed weight basis (w/w) of at least 36%, at least 37%, at least 38%, at least 39%, at least 40%, at least 41%, at least 42%, at least 43%, at least 44%, 35 or at least 45%.

In another embodiment, the invention provides a process for producing extracted flaxseed oil, the process comprising the steps of:

- i) obtaining flax seed comprising at least 36% seedoil on a weight basis,

- ii) extracting oil from the flax seed, and
- iii) recovering the oil, wherein the recovered oil comprises at least 90% (w/w) triacylglycerols (TAG),

thereby producing the flaxseed oil. In a preferred embodiment, the flax seed have an oil content on a seed weight basis (w/w) of at least 37%, at least 38%, at least 39%, or at least 40%.

In another embodiment, the invention provides a process for producing extracted Camelina oil, the process comprising the steps of:

- i) obtaining *Camelina sativa* seed comprising at least 36% seedoil on a weight basis,
- ii) extracting oil from the *Camelina sativa* seed, and
- iii) recovering the oil, wherein the recovered oil comprises at least 90% (w/w) triacylglycerols (TAG),

thereby producing the Camelina oil. In a preferred embodiment, the *Camelina sativa* seed have an oil content on a seed weight basis (w/w) of at least 37%, at least 38%, at least 39%, at least 40%, at least 41%, at least 42%, at least 43%, at least 44%, or at least 45%.

The process of the third aspect may also comprise measuring the oil and/or protein content of the seed by near-infrared reflectance spectroscopy as described in Hom et al. (2007).

In an embodiment, the process of the third aspect of the invention comprises partially or completely drying the vegetative plant part, or the non-human organism, or part thereof, or the seed, and/or one or more of rolling, pressing such as flaking, crushing or grinding the vegetative plant part, or the non-human organism or part thereof, or the seed, or any combination of these methods, in the extraction process. The process may use an organic solvent (e.g., hexane such as n-hexane or a combination of n-hexane with isohexane, or butane alone or in combination with hexane) in the extraction process to extract the lipid or oil or to increase the efficiency of the extraction process, particularly in combination with a prior drying process to reduce the moisture content.

In an embodiment, the process comprises recovering the extracted lipid or oil by collecting it in a container, and/or purifying the extracted lipid or seedoil, such as, for example, by degumming, deodorising, decolourising, drying and/or fractionating the extracted lipid or oil, and/or removing at least some, preferably substantially all, waxes and/or wax esters from the extracted lipid or oil. The process may comprise analysing the fatty acid composition of the extracted lipid or oil, such as, for example, by converting the fatty acids in the extracted lipid or oil to fatty acid methyl esters and analysing these using GC to determine the fatty acid composition. The fatty acid

composition of the lipid or oil is determined prior to any fractionation of the lipid or oil that alters its fatty acid composition. The extracted lipid or oil may comprise a mixture of lipid types and/or one or more derivatives of the lipids, such as free fatty acids.

5 In an embodiment, the process of the third aspect of the invention results in substantial quantities of extracted lipid or oil. In an embodiment, the volume of the extracted lipid or oil is at least 1 litre, preferably at least 10 litres. In a preferred embodiment, the extracted lipid or oil is packaged ready for transportation or sale.

10 In an embodiment, the extracted lipid or oil comprises at least 91%, at least 92%, at least 93%, at least 94%, at least 95% or at least 96% TAG on a weight basis. The extracted lipid or oil may comprise phospholipid as a minor component, up to about 8% by weight, preferably less than 5% by weight, and more preferably less than 3% by weight.

15 In an embodiment, the process results in extracted lipid or oil wherein one or more or all of the following features apply:

(i) triacylglycerols comprise at least 90%, preferably at least 95% or 96%, of the extracted lipid or oil,

(ii) the extracted lipid or oil comprises free sterols, steroyl esters, steroyl glycosides, waxes or wax esters, or any combination thereof, and

20 (iii) the total sterol content and/or composition in the extracted lipid or oil is significantly different to the sterol content and/or composition in the extracted lipid or oil produced from a corresponding non-human organism or part thereof, or seed.

In an embodiment, the process further comprises converting the extracted lipid or oil to an industrial product. That is, the extracted lipid or oil is converted post-extraction to another chemical form which is an industrial product. Preferably, the industrial product is a hydrocarbon product such as fatty acid esters, preferably fatty acid methyl esters and/or fatty acid ethyl esters, an alkane such as methane, ethane or a longer-chain alkane, a mixture of longer chain alkanes, an alkene, a biofuel, carbon monoxide and/or hydrogen gas, a bioalcohol such as ethanol, propanol, or butanol, biochar, or a combination of carbon monoxide, hydrogen and biochar.

25 In the process of either the first, second or third aspects of the invention, the vegetative plant part, or the part of the non-human organism may be an aerial plant part or a green plant part such as a plant leaf or stem, a woody part such as a stem, branch or trunk, or a root or tuber. Preferably, the plants are grown in a field and the parts such as seed harvested from the plants in the field.

35 In an embodiment, the process further comprises a step of harvesting the vegetative plant part, non-human organism or part thereof, preferably with a mechanical harvester.

Preferably, the vegetative plant parts are harvested at a time when the yield of non-polar lipids are at their highest. In one embodiment, the vegetative plant parts are harvested about at the time of flowering. In another embodiment, the vegetative plant parts are harvested from about at the time of flowering to about the beginning of senescence. In another embodiment, the vegetative plant parts are harvested when the plants are at least about 1 month of age.

If the organism is an algal or fungal organism, the cells may be grown in an enclosed container or in an open-air system such as a pond. The resultant organisms comprising the non-polar lipid may be harvested, such as, for example, by a process comprising filtration, centrifugation, sedimentation, flotation or flocculation of algal or fungal organisms such as by adjusting pH of the medium. Sedimentation is less preferred.

In the process of the third aspect of the invention, the total non-polar lipid content of the non-human organism or part thereof, such a vegetative plant part or seed, is increased relative to a corresponding vegetative plant part, non-human organism or part thereof, or seed.

In an embodiment, the vegetative plant part, or non-human organism or part thereof, or seed of the first, second or third aspects of the invention is further defined by three features, namely Feature (i), Feature (ii) and Feature (iii), singly or in combination:

Feature (i) quantifies the extent of the increased level of the one or more non-polar lipids or the total non-polar lipid content of the vegetative plant part, or non-human organism or part thereof, or seed which may be expressed as the extent of increase on a weight basis (dry weight basis or seed weight basis), or as the relative increase compared to the level in the corresponding vegetative plant part, or non-human organism or part thereof, or seed. Feature (ii) specifies the plant genus or species, or the fungal or algal species, or other cell type, and Feature (iii) specifies one or more specific lipids that are increased in the non-polar lipid content.

For Feature (i), in an embodiment, the extent of the increase of the one or more non-polar lipids is at least 0.5%, at least 1%, at least 2%, at least 3%, at least 4%, at least 5%, at least 6%, at least 7%, at least 8%, at least 9%, at least 10%, at least 11%, at least 12%, at least 13%, at least 14%, at least 15%, at least 16%, at least 17%, at least 18%, at least 19%, at least 20%, at least 21%, at least 22%, at least 23%, at least 24%, at least 25% or at least 26% greater on a dry weight or seed weight basis than the corresponding vegetative plant part, or non-human organism or part thereof.

Also for Feature (i), in a preferred embodiment, the total non-polar lipid content of the vegetative plant part, or non-human organism or part thereof, or seed is increased when compared to the corresponding vegetative plant part, or non-human

organism or part thereof, or seed. In an embodiment, the total non-polar lipid content is increased by at least 0.5%, at least 1%, at least 2%, at least 3%, at least 4%, at least 5%, at least 6%, at least 7%, at least 8%, at least 9%, at least 10%, at least 11%, at least 12%, at least 13%, at least 14%, at least 15%, at least 16%, at least 17%, at least 18%, at least 19%, at least 20%, at least 21%, at least 22%, at least 23%, at least 24%, at least 25% or at least 26% greater on a dry weight or seed weight basis than the corresponding vegetative plant part, or non-human organism or part thereof, or seed.

Further, for Feature (i), in an embodiment, the level of the one or more non-polar lipids and/or the total non-polar lipid content is at least 1%, at least 2%, at least 3%, at least 4%, at least 5%, at least 6%, at least 7%, at least 8%, at least 9%, at least 10%, at least 11%, at least 12%, at least 13%, at least 14%, at least 15%, at least 16%, at least 17%, at least 18%, at least 19%, at least 20%, at least 21%, at least 22%, at least 23%, at least 24%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, or at least 100% greater on a relative basis than the corresponding vegetative plant part, or non-human organism or part thereof, or seed.

Also for Feature (i), the extent of increase in the level of the one or more non-polar lipids and/or the total non-polar lipid content may be at least 2-fold, at least 3-fold, at least 4-fold, at least 5-fold, at least 6-fold, at least 7-fold, at least 8-fold, at least 9-fold, at least 10-fold, or at least 12-fold, preferably at least about 13-fold or at least about 15-fold greater on a relative basis than the corresponding vegetative plant part, or non-human organism or part thereof, or seed.

As a result of the increase in the level of the one or more non-polar lipids and/or the total non-polar lipid content as defined in Feature (i), the total non-polar lipid content of the vegetative plant part, or non-human organism or part thereof, or seed is preferably between 5% and 25%, between 7% and 25%, between 10% and 25%, between 12% and 25%, between 15% and 25%, between 7% and 20%, between 10% and 20%, about 10%, about 11%, about 12%, about 13%, about 14%, about 15%, about 16%, about 17%, about 18%, about 20%, or about 22%, each as a percentage of dry weight or seed weight.

For Feature (ii), in an embodiment, the non-human organism is a plant, alga, or an organism suitable for fermentation such as a yeast or other fungus, preferably an oleaginous fungus such as an oleaginous yeast. The plant may be, or the vegetative plant part may be from, for example, a plant which is *Acrocomia aculeata* (macauba palm), *Arabidopsis thaliana*, *Aracinis hypogaea* (peanut), *Astrocaryum murumuru* (murumuru), *Astrocaryum vulgare* (tucumã), *Attalea geraensis* (Indaiá-rateiro), *Attalea humilis* (American oil palm), *Attalea oleifera* (andaiá), *Attalea phalerata* (uricuri), *Attalea speciosa* (babassu), *Avena sativa* (oats), *Beta vulgaris* (sugar beet),

Brassica sp. such as *Brassica carinata*, *Brassica juncea*, *Brassica napobrassica*,
Brassica napus (canola), *Camelina sativa* (false flax), *Cannabis sativa* (hemp),
Carthamus tinctorius (safflower), *Caryocar brasiliense* (pequi), *Cocos nucifera*
 (Coconut), *Crambe abyssinica* (Abyssinian kale), *Cucumis melo* (melon), *Elaeis*
 5 *guineensis* (African palm), *Glycine max* (soybean), *Gossypium hirsutum* (cotton),
Helianthus sp. such as *Helianthus annuus* (sunflower), *Hordeum vulgare* (barley),
Jatropha curcas (physic nut), *Joannesia princeps* (arara nut-tree), *Lemna* sp.
 (duckweed) such as *Lemna aequinoctialis*, *Lemna disperma*, *Lemna ecuadoriensis*,
Lemna gibba (swollen duckweed), *Lemna japonica*, *Lemna minor*, *Lemna minuta*,
 10 *Lemna obscura*, *Lemna paucicostata*, *Lemna perpusilla*, *Lemna tenera*, *Lemna*
trisulca, *Lemna turionifera*, *Lemna valdiviana*, *Lemna yungensis*, *Licania rigida*
 (oitica), *Linum usitatissimum* (flax), *Lupinus angustifolius* (lupin), *Mauritia*
flexuosa (buriti palm), *Maximiliana maripa* (inaja palm), *Miscanthus* sp. such as
Miscanthus x giganteus and *Miscanthus sinensis*, *Nicotiana* sp. (tobacco) such as
 15 *Nicotiana tabacum* or *Nicotiana benthamiana*, *Oenocarpus bacaba* (bacaba-do-
 azeite), *Oenocarpus bataua* (pataúã), *Oenocarpus distichus* (bacaba-de-leque), *Oryza*
sp. (rice) such as *Oryza sativa* and *Oryza glaberrima*, *Panicum virgatum*
 (switchgrass), *Paraqueiba paraensis* (mari), *Persea amencana* (avocado), *Pongamia*
pinnata (Indian beech), *Populus trichocarpa*, *Ricinus communis* (castor), *Saccharum*
 20 *sp.* (sugarcane), *Sesamum indicum* (sesame), *Solanum tuberosum* (potato), *Sorghum*
sp. such as *Sorghum bicolor*, *Sorghum vulgare*, *Theobroma grandiflorum* (cupuassu),
Trifolium sp., *Trithrinax brasiliensis* (Brazilian needle palm), *Triticum* sp. (wheat)
 such as *Triticum aestivum* and *Zea mays* (corn). In an embodiment, the *Brassica*
napus plant is of the variety Westar. In an alternative embodiment, if the plant is
 25 *Brassica napus*, it is of a variety or cultivar other than Westar. In an embodiment, the
 plant is of a species other than *Arabidopsis thaliana*. In another embodiment, the
 plant is of a species other than *Nicotiana tabacum*. In another embodiment, the plant
 is of a species other than *Nicotiana benthamiana*. In one embodiment, the plant is a
 perennial, for example, a switchgrass. Each of the features described for the plant of
 30 the third aspect can be applied *mutatis mutandis* to the vegetative plant part of the first
 and/or second aspect.

For Feature (iii), TAG, DAG, TAG and DAG, MAG, total polyunsaturated
 fatty acid (PUFA), or a specific PUFA such as eicosadienoic acid (EDA), arachidonic
 acid (ARA), alpha linolenic acid (ALA), stearidonic acid (SDA), eicosatrienoic acid
 35 (ETE), eicosatetraenoic acid (ETA), eicosapentaenoic acid (EPA), docosapentaenoic
 acid (DPA), docosahexaenoic acid (DHA), or a fatty acid which comprises a hydroxyl
 group, an epoxy group, a cyclopropane group, a double carbon-carbon bond, a triple
 carbon-carbon bond, conjugated double bonds, a branched chain such as a methylated

or hydroxylated branched chain, or a combination of two or more thereof, or any of two, three, four, five or six of the aforementioned groups, bonds or branched chains, is/are increased or decreased. The extent of the increase of TAG, DAG, TAG and DAG, MAG, PUFA, specific PUFA, or fatty acid, is as defined in Feature (i) above.

5 In a preferred embodiment, the MAG is 2-MAG. Preferably, DAG and/or TAG, more preferably the total of DAG and TAG, or MAG and TAG, are increased. In an embodiment, TAG levels are increased without increasing the MAG and/or DAG content.

Also for Feature (iii), in an embodiment, the total fatty acid content and/or TAG content of the total non-polar lipid content comprises (a) at least 2% more, preferably at least 5% more, more preferably at least 7% more, most preferably at least 10% more, at least 15% more, at least 20% more, at least 25% more oleic acid, or at least 30% more relative to the non-polar lipid(s) in the corresponding vegetative plant part, or non-human organism or part thereof, or seed lacking the one or more exogenous polynucleotides. In an embodiment, the total fatty acid content in the non-polar lipid(s) comprises (b) at least 2% less, preferably at least 4% less, more preferably at least 7% less, at least 10% less, at least 15% less, or at least 20% less palmitic acid relative to the non-polar lipid(s) in the corresponding vegetative plant part, or non-human organism or part thereof, or seed lacking the one or more exogenous polynucleotides. In an embodiment, the total fatty acid content of the total non-polar lipid content comprises (c) at least 2% less, preferably at least 4% less, more preferably at least 7% less, at least 10% less, or at least 15% less ALA relative to the non-polar lipid(s) in the corresponding vegetative plant part, or non-human organism or part thereof, or seed lacking the one or more exogenous polynucleotides. In an embodiment, the total fatty acid content of the total non-polar lipid content comprises (d) at least 2% more, preferably at least 5% more, more preferably at least 7% more, most preferably at least 10% more, or at least 15% more, LA, relative to the non-polar lipid(s) in the corresponding vegetative plant part, or non-human organism or part thereof, or seed lacking the one or more exogenous polynucleotides. Most preferably, the total fatty acid and/or TAG content of the total non-polar lipid content has an increased oleic acid level according to a figure defined in (a) and a decreased palmitic acid content according to a figure defined in (b). In an embodiment, the total sterol content is increased by at least 10% relative to seedoil from a corresponding seed. In an embodiment, the extracted lipid or oil comprises at least 10ppm chlorophyll, preferably at least 30ppm chlorophyll. The chlorophyll may subsequently be removed by de-colourising the extracted lipid or oil.

In preferred embodiments, the one or more non-polar lipids and/or the total non-polar lipid content is defined by the combination of Features (i), (ii) and (iii), or Features (i) and (ii), or Features (i) and (iii), or Features (ii) and (iii).

The process of the third aspect of the invention provides, in an embodiment, that one or more or all of the following features apply:

(i) the level of one or more non-polar lipids in the vegetative plant part, or non-human organism or part thereof, or seed is at least 0.5% greater on a weight basis than the level in a corresponding vegetative plant part, non-human organism or part thereof, or seed, respectively, lacking the one or more exogenous polynucleotide(s), or preferably as further defined in Feature (i),

(ii) the level of one or more non-polar lipids in the vegetative plant part, non-human organism or part thereof, or seed is at least 1% greater on a relative basis than in a corresponding vegetative plant part, non-human organism or part thereof, or seed, respectively, lacking the one or more exogenous polynucleotide(s), or preferably as further defined in Feature (i),

(iii) the total non-polar lipid content in the vegetative plant part, non-human organism or part thereof, or seed is at least 0.5% greater on a weight basis than the level in a corresponding vegetative plant part, non-human organism or part thereof, or seed, respectively, lacking the one or more exogenous polynucleotide(s), or preferably as further defined in Feature (i),

(iv) the total non-polar lipid content in the vegetative plant part, non-human organism or part thereof, or seed is at least 1% greater on a relative basis than in a corresponding vegetative plant part, non-human organism or part thereof, or seed, respectively, lacking the one or more exogenous polynucleotide(s), or preferably as further defined in Feature (i),

(v) the level of one or more non-polar lipids and/or the total non-polar lipid content of the vegetative plant part, non-human organism or part thereof, or seed, is at least 0.5% greater on a weight basis and/or at least 1% greater on a relative basis than a corresponding vegetative plant part, non-human organism or a part thereof, or seed, respectively, which is lacking the one or more exogenous polynucleotides and which comprises an exogenous polynucleotide encoding an *Arabidopsis thaliana* DGAT1, or preferably as further defined in Feature (i),

(vi) the TAG, DAG, TAG and DAG, or MAG content in the lipid in the vegetative plant part, non-human organism or part thereof, or seed, and/or in the extracted lipid therefrom, is at least 10% greater on a relative basis than the TAG, DAG, TAG and DAG, or MAG content in the lipid in a corresponding vegetative plant part, non-human organism or a part thereof, or seed lacking the one or more

exogenous polynucleotide(s), or a corresponding extracted lipid therefrom, respectively, or preferably as further defined in Feature (i), and

(vii) the total polyunsaturated fatty acid (PUFA) content in the lipid in the vegetative plant part, non-human organism or part thereof, or seed and/or in the extracted lipid therefrom, is increased (e.g., in the presence of a MGAT) or decreased (e.g., in the absence of a MGAT) relative to the total PUFA content in the lipid in a corresponding vegetative plant part, non-human organism or part thereof, or seed lacking the one or more exogenous polynucleotide(s), or a corresponding extracted lipid therefrom, respectively, or preferably as further defined in Feature (i) or Feature (iii).

In an embodiment, the level of a PUFA in the vegetative plant part, non-human organism or part thereof, or seed and/or the extracted lipid therefrom, is increased relative to the level of the PUFA in a corresponding vegetative plant part, non-human organism or part thereof, or seed, or a corresponding extracted lipid therefrom, respectively, wherein the polyunsaturated fatty acid is eicosadienoic acid, arachidonic acid (ARA), alpha linolenic acid (ALA), stearidonic acid (SDA), eicosatrienoic acid (ETE), eicosatetraenoic acid (ETA), eicosapentaenoic acid (EPA), docosapentaenoic acid (DPA), docosahexaenoic acid (DHA), or a combination of two or more thereof. Preferably, the extent of the increase is as defined in Feature (i).

In an embodiment of the third aspect, the corresponding vegetative plant part, or non-human organism or part thereof, or seed is a non-transgenic vegetative plant part, or non-human organism or part thereof, or seed, respectively. In a preferred embodiment, the corresponding vegetative plant part, or non-human organism or part thereof, or seed is of the same cultivar, strain or variety but lacking the one or more exogenous polynucleotides. In a further preferred embodiment, the corresponding vegetative plant part, or non-human organism or part thereof, or seed is at the same developmental stage, for example, flowering, as the vegetative plant part, or non-human organism or part thereof, or seed. In another embodiment, the vegetative plant parts are harvested from about at the time of flowering to about the beginning of senescence. In another embodiment, the seed is harvested when the plants are at least about 1 month of age.

In an embodiment, part of the non-human organism is seed and the total oil content, or the total fatty acid content, of the seed is at least 0.5% to 25%, or at least 1.0% to 24%, greater on a weight basis than a corresponding seed lacking the one or more exogenous polynucleotides.

In an embodiment, the relative DAG content of the seedoil is at least 10%, at least 10.5%, at least 11%, at least 11.5%, at least 12%, at least 12.5%, at least 13%, at least 13.5%, at least 14%, at least 14.5%, at least 15%, at least 15.5%, at least 16%, at

least 16.5%, at least 17%, at least 17.5%, at least 18%, at least 18.5%, at least 19%, at least 19.5%, at least 20% greater on a relative basis than of seedoil from a corresponding seed. In an embodiment, the DAG content of the seed is increased by an amount as defined in Feature (i) and the seed is from a genus and/or species as defined in Feature (ii).

In an embodiment, the relative TAG content of the seed is at least 5%, at least 5.5%, at least 6%, at least 6.5%, at least 7%, at least 7.5%, at least 8%, at least 8.5%, at least 9%, at least 9.5%, at least 10%, or at least 11% greater on an absolute basis relative to a corresponding seed. In an embodiment, the TAG content of the seed is increased by an amount as defined in Feature (i) and the seed is from a genus and/or species as defined in Feature (ii).

In another embodiment, the part of the non-human organism is a vegetative plant part and the TAG, DAG, TAG and DAG, or MAG content of the vegetative plant part is at least 10%, at least 11%, at least 12%, at least 13%, at least 14%, at least 15%, at least 16%, at least 17%, at least 18%, at least 19%, at least 20%, at least 21%, at least 22%, at least 23%, at least 24%, at least 25%, at least 30% at least 35%, at least 40%, at least 45%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, or at least 100% greater on a relative basis than the TAG, DAG, TAG and DAG, or MAG content of a corresponding vegetative plant part lacking the one or more exogenous polynucleotides. In a preferred embodiment, the MAG is 2-MAG. In an embodiment, the TAG, DAG, TAG and DAG, or MAG content of the vegetative plant part is determined from the amount of these lipid components in the extractable lipid of the vegetative plant part. In a further embodiment, the TAG, DAG, TAG and DAG, or MAG content of the transgenic vegetative plant part is increased by an amount as defined in Feature (i).

In an embodiment, at least 20% (mol%), at least 22% (mol%), at least 30% (mol%), at least 40% (mol%), at least 50% (mol%) or at least 60% (mol%), preferably at least 65% (mol%), more preferably at least 66% (mol%), at least 67% (mol%), at least 68% (mol%), at least 69% (mol%) or at least 70% (mol%) of the fatty acid content of the total non-polar lipid content of the vegetative plant part, non-human organism or part thereof, or seed, or of the lipid or oil extracted therefrom, preferably of the TAG fraction, is oleic acid. Such high oleic contents are preferred for use in biodiesel applications.

In another embodiment, the PUFA content of the vegetative plant part, or non-human organism or part thereof, or seed is increased (e.g., in the presence of a MGAT) or decreased (e.g., in the absence of a MGAT) when compared to the corresponding vegetative plant part, or non-human organism or part thereof, or seed. In this context, the PUFA content includes both esterified PUFA (including TAG,

DAG, etc.) and non-esterified PUFA. In an embodiment, the PUFA content of the vegetative plant part, or non-human organism or part thereof, or seed is preferably determined from the amount of PUFA in the extractable lipid of the vegetative plant part, or non-human organism or part thereof, or seed. The extent of the increase in PUFA content may be as defined in Feature (i). The PUFA content may comprise EDA, ARA, ALA, SDA, ETE, ETA, EPA, DPA, DHA, or a combination of two of more thereof.

In another embodiment, the level of a PUFA in the vegetative plant part, non-human organism or part thereof, or seed, or the lipid or oil extracted therefrom is increased or decreased when compared to the corresponding vegetative plant part, non-human organism or part thereof, or seed, or the lipid or oil extracted therefrom. The PUFA may be EDA, ARA, ALA, SDA, ETE, ETA, EPA, DPA, DHA, or a combination of two of more thereof. The extent of the increase in the PUFA may be as defined in Feature (i).

In another embodiment, the level of a fatty acid in the extracted lipid or oil is increased when compared to the lipid extracted from the corresponding vegetative plant part, or non-human organism or part thereof, or seed and wherein the fatty acid comprises a hydroxyl group, an epoxy group, a cyclopropane group, a double carbon-carbon bond, a triple carbon-carbon bond, conjugated double bonds, a branched chain such as a methylated or hydroxylated branched chain, or a combination of two or more thereof, or any of two, three, four, five or six of the aforementioned groups, bonds or branched chains. The extent of the increase in the fatty acid may be as defined in Feature (i).

In an embodiment, the level of the one or more non-polar lipids (such as TAG, DAG, TAG and DAG, MAG, PUFA, or a specific PUFA, or a specific fatty acid) and/or the total non-polar lipid content is determinable by analysis by using gas chromatography of fatty acid methyl esters obtained from the extracted lipid. Alternate methods for determining any of these contents are known in the art, and include methods which do not require extraction of lipid from the organism or part thereof, for example, analysis by near infrared (NIR) or nuclear magnetic resonance (NMR).

In a further embodiment, the level of the one or more non-polar lipids and/or the total non-polar lipid content of the vegetative plant part, or non-human organism or part thereof, or seed is at least 0.5% greater on a dry weight or seed weight basis and/or at least 1% greater on a relative basis, preferably at least 1% or 2% greater on a dry weight or seed weight basis, than a corresponding vegetative plant part, or non-human organism or a part thereof, or seed lacking the one or more exogenous

polynucleotides but comprising an exogenous polynucleotide encoding an *Arabidopsis thaliana* DGAT1 (SEQ ID NO:83).

In yet a further embodiment, the vegetative plant part or the non-human organism or part thereof, or seed further comprises (i) one or more introduced mutations, and/or (ii) an exogenous polynucleotide which down-regulate the production and/or activity of an endogenous enzyme of the vegetative plant part or the non-human organism or part thereof, the endogenous enzyme being selected from a fatty acid acyltransferase such as DGAT, *sn*-1 glycerol-3-phosphate acyltransferase (*sn*-1 GPAT), 1-acyl-glycerol-3-phosphate acyltransferase (LPAAT), acyl-CoA:lysophosphatidylcholine acyltransferase (LPCAT), phosphatidic acid phosphatase (PAP), an enzyme involved in starch biosynthesis such as (ADP)-glucose pyrophosphorylase (AGPase), a fatty acid desaturase such as a Δ 12 fatty acid desaturase (FAD2), a polypeptide involved in the degradation of lipid and/or which reduces lipid content such as a lipase such as CGi58 polypeptide or SUGAR-DEPENDENT1 triacylglycerol lipase, or a combination of two or more thereof. In an alternative embodiment, the vegetative plant part or the non-human organism or part thereof does not comprise (i) above, or does not comprise (ii) above, or does not comprise (i) above and does not comprise (ii) above. In an embodiment, the exogenous polynucleotide which down-regulates the production of AGPase is not the polynucleotide disclosed in Sanjaya et al. (2011). In an embodiment, the exogenous polynucleotides in the vegetative plant part or the non-human organism or part thereof, or seed does not consist of an exogenous polynucleotide encoding a WRI1 and an exogenous polynucleotide encoding an RNA molecule which inhibits expression of a gene encoding an AGPase.

In the process of either the first, second or third aspects, the vegetative plant part, or non-human organism or part thereof, or seed, or the extracted lipid or oil, is further defined in preferred embodiments. Therefore, in an embodiment one or more or all of the following features apply

(i) oleic acid comprises at least 20% (mol%), at least 22% (mol%), at least 30% (mol%), at least 40% (mol%), at least 50% (mol%), or at least 60% (mol%), preferably at least 65% (mol%) or at least 66% (mol%) of the total fatty acid content in the non-polar lipid or oil in the vegetative plant part, non-human organism or part thereof, or seed,

ii) oleic acid comprises at least 20% (mol%), at least 22% (mol%), at least 30% (mol%), at least 40% (mol%), at least 50% (mol%), or at least 60% (mol%), preferably at least 65% (mol%) or at least 66% (mol%) of the total fatty acid content in the extracted lipid or oil,

(iii) the non-polar lipid or oil in the vegetative plant part, non-human organism or part thereof, or seed comprises a fatty acid which comprises a hydroxyl group, an epoxy group, a cyclopropane group, a double carbon-carbon bond, a triple carbon-carbon bond, conjugated double bonds, a branched chain such as a methylated or hydroxylated branched chain, or a combination of two or more thereof, or any of two, three, four, five or six of the aforementioned groups, bonds or branched chains, and

(iv) the extracted lipid or oil comprises a fatty acid which comprises a hydroxyl group, an epoxy group, a cyclopropane group, a double carbon-carbon bond, a triple carbon-carbon bond, conjugated double bonds, a branched chain such as a methylated or hydroxylated branched chain, or a combination of two or more thereof, or any of two, three, four, five or six of the aforementioned groups, bonds or branched chains. The fatty acid composition in this embodiment is measured prior to any modification of the fatty acid composition, such as, for example, by fractionating the extracted lipid or oil to alter the fatty acid composition. In preferred embodiments, the extent of the increase is as defined in Feature (i).

In an embodiment, the level of a lipid in the vegetative plant part, non-human organism or part thereof, or seed and/or in the extracted lipid or oil is determinable by analysis by using gas chromatography of fatty acid methyl esters prepared from the extracted lipid or oil. The method of analysis is preferably as described in Example 1 herein.

Again with respect to either the first, second or third aspects, the invention provides for one or more exogenous polynucleotides in the vegetative plant part, or non-human organism or part thereof, or seed used in the process. Therefore, in an embodiment, the vegetative plant part, or the non-human organism or part thereof, or the seed comprises a first exogenous polynucleotide which encodes an RNA or preferably a transcription factor polypeptide that increases the expression of one or more glycolytic or fatty acid biosynthetic genes in a vegetative plant part, or a non-human organism or a part thereof, or a seed, respectively, and a second exogenous polynucleotide which encodes an RNA or a polypeptide involved in biosynthesis of one or more non-polar lipids, wherein the first and second exogenous polynucleotides are each operably linked to a promoter which is capable of directing expression of the polynucleotide in a vegetative plant part, or a non-human organism or a part thereof, or a seed, respectively. That is, the first and second exogenous polynucleotides encode different factors which together provide for the increase in the non-polar lipid content in the vegetative plant part, or the non-human organism or part thereof, or the seed.

The increase is preferably additive, more preferably synergistic, relative to the presence of either the first or second exogenous polynucleotide alone. The factors

encoded by the first and second polynucleotides operate by different mechanisms. Preferably, the transcription factor polypeptide increases the availability of substrates for non-polar lipid synthesis, such as, for example, increasing glycerol-3-phosphate and/or fatty acids preferably in the form of acyl-CoA, by increasing expression of genes, for example at least 5 or at least 8 genes, involved in glycolysis or fatty acid biosynthesis (such as, but not limited to, one or more of ACCase, sucrose transporters (SuSy, cell wall invertases), ketoacyl synthase (KAS), phosphofructokinase (PFK), pyruvate kinase (PK) (for example, (At5g52920, At3g22960), pyruvate dehydrogenase, hexose transporters (for example, GPT2 and PPT1), cytosolic fructokinase, cytosolic phosphoglycerate mutase, enoyl-ACP reductase (At2g05990), and phosphoglycerate mutase (At1g22170)) preferably more than one gene for each category. In an embodiment, the first exogenous polynucleotide encodes a Wrinkled 1 (WRI1) transcription factor, a Leafy Cotyledon 1 (Lec1) transcription factor, a Leafy Cotyledon 2 (LEC2) transcription factor, a Fus3 transcription factor, an ABI3 transcription factor, a Dof4 transcription factor, a BABY BOOM (BBM) transcription factor or a Dof11 transcription factor. In one embodiment, the LEC2 is not an *Arabidopsis* LEC2. As part of this embodiment, or separately, the second exogenous polynucleotide may encode a polypeptide having a fatty acid acyltransferase activity, for example, monoacylglycerol acyltransferase (MGAT) activity and/or diacylglycerol acyltransferase (DGAT) activity, or glycerol-3-phosphate acyltransferase (GPAT) activity. In one embodiment, the DGAT is not an *Arabidopsis* DGAT.

In a preferred embodiment, the vegetative plant part, or non-human organism or a part thereof, or the seed, of the first, second or third aspects of the invention comprises two or more exogenous polynucleotide(s), one of which encodes a transcription factor polypeptide that increases the expression of one or more glycolytic or fatty acid biosynthetic genes in the vegetative plant part, or non-human organism or a part thereof, or seed such as a Wrinkled 1 (WRI1) transcription factor, and a second of which encodes a polypeptide involved in biosynthesis of one or more non-polar lipids such as a DGAT.

In an embodiment, the vegetative plant part, non-human organism or a part thereof, or the seed of the first, second or third aspects of the invention may further comprise a third, or more, exogenous polynucleotide(s). The third, or more, exogenous polynucleotide(s) may encode one or more or any combination of:

i) a further RNA or transcription factor polypeptide that increases the expression of one or more glycolytic or fatty acid biosynthetic genes in a non-human organism or a part thereof (for example, if the first exogenous polynucleotide encodes a Wrinkled 1 (WRI1) transcription factor, the third exogenous polynucleotide may

encode a LEC2 or BBM transcription factor (preferably, LEC2 or BBM expression controlled by an inducible promoter or a promoter which does not result in high transgene expression levels),

ii) a further RNA or polypeptide involved in biosynthesis of one or more non-polar lipids (for example, if the second exogenous polynucleotide encodes a DGAT, the third exogenous polynucleotide may encode a MGAT or GPAT, or two further exogenous polynucleotides may be present encoding an MGAT and a GPAT),

iii) a polypeptide that stabilizes the one or more non-polar lipids, preferably an oleosin, such as a polyoleosin or a caleosin, more preferably a polyoleosin,

iv) an RNA molecule which inhibits expression of a gene encoding a polypeptide involved in starch biosynthesis such as a AGPase polypeptide,

v) an RNA molecule which inhibits expression of a gene encoding a polypeptide involved in the degradation of lipid and/or which reduces lipid content such as a lipase such as CGi58 polypeptide or SUGAR-DEPENDENT1 triacylglycerol lipase, or

vi) a silencing suppressor polypeptide,

wherein the third, or more, exogenous polynucleotide(s) is operably linked to a promoter which is capable of directing expression of the polynucleotide(s) in a vegetative plant part, or a non-human organism or a part thereof, or a seed, respectively.

A number of specific combinations of genes are shown herein to be effective for increasing non-polar lipid contents. Therefore, regarding the process of either the first, second or third aspects of the invention, in an embodiment, the vegetative plant part, or the non-human organism or part thereof, or the seed comprises one or more exogenous polynucleotide(s) which encode:

i) a Wrinkled 1 (WRI1) transcription factor and a DGAT,

ii) a WRI1 transcription factor and a DGAT and an Oleosin,

iii) a WRI1 transcription factor, a DGAT, a MGAT and an Oleosin,

iv) a monoacylglycerol acyltransferase (MGAT),

v) a diacylglycerol acyltransferase 2 (DGAT2),

vi) a MGAT and a glycerol-3-phosphate acyltransferase (GPAT),

vii) a MGAT and a DGAT,

viii) a MGAT, a GPAT and a DGAT,

ix) a WRI1 transcription factor and a MGAT,

x) a WRI1 transcription factor, a DGAT and a MGAT,

xi) a WRI1 transcription factor, a DGAT, a MGAT, an Oleosin and a GPAT,

xii) a DGAT and an Oleosin, or

xiii) a MGAT and an Oleosin, and

xiv) optionally, a silencing suppressor polypeptide,

wherein each of the one or more exogenous polynucleotide(s) is operably linked to a promoter which is capable of directing expression of the polynucleotide in a vegetative plant part, or a non-human organism or part thereof, or seed, respectively.

5 Preferably the one or more exogenous polynucleotides are stably integrated into the genome of the vegetative plant part, or the non-human organism or part thereof, or the seed, and more preferably are present in a homozygous state. The polynucleotide may encode an enzyme having an amino acid sequence which is the same as a sequence of a naturally occurring enzyme of, for example, plant, yeast or animal origin. Further,
10 the polynucleotide may encode an enzyme having one or more conservative mutations when compared to the naturally occurring enzyme.

In an embodiment,

(i) the GPAT also has phosphatase activity to produce MAG, such as a polypeptide having an amino acid sequence of *Arabidopsis* GPAT4 or GPAT6, and/or

15 (ii) the DGAT is a DGAT1 or a DGAT2, and/or

(iii) the MGAT is an MGAT1 or an MGAT2.

In a preferred embodiment, the vegetative plant part, the non-human organism or part thereof, or the seed comprises a first exogenous polynucleotide encoding a WRI1 and a second exogenous polynucleotide encoding a DGAT, preferably a
20 DGAT1.

In another preferred embodiment, the vegetative plant part, the non-human organism or part thereof, or the seed comprises a first exogenous polynucleotide encoding a WRI1, a second exogenous polynucleotide encoding a DGAT, preferably a DGAT1, and a third exogenous polynucleotide encoding an oleosin.

25 In a further embodiment, the vegetative plant part, the non-human organism or part thereof, or the seed comprises a first exogenous polynucleotide encoding a WRI1, a second exogenous polynucleotide encoding a DGAT, preferably a DGAT1, a third exogenous polynucleotide encoding an oleosin, and a fourth exogenous polynucleotide encoding an MGAT, preferably an MGAT2.

30 In a further embodiment, the vegetative plant part, the non-human organism or part thereof, or the seed comprises a first exogenous polynucleotide encoding a WRI1, a second exogenous polynucleotide encoding a DGAT, preferably a DGAT1, a third exogenous polynucleotide encoding an oleosin, and a fourth exogenous polynucleotide encoding LEC2 or BBM.

35 In a further embodiment, the vegetative plant part, the non-human organism or part thereof, or the seed comprises a first exogenous polynucleotide encoding a WRI1, a second exogenous polynucleotide encoding a DGAT, preferably a DGAT1, a third exogenous polynucleotide encoding an oleosin, a fourth exogenous polynucleotide

encoding an MGAT, preferably an MGAT2, and a fifth exogenous polynucleotide encoding LEC2 or BBM.

In a further embodiment, the vegetative plant part, the non-human organism or part thereof, or the seed comprises a first exogenous polynucleotide encoding a WRI1, a second exogenous polynucleotide encoding a DGAT, preferably a DGAT1, a third exogenous polynucleotide encoding an oleosin, and a fourth exogenous polynucleotide encoding an RNA molecule which inhibits expression of a gene encoding a lipase such as CGi58 polypeptide.

In a further embodiment, the vegetative plant part, the non-human organism or part thereof, or the seed comprises a first exogenous polynucleotide encoding a WRI1, a second exogenous polynucleotide encoding a DGAT, preferably a DGAT1, a third exogenous polynucleotide encoding an oleosin, a fourth exogenous polynucleotide encoding an RNA molecule which inhibits expression of a gene encoding a lipase such as a CGi58 polypeptide, and a fifth exogenous polynucleotide encoding LEC2 or BBM.

In a further embodiment, the vegetative plant part, the non-human organism or part thereof, or the seed comprises a first exogenous polynucleotide encoding a WRI1, a second exogenous polynucleotide encoding a DGAT, preferably a DGAT1, a third exogenous polynucleotide encoding an oleosin, a fourth exogenous polynucleotide encoding an RNA molecule which inhibits expression of a gene encoding a lipase such as a CGi58 polypeptide, and a fifth exogenous polynucleotide encoding an MGAT, preferably an MGAT2.

In a further embodiment, the vegetative plant part, the non-human organism or part thereof, or the seed comprises a first exogenous polynucleotide encoding a WRI1, a second exogenous polynucleotide encoding a DGAT, preferably a DGAT1, a third exogenous polynucleotide encoding an oleosin, a fourth exogenous polynucleotide encoding an RNA molecule which inhibits expression of a gene encoding a lipase such as a CGi58 polypeptide, a fifth exogenous polynucleotide encoding an MGAT, preferably an MGAT2, and a sixth exogenous polynucleotide encoding LEC2 or BBM.

In an embodiment, the seed comprises a first exogenous polynucleotide encoding a WRI1, a second exogenous polynucleotide encoding a DGAT, preferably a DGAT1, a third exogenous polynucleotide encoding an oleosin, and a fourth exogenous polynucleotide encoding an MGAT, preferably an MGAT2. Preferably, the seed further comprises a fifth exogenous polynucleotide encoding a GPAT.

Where relevant, instead of a polynucleotide encoding an RNA molecule which inhibits expression of a gene encoding a lipase such as a CGi58 polypeptide, the vegetative plant part, the non-human organism or part thereof, or the seed has one or

more introduced mutations in the lipase gene such as a CGi58 gene which confers reduced levels of the lipase polypeptide when compared to a corresponding vegetative plant part, non-human organism or part thereof, or seed lacking the mutation.

In a preferred embodiment, the exogenous polynucleotides encoding the DGAT and oleosin are operably linked to a constitutive promoter, or a promoter active in green tissues of a plant at least before and up until flowering, which is capable of directing expression of the polynucleotides in the vegetative plant part, the non-human organism or part thereof, or the seed. In a further preferred embodiment, the exogenous polynucleotide encoding WRI1, and RNA molecule which inhibits expression of a gene encoding a lipase such as a CGi58 polypeptide, is operably linked to a constitutive promoter, a promoter active in green tissues of a plant at least before and up until flowering, or an inducible promoter, which is capable of directing expression of the polynucleotides in the vegetative plant part, the non-human organism or part thereof, or the seed. In yet a further preferred embodiment, the exogenous polynucleotides encoding LEC2, BBM and/or MGAT2 are operably linked to an inducible promoter which is capable of directing expression of the polynucleotides in the vegetative plant part, the non-human organism or part thereof, or the seed.

In each of the above embodiments, the polynucleotides may be provided as separate molecules or may be provided as a contiguous single molecule, such as on a single T-DNA molecule. In an embodiment, the orientation of transcription of at least one gene on the T-DNA molecule is opposite to the orientation of transcription of at least one other gene on the T-DNA molecule.

In each of the above embodiments, the total non-polar lipid content of the vegetative plant part, or non-human organism or part thereof, or the seed, preferably a plant leaf or part thereof, stem or tuber, is at least about 3%, more preferably at least about 5%, preferably at least about 7%, more preferably at least about 10%, more preferably at least about 11%, more preferably at least about 12%, more preferably at least about 13%, more preferably at least about 14%, or more preferably at least about 15% (w/w dry weight). In a further preferred embodiment, the total non-polar lipid content is between 5% and 25%, between 7% and 25%, between 10% and 25%, between 12% and 25%, between 15% and 25%, between 7% and 20%, between 10% and 20%, between 10% and 15%, between 15% and 20%, between 20% and 25%, about 10%, about 11%, about 12%, about 13%, about 14%, about 15%, about 16%, about 17%, about 18%, about 20%, or about 22%, each as a percentage of dry weight or seed weight. In a particularly preferred embodiment, the vegetative plant part is a leaf (or leaves) or a portion thereof. In a more preferred embodiment, the vegetative plant part is a leaf portion having a surface area of at least 1 cm².

Furthermore, in each of the above embodiments, the total TAG content of the vegetative plant part, or non-human organism or part thereof, or the seed, preferably a plant leaf or part thereof, stem or tuber, is at least about 3%, more preferably at least about 5%, preferably at least about 7%, more preferably at least about 10%, more preferably at least about 11%, more preferably at least about 12%, more preferably at least about 13%, more preferably at least about 14%, more preferably at least about 15%, or more preferably at least about 17% (w/w dry weight). In a further preferred embodiment, the total TAG content is between 5% and 30%, between 7% and 30%, between 10% and 30%, between 12% and 30%, between 15% and 30%, between 7% and 30%, between 10% and 30%, between 20% and 28%, between 18% and 25%, between 22% and 30%, about 10%, about 11%, about 12%, about 13%, about 14%, about 15%, about 16%, about 17%, about 18%, about 20%, or about 22%, each as a percentage of dry weight or seed weight. In a particularly preferred embodiment, the vegetative plant part is a leaf (or leaves) or a portion thereof. In a more preferred embodiment, the vegetative plant part is a leaf portion having a surface area of at least 1 cm².

Furthermore, in each of the above embodiments, the total lipid content of the vegetative plant part, or non-human organism or part thereof, or the seed, preferably a plant leaf or part thereof, stem or tuber, is at least about 3%, more preferably at least about 5%, preferably at least about 7%, more preferably at least about 10%, more preferably at least about 11%, more preferably at least about 12%, more preferably at least about 13%, more preferably at least about 14%, more preferably at least about 15%, more preferably at least about 17% (w/w dry weight), more preferably at least about 20%, more preferably at least about 25%. In a further preferred embodiment, the total lipid content is between 5% and 35%, between 7% and 35%, between 10% and 35%, between 12% and 35%, between 15% and 35%, between 7% and 35%, between 10% and 20%, between 18% and 28%, between 20% and 28%, between 22% and 28%, about 10%, about 11%, about 12%, about 13%, about 14%, about 15%, about 16%, about 17%, about 18%, about 20%, about 22%, or about 25%, each as a percentage of dry weight. Typically, the total lipid content of the vegetative plant part, or non-human organism or part thereof is about 2-3% higher than the non-polar lipid content. In a particularly preferred embodiment, the vegetative plant part is a leaf (or leaves) or a portion thereof. In a more preferred embodiment, the vegetative plant part is a leaf portion having a surface area of at least 1 cm².

In an embodiment, the vegetative plant part, the non-human organism or part thereof, or the seed, preferably the vegetative plant part, comprises a first exogenous polynucleotide encoding a WRI1, a second exogenous polynucleotide encoding a DGAT, preferably a DGAT1, a third exogenous polynucleotide encoding an MGAT,

preferably an MGAT2, and a fourth exogenous polynucleotide encoding an oleosin, wherein the vegetative plant part, non-human organism or part thereof, or seed has one or more or all of the following features:

i) a total lipid content of at least 8%, at least 10%, at least 12%, at least 14%, or at least 15.5% (% weight),

ii) at least a 3 fold, at least a 5 fold, at least a 7 fold, at least an 8 fold, or least a 10 fold, at higher total lipid content in the vegetative plant part or non-human organism relative to a corresponding vegetative plant part or non-human organism lacking the exogenous polynucleotides,

iii) a total TAG content of at least 5%, at least 6%, at least 6.5% or at least 7% (% weight of dry weight or seed weight),

iv) at least a 40 fold, at least a 50 fold, at least a 60 fold, or at least a 70 fold, or at least a 100 fold, higher total TAG content relative to a corresponding vegetative plant part or non-human organism lacking the exogenous polynucleotides,

v) oleic acid comprises at least 15%, at least 19% or at least 22% (% weight) of the fatty acids in TAG,

vi) at least a 10 fold, at least a 15 fold or at least a 17 fold higher level of oleic acid in TAG relative to a corresponding vegetative plant part or non-human organism lacking the exogenous polynucleotides,

vii) palmitic acid comprises at least 20%, at least 25%, at least 30% or at least 33% (% weight) of the fatty acids in TAG,

viii) at least a 1.5 fold higher level of palmitic acid in TAG relative to a corresponding vegetative plant part or non-human organism lacking the exogenous polynucleotides,

ix) linoleic acid comprises at least 22%, at least 25%, at least 30% or at least 34% (% weight) of the fatty acids in TAG,

x) α -linolenic acid comprises less than 20%, less than 15%, less than 11% or less than 8% (% weight) of the fatty acids in TAG, and

xi) at least a 5 fold, or at least an 8 fold, lower level of α -linolenic acid in TAG relative to a corresponding vegetative plant part or non human organism lacking the exogenous polynucleotides. In this embodiment, preferably the vegetative plant part at least has feature(s), i), ii) iii), iv), i) and ii), i) and iii), i) and iv), i) to iii), i), iii) and iv), i) to iv), ii) and iii), ii) and iv), ii) to iv), or iii) and iv). In an embodiment, % dry weight is % leaf dry weight.

In a further embodiment, the vegetative plant part, the non-human organism or part thereof, or the seed, preferably the vegetative plant part, comprises a first exogenous polynucleotide encoding a WRI1, a second exogenous polynucleotide encoding a DGAT, preferably a DGAT1, a third exogenous polynucleotide encoding

an oleosin, wherein the vegetative plant part, non-human organism or part thereof, or seed has one or more or all of the following features:

i) a total TAG content of at least 10%, at least 12.5%, at least 15% or at least 17% (% weight of dry weight or seed weight),

ii) at least a 40 fold, at least a 50 fold, at least a 60 fold, or at least a 70 fold, or at least a 100 fold, higher total TAG content in the vegetative plant part or non-human organism relative to a corresponding vegetative plant part or non human organism lacking the exogenous polynucleotides,

iii) oleic acid comprises at least 19%, at least 22%, or at least 25% (% weight) of the fatty acids in TAG,

iv) at least a 10 fold, at least a 15 fold, at least a 17 fold, or at least a 19 fold, higher level of oleic acid in TAG in the vegetative plant part or non-human organism relative to a corresponding vegetative plant part or non-human organism lacking the exogenous polynucleotides,

v) palmitic acid comprises at least 20%, at least 25%, or at least 28% (% weight) of the fatty acids in TAG,

vi) at least a 1.25 fold higher level of palmitic acid in TAG in the vegetative plant part or non-human organism relative to a corresponding vegetative plant part or non-human organism lacking the exogenous polynucleotides,

vii) linoleic acid comprises at least 15%, or at least 20%, (% weight) of the fatty acids in TAG,

viii) α -linolenic acid comprises less than 15%, less than 11% or less than 8% (% weight) of the fatty acids in TAG, and

ix) at least a 5 fold, or at least an 8 fold, lower level of α -linolenic acid in TAG in the vegetative plant part or non-human organism relative to a corresponding vegetative plant part or non-human organism lacking the exogenous polynucleotides.

In this embodiment, preferably the vegetative plant part at least has feature(s), i), ii), or i) and ii). In an embodiment, % dry weight is % leaf dry weight.

Preferably, the defined features for the two above embodiments are as at the flowering stage of the plant.

In an alternate embodiment, the vegetative plant part, the non-human organism or part thereof, or the seed consists of one or more exogenous polynucleotides encoding a DGAT1 and a LEC2.

In a preferred embodiment, the exogenous polynucleotide encoding WRI1 comprises one or more of the following:

i) nucleotides whose sequence is set forth as any one of SEQ ID NOs:231 to 278,

ii) nucleotides encoding a polypeptide comprising amino acids whose sequence is set forth as any one of SEQ ID NOs:279 to 337, or a biologically active fragment thereof,

iii) nucleotides whose sequence is at least 30% identical to i) or ii), and

iv) nucleotides which hybridize to any one of i) to iii) under stringent conditions.

In a preferred embodiment, the exogenous polynucleotide encoding DGAT comprises one or more of the following:

i) nucleotides whose sequence is set forth as any one of SEQ ID NOs:204 to 211, 338 to 346,

ii) nucleotides encoding a polypeptide comprising amino acids whose sequence is set forth as any one of SEQ ID NOs:83, 212 to 219, 347 to 355, or a biologically active fragment thereof,

iii) nucleotides whose sequence is at least 30% identical to i) or ii), and

iv) a polynucleotide which hybridizes to any one of i) to iii) under stringent conditions.

In another preferred embodiment, the exogenous polynucleotide encoding MGAT comprises one or more of the following:

i) nucleotides whose sequence is set forth as any one of SEQ ID NOs:1 to 44,

ii) nucleotides encoding a polypeptide comprising amino acids whose sequence is set forth as any one of SEQ ID NOs:45 to 82, or a biologically active fragment thereof,

iii) nucleotides whose sequence is at least 30% identical to i) or ii), and

iv) a polynucleotide which hybridizes to any one of i) to iii) under stringent conditions.

In another preferred embodiment, the exogenous polynucleotide encoding GPAT comprises one or more of the following:

i) nucleotides whose sequence is set forth as any one of SEQ ID NOs:84 to 143,

ii) nucleotides encoding a polypeptide comprising amino acids whose sequence is set forth as any one of SEQ ID NOs:144 to 203, or a biologically active fragment thereof,

iii) nucleotides whose sequence is at least 30% identical to i) or ii), and

iv) a polynucleotide which hybridizes to any one of i) to iii) under stringent conditions.

In another preferred embodiment, the exogenous polynucleotide encoding DGAT2 comprises one or more of the following:

i) nucleotides whose sequence is set forth as any one of SEQ ID NOs:204 to 211,

ii) nucleotides encoding a polypeptide comprising amino acids whose sequence is set forth as any one of SEQ ID NOs:212 to 219, or a biologically active fragment thereof,

iii) nucleotides whose sequence is at least 30% identical to i) or ii), and

iv) a polynucleotide which hybridizes to any one of i) to iii) under stringent conditions.

In another preferred embodiment, the exogenous polynucleotide encoding an oleosin comprises one or more of the following:

i) nucleotides whose sequence is set forth as any one of SEQ ID NOs:389 to 408,

ii) nucleotides encoding a polypeptide comprising amino acids whose sequence is set forth as any one of SEQ ID NOs:362 to 388, or a biologically active fragment thereof,

iii) nucleotides whose sequence is at least 30% identical to i) or ii), and

iv) a sequence of nucleotides which hybridizes to any one of i) to iii) under stringent conditions.

In an embodiment, the CGi58 polypeptide comprises one or more of the following:

i) nucleotides whose sequence is set forth as any one of SEQ ID NOs:422 to 428,

ii) nucleotides encoding a polypeptide comprising amino acids whose sequence is set forth as any one of SEQ ID NOs:429 to 436, or a biologically active fragment thereof,

iii) nucleotides whose sequence is at least 30% identical to i) or ii), and

iv) a sequence of nucleotides which hybridizes to any one of i) to iii) under stringent conditions.

In another embodiment, the exogenous polynucleotide encoding LEC2 comprises one or more of the following:

i) nucleotides whose sequence is set forth as any one of SEQ ID NOs:437 to 439,

ii) nucleotides encoding a polypeptide comprising amino acids whose sequence is set forth as any one of SEQ ID NOs:442 to 444, or a biologically active fragment thereof,

iii) nucleotides whose sequence is at least 30% identical to i) or ii), and

iv) a sequence of nucleotides which hybridizes to any one of i) to iii) under stringent conditions.

In a further embodiment, the exogenous polynucleotide encoding BBM comprises one or more of the following:

i) nucleotides whose sequence is set forth as any one of SEQ ID NOs:440 or 441

ii) nucleotides encoding a polypeptide comprising amino acids whose sequence is set forth as any one of SEQ ID NOs:445 or 446, or a biologically active fragment thereof,

iii) nucleotides whose sequence is at least 30% identical to i) or ii), and

iv) a sequence of nucleotides which hybridizes to any one of i) to iii) under stringent conditions.

Clearly, sequences preferred in one embodiment can be combined with sequences preferred in another embodiment and more advantageously further combined with a sequence preferred in yet another embodiment.

In one embodiment, the one or more exogenous polynucleotides encode a mutant MGAT and/or DGAT and/or GPAT. For example, the one or more exogenous polynucleotides may encode a MGAT and/or DGAT and/or GPAT having one, or more than one, conservative amino acid substitutions as exemplified in Table 1 relative to a wildtype MGAT and/or DGAT and/or GPAT as defined by a SEQ ID NO herein. Preferably the mutant polypeptide has an equivalent or greater activity relative to the non-mutant polypeptide.

In an embodiment, the vegetative plant part, non-human organism or part thereof, or seed comprises a first exogenous polynucleotide that encodes a MGAT and a second exogenous polynucleotide that encodes a GPAT. The first and second polynucleotides may be provided as separate molecules or may be provided as a contiguous single molecule, such as on a single T-DNA molecule. In an embodiment, the orientation of transcription of at least one gene on the T-DNA molecule is opposite to the orientation of transcription of at least one other gene on the T-DNA molecule. In a preferred embodiment, the GPAT is a GPAT having phosphatase activity such as an *Arabidopsis* GPAT4 or GPAT6. The GPAT having phosphatase activity acts to catalyze the formation of MAG from G-3-P (i.e., acylates G-3-P to form LPA and subsequently removes a phosphate group to form MAG) in the non-human organism or part thereof. The MGAT then acts to catalyze the formation of DAG in the non-human organism or part thereof by acylating the MAG with an acyl group derived from fatty acyl-CoA. The MGAT such as *A. thaliana* MGAT1 may also act to catalyze the formation of TAG in the non-human organism or part thereof if it also has DGAT activity.

The vegetative plant part, non-human organism or part thereof, or seed may comprise a third exogenous polynucleotide encoding, for example, a DGAT. The

first, second and third polynucleotides may be provided as separate molecules or may be provided as a contiguous single molecule, such as on a single T-DNA molecule. The DGAT acts to catalyse the formation of TAG in the transgenic vegetative plant part, non-human organism or part thereof, or seed by acylating the DAG (preferably produced by the MGAT pathway) with an acyl group derived from fatty acyl-CoA. In an embodiment, the orientation of transcription of at least one gene on the T-DNA molecule is opposite to the orientation of transcription of at least one other gene on the T-DNA molecule.

In another embodiment, the vegetative plant part, non-human organism or part thereof, or seed comprises a first exogenous polynucleotide that encodes a MGAT and a second exogenous polynucleotide that encodes a DGAT. The first and second polynucleotides may be provided as separate molecules or may be provided as a contiguous single molecule, such as on a single T-DNA molecule. In an embodiment, the orientation of transcription of at least one gene on the T-DNA molecule is opposite to the orientation of transcription of at least one other gene on the T-DNA molecule. The vegetative plant part, non-human organism or part thereof, or seed may comprise a third exogenous polynucleotide encoding, for example, a GPAT, preferably a GPAT having phosphatase activity such as an *Arabidopsis* GPAT4 or GPAT6. The first, second and third polynucleotides may be provided as separate molecules or may be provided as a contiguous single molecule.

Furthermore, an endogenous gene activity in the plant, vegetative plant part, or the non-human organism or part thereof, or the seed may be down-regulated. Therefore, in an embodiment, the vegetative plant part, the non-human organism or part thereof, or the seed comprises one or more of:

(i) one or more introduced mutations in a gene which encodes an endogenous enzyme of the plant, vegetative plant part, non-human organism or part thereof, or seed, respectively, or

(ii) an exogenous polynucleotide which down-regulates the production and/or activity of an endogenous enzyme of the plant, vegetative plant part, non-human organism or part thereof, or seed, respectively,

wherein each endogenous enzyme is selected from the group consisting of a fatty acid acyltransferase such as DGAT, an *sn*-1 glycerol-3-phosphate acyltransferase (*sn*-1 GPAT), a 1-acyl-glycerol-3-phosphate acyltransferase (LPAAT), an acyl-CoA:lysophosphatidylcholine acyltransferase (LPCAT), a phosphatidic acid phosphatase (PAP), an enzyme involved in starch biosynthesis such as (ADP)-glucose pyrophosphorylase (AGPase), a fatty acid desaturase such as a Δ 12 fatty acid desaturase (FAD2), a polypeptide involved in the degradation of lipid and/or which reduces lipid content such as a lipase such as a CGI58 polypeptide or SUGAR-

DEPENDENT1 triacylglycerol lipase, or a combination of two or more thereof. In an embodiment, the exogenous polynucleotide is selected from the group consisting of an antisense polynucleotide, a sense polynucleotide, a catalytic polynucleotide, a microRNA, a polynucleotide which encodes a polypeptide which binds the endogenous enzyme, a double stranded RNA molecule or a processed RNA molecule derived therefrom. In an embodiment, the exogenous polynucleotide which down-regulates the production of AGPase is not the polynucleotide disclosed in Sanjaya et al. (2011). In an embodiment, the exogenous polynucleotides in the vegetative plant part or the non-human organism or part thereof, or seed does not consist of an exogenous polynucleotide encoding a WRI1 and an exogenous polynucleotide encoding an RNA molecule which inhibits expression of a gene encoding an AGPase.

Increasing the level of non-polar lipids is important for applications involving particular fatty acids. Therefore, in an embodiment, the total non-polar lipid, the extracted lipid or oil comprises:

- (i) non-polar lipid which is TAG, DAG, TAG and DAG, or MAG, and
- (ii) a specific PUFA which is EDA, ARA, SDA, ETE, ETA, EPA, DPA, DHA, the specific PUFA being at a level of at least 1% of the total fatty acid content in the non-polar lipid, or a combination of two or more of the specific PUFA, or
- (iii) a fatty acid which is present at a level of at least 1% of the total fatty acid content in the non-polar lipid and which comprises a hydroxyl group, an epoxy group, a cyclopropane group, a double carbon-carbon bond, a triple carbon-carbon bond, conjugated double bonds, a branched chain such as a methylated or hydroxylated branched chain, or a combination of two or more thereof, or any of two, three, four, five or six of the aforementioned groups, bonds or branched chains.

In a fourth aspect, the invention provides non-human organisms, preferably plants, or parts thereof such as vegetative plant parts or seed, which are useful in the processes of the first, second or third aspects or in further aspects described hereafter. Each of the features in the embodiments described for the first, second or third aspects can be applied *mutatis mutandis* to the non-human organisms, preferably plants, or parts thereof such as vegetative plant parts or seed of the fourth aspect. Particular embodiments are emphasized as follows.

In an embodiment of the fourth aspect, the invention provides a plant comprising a vegetative part, or the vegetative part thereof, wherein the vegetative part has a total non-polar lipid content of at least about 3%, more preferably at least about 5%, preferably at least about 7%, more preferably at least about 10%, more preferably at least about 11%, more preferably at least about 12%, more preferably at least about 13%, more preferably at least about 14%, or more preferably at least about 15% (w/w dry weight). In a further preferred embodiment, the total non-polar lipid

content is between 5% and 25%, between 7% and 25%, between 10% and 25%, between 12% and 25%, between 15% and 25%, between 7% and 20%, between 10% and 20%, about 10%, about 11%, about 12%, about 13%, about 14%, about 15%, about 16%, about 17%, about 18%, about 20%, or about 22%, each as a percentage of dry weight. In a particularly preferred embodiment, the vegetative plant part is a leaf (or leaves) or a portion thereof. In a more preferred embodiment, the vegetative plant part is a leaf portion having a surface area of at least 1 cm². In a further embodiment, the non-polar lipid comprises at least 90% triacylglycerols (TAG). Preferably the plant is fertile, morphologically normal, and/or agronomically useful. Seed of the plant preferably germinates at a rate substantially the same as for a corresponding wild-type plant. Preferably the vegetative part is a leaf or a stem, or a combination of the two, or a root or tuber such as, for example, potato tubers.

In another embodiment, the non-human organism, preferably plant, or part thereof such as vegetative plant part or seed comprises one or more exogenous polynucleotides as defined herein and has an increased level of the one or more non-polar lipids and/or the total non-polar lipid content which is at least 2-fold, at least 3-fold, at least 4-fold, at least 5-fold, at least 6-fold, at least 7-fold, at least 8-fold, at least 9-fold, at least 10-fold, or at least 12-fold, preferably at least about 13-fold or at least about 15-fold greater on a relative basis than a corresponding non-human organism, preferably plant, or part thereof such as vegetative plant part or seed lacking the one or more exogenous polynucleotides.

In an embodiment, the invention provides a canola plant comprising canola seed whose oil content is at least 45% on a weight basis. Preferably, the canola plant or its seed have features as described in the first, second or third aspects of the invention.

In an embodiment, the invention provides a corn plant comprising corn seed whose oil content is at least 5% on a weight basis. Preferably, the corn plant or its seed have features as described in the first, second or third aspects of the invention.

In an embodiment, the invention provides a soybean plant comprising soybean seed whose oil content is at least 20% on a weight basis. Preferably, the soybean plant or its seed have features as described in the first, second or third aspects of the invention.

In an embodiment, the invention provides a lupin plant comprising lupin seed whose oil content is at least 10% on a weight basis. Preferably, the lupin plant or its seed have features as described in the first, second or third aspects of the invention.

In an embodiment, the invention provides a peanut plant comprising peanuts whose oil content is at least 50% on a weight basis. Preferably, the peanut plant or its seed have features as described in the first, second or third aspects of the invention.

In an embodiment, the invention provides a sunflower plant comprising sunflower seed whose oil content is at least 50% on a weight basis. Preferably, the sunflower plant or its seed have features as described in the first, second or third aspects of the invention.

5 In an embodiment, the invention provides a cotton plant comprising cotton seed whose oil content is at least 41% on a weight basis. Preferably, the cotton plant or its seed have features as described in the first, second or third aspects of the invention.

10 In an embodiment, the invention provides a safflower plant comprising safflower seed whose oil content is at least 35% on a weight basis. Preferably, the safflower plant or its seed have features as described in the first, second or third aspects of the invention.

15 In an embodiment, the invention provides a flax plant comprising flax seed whose oil content is at least 36% on a weight basis. Preferably, the flax plant or its seed have features as described in the first, second or third aspects of the invention.

In an embodiment, the invention provides a *Camelina sativa* plant comprising *Camelina sativa* seed whose oil content is at least 36% on a weight basis. Preferably, the *Camelina sativa* plant or its seed have features as described in the first, second or third aspects of the invention.

20 In embodiments, the plants may be further defined by Features (i), (ii) and (iii) as described hereinbefore. In a preferred embodiment, the plant or the vegetative part comprises one or more or all of the following features:

(i) oleic acid in a vegetative part or seed of the plant, the oleic acid being in an esterified or non-esterified form, wherein at least 20% (mol%), at least 22% (mol%),
25 at least 30% (mol%), at least 40% (mol%), at least 50% (mol%), or at least 60% (mol%), preferably at least 65% (mol%) or at least 66% (mol%) of the total fatty acids in the lipid content of the vegetative part or seed is oleic acid,

(ii) oleic acid in a vegetative part or seed of the plant, the oleic acid being in an esterified form in non-polar lipid, wherein at least 20% (mol%), at least 22% (mol%),
30 at least 30% (mol%), at least 40% (mol%), at least 50% (mol%), or at least 60% (mol%), preferably at least 65% (mol%) or at least 66% (mol%) of the total fatty acids in the non-polar lipid content of the vegetative part or seed is oleic acid,

(iii) a modified fatty acid in a vegetative part or seed of the plant, the modified fatty acid being in an esterified or non-esterified form, preferably in an esterified form
35 in non-polar lipids of the vegetative part or seed, wherein the modified fatty acid comprises a hydroxyl group, an epoxy group, a cyclopropane group, a double carbon-carbon bond, a triple carbon-carbon bond, conjugated double bonds, a branched chain such as a methylated or hydroxylated branched chain, or a combination of two or

more thereof, or any of two, three, four, five or six of the aforementioned groups, bonds or branched chains, and

(iv) waxes and/or wax esters in the non-polar lipid of the vegetative part or seed of the plant.

5 In an embodiment, the plant or the vegetative plant part is a member of a population or collection of at least about 1000 such plants or parts. That is, each plant or plant part in the population or collection has essentially the same properties or comprise the same exogenous nucleic acids as the other members of the population or collection. Preferably, the plants are homozygous for the exogenous polynucleotides, 10 which provides a degree of uniformity. Preferably, the plants are growing in a field. The collection of vegetative plants parts have preferably been harvested from plants growing in a field. Preferably, the vegetative plant parts have been harvested at a time when the yield of non-polar lipids are at their highest. In one embodiment, the vegetative plant parts have been harvested about at the time of flowering. In another 15 embodiment, the vegetative plant parts are harvested when the plants are at least about 1 month of age. In another embodiment, the vegetative plant parts are harvested from about at the time of flowering to about the beginning of senescence. In another embodiment, the vegetative plant parts are harvested at least about 1 month after induction of expression of inducible genes.

20 In a further embodiment of the fourth aspect, the invention provides a vegetative plant part, non-human organism or a part thereof, or seed comprising one or more exogenous polynucleotide(s) and an increased level of one or more non-polar lipid(s) relative to a corresponding vegetative plant part, non-human organism or a part thereof, or seed lacking the one or more exogenous polynucleotide(s), wherein 25 each of the one or more exogenous polynucleotides is operably linked to a promoter which is capable of directing expression of the polynucleotide in a vegetative plant part, non-human organism or part thereof, or seed and wherein one or more or all of the following features apply:

30 (i) the one or more exogenous polynucleotide(s) comprise a first exogenous polynucleotide which encodes an RNA or transcription factor polypeptide that increases the expression of one or more glycolytic or fatty acid biosynthetic genes in a vegetative plant part, non-human organism or a part thereof, or seed and a second exogenous polynucleotide which encodes an RNA or polypeptide involved in biosynthesis of one or more non-polar lipids,

35 (ii) if the non-human organism is a plant, a vegetative part of the plant has a total non-polar lipid content of at least about 3%, more preferably at least about 5%, preferably at least about 7%, more preferably at least about 10%, more preferably at least about 11%, more preferably at least about 12%, more preferably at least about

13%, more preferably at least about 14%, or more preferably at least about 15% (w/w dry weight),

5 (iii) the non-human organism is an alga selected from the group consisting of diatoms (bacillariophytes), green algae (chlorophytes), blue-green algae (cyanophytes), golden-brown algae (chrysophytes), haptophytes, brown algae and heterokont algae,

10 (iv) the non-polar lipid(s) comprise a fatty acid which comprises a hydroxyl group, an epoxy group, a cyclopropane group, a double carbon-carbon bond, a triple carbon-carbon bond, conjugated double bonds, a branched chain such as a methylated or hydroxylated branched chain, or a combination of two or more thereof, or any of two, three, four, five or six of the aforementioned groups, bonds or branched chains,

15 (v) the vegetative plant part, non-human organism or part thereof, or seed comprises oleic acid in an esterified or non-esterified form in its lipid, wherein at least 20% (mol%), at least 22% (mol%), at least 30% (mol%), at least 40% (mol%), at least 50% (mol%), or at least 60% (mol%), preferably at least 65% (mol%) or at least 66% (mol%) of the total fatty acids in the lipid of the vegetative plant part, non-human organism or part thereof, or seed is oleic acid,

20 (vi) the vegetative plant part, non-human organism or part thereof, or seed comprises oleic acid in an esterified form in its non-polar lipid, wherein at least 20% (mol%), at least 22% (mol%), at least 30% (mol%), at least 40% (mol%), at least 50% (mol%), or at least 60% (mol%), preferably at least 65% (mol%) or at least 66% (mol%) of the total fatty acids in the non-polar lipid of the vegetative plant part, non-human organism or part thereof, or seed is oleic acid,

25 (vii) the total fatty acid content in the lipid of the vegetative plant part, non-human organism or part thereof, or seed comprises at least 2% more oleic acid and/or at least 2% less palmitic acid than the lipid in the corresponding vegetative plant part, non-human organism or part thereof, or seed lacking the one or more exogenous polynucleotides, and/or

30 (viii) the total fatty acid content in the non-polar lipid of the vegetative plant part, non-human organism or part thereof, or seed comprises at least 2% more oleic acid and/or at least 2% less palmitic acid than the non-polar lipid in the corresponding vegetative plant part, non-human organism or part thereof, or seed lacking the one or more exogenous polynucleotides,

35 (ix) the non-polar lipid(s) comprise a modified level of total sterols, preferably free sterols, steroyl esters and/or steroyl glycosides,

(x) the non-polar lipid(s) comprise waxes and/or wax esters, and

(xi) the non-human organism or part thereof is one member of a population or collection of at least about 1000 such non-human organisms or parts thereof.

In an embodiment, the one or more exogenous polynucleotide(s) comprise the first exogenous polynucleotide and the second exogenous polynucleotide, and wherein one or more or all of the features (ii) to (xi) apply.

In an embodiment of (ii) above, the total non-polar lipid content is between 5% and 25%, between 7% and 25%, between 10% and 25%, between 12% and 25%, between 15% and 25%, between 7% and 20%, between 10% and 20%, about 10%, about 11%, about 12%, about 13%, about 14%, about 15%, about 16%, about 17%, about 18%, about 20%, or about 22%, each as a percentage of dry weight. In a more preferred embodiment, the vegetative plant part is a leaf portion having a surface area of at least 1 cm².

In preferred embodiments, the non-human organism or part thereof is a plant, an alga or an organism suitable for fermentation such as a fungus. The part of the non-human organism may be a seed, fruit, or a vegetative part of a plant such as an aerial plant part or a green part such as a leaf or stem. In another embodiment, the part is a cell of a multicellular organism. With respect to the part of the non-human organism, the part comprises at least one cell of the non-human organism. In further preferred embodiments, the non-human organism or part thereof is further defined by features as defined in any of the embodiments described in the first, second or third aspects of the invention, including but not limited to Features (i), (ii) and (iii), and the exogenous polynucleotides or combinations of exogenous polynucleotides as defined in any of the embodiments described in the first, second or third aspects of the invention.

In an embodiment, the plant, vegetative plant part, non-human organism or part thereof, or seed comprises one or more exogenous polynucleotides which encode:

- i) a Wrinkled 1 (WRI1) transcription factor and a DGAT,
- ii) a WRI1 transcription factor and a DGAT and an Oleosin,
- iii) a WRI1 transcription factor, a DGAT, a MGAT and an Oleosin,
- iv) a monoacylglycerol acyltransferase (MGAT),
- v) a diacylglycerol acyltransferase 2 (DGAT2),
- vi) a MGAT and a glycerol-3-phosphate acyltransferase (GPAT),
- vii) a MGAT and a DGAT,
- viii) a MGAT, a GPAT and a DGAT,
- ix) a WRI1 transcription factor and a MGAT,
- x) a WRI1 transcription factor, a DGAT and a MGAT,
- xi) a WRI1 transcription factor, a DGAT, a MGAT, an Oleosin and a GPAT,
- xii) a DGAT and an Oleosin, or
- xiii) a MGAT and an Oleosin, and
- xiv) optionally, a silencing suppressor polypeptide,

wherein each exogenous polynucleotide is operably linked to a promoter which is capable of directing expression of the polynucleotide in a plant, vegetative plant part, non-human organism or part thereof, or seed, respectively. The one or more exogenous polynucleotides may comprise nucleotides whose sequence is defined herein. Preferably, the plant, vegetative plant part, non-human organism or part thereof, or seed is homozygous for the one or more exogenous polynucleotides. Preferably, the exogenous polynucleotides are integrated into the genome of the plant, vegetative plant part, non-human organism or part thereof, or seed. The one or more polynucleotides may be provided as separate molecules or may be provided as a contiguous single molecule. Preferably, the exogenous polynucleotides are integrated in the genome of the plant or organism at a single genetic locus or genetically linked loci, more preferably in the homozygous state. More preferably, the integrated exogenous polynucleotides are genetically linked with a selectable marker gene such as an herbicide tolerance gene.

In a preferred embodiment, the vegetative plant part, the non-human organism or part thereof, or the seed comprises a first exogenous polynucleotide encoding a WRI1 and a second exogenous polynucleotide encoding a DGAT, preferably a DGAT1.

In another preferred embodiment, the vegetative plant part, the non-human organism or part thereof, or the seed comprises a first exogenous polynucleotide encoding a WRI1, a second exogenous polynucleotide encoding a DGAT, preferably a DGAT1, and a third exogenous polynucleotide encoding an oleosin.

In a further embodiment, the vegetative plant part, the non-human organism or part thereof, or the seed comprises a first exogenous polynucleotide encoding a WRI1, a second exogenous polynucleotide encoding a DGAT, preferably a DGAT1, a third exogenous polynucleotide encoding an oleosin, and a fourth exogenous polynucleotide encoding an MGAT, preferably an MGAT2.

In a further embodiment, the vegetative plant part, the non-human organism or part thereof, or the seed comprises a first exogenous polynucleotide encoding a WRI1, a second exogenous polynucleotide encoding a DGAT, preferably a DGAT1, a third exogenous polynucleotide encoding an oleosin, and a fourth exogenous polynucleotide encoding LEC2 or BBM.

In a further embodiment, the vegetative plant part, the non-human organism or part thereof, or the seed comprises a first exogenous polynucleotide encoding a WRI1, a second exogenous polynucleotide encoding a DGAT, preferably a DGAT1, a third exogenous polynucleotide encoding an oleosin, a fourth exogenous polynucleotide encoding an MGAT, preferably an MGAT2, and a fifth exogenous polynucleotide encoding LEC2 or BBM.

In a further embodiment, the vegetative plant part, the non-human organism or part thereof, or the seed comprises a first exogenous polynucleotide encoding a WRI1, a second exogenous polynucleotide encoding a DGAT, preferably a DGAT1, a third exogenous polynucleotide encoding an oleosin, and a fourth exogenous polynucleotide encoding an RNA molecule which inhibits expression of a gene encoding a lipase such as a CGi58 polypeptide.

In a further embodiment, the vegetative plant part, the non-human organism or part thereof, or the seed comprises a first exogenous polynucleotide encoding a WRI1, a second exogenous polynucleotide encoding a DGAT, preferably a DGAT1, a third exogenous polynucleotide encoding an oleosin, a fourth exogenous polynucleotide encoding an RNA molecule which inhibits expression of a gene encoding a lipase such as a CGi58 polypeptide, and a fifth exogenous polynucleotide encoding LEC2 or BBM.

In a further embodiment, the vegetative plant part, the non-human organism or part thereof, or the seed comprises a first exogenous polynucleotide encoding a WRI1, a second exogenous polynucleotide encoding a DGAT, preferably a DGAT1, a third exogenous polynucleotide encoding an oleosin, a fourth exogenous polynucleotide encoding an RNA molecule which inhibits expression of a gene encoding a lipase such as a CGi58 polypeptide, and a fifth exogenous polynucleotide encoding an MGAT, preferably an MGAT2.

In a further embodiment, the vegetative plant part, the non-human organism or part thereof, or the seed comprises a first exogenous polynucleotide encoding a WRI1, a second exogenous polynucleotide encoding a DGAT, preferably a DGAT1, a third exogenous polynucleotide encoding an oleosin, a fourth exogenous polynucleotide encoding an RNA molecule which inhibits expression of a gene encoding a lipase such as a CGi58 polypeptide, a fifth exogenous polynucleotide encoding an MGAT, preferably an MGAT2, and a sixth exogenous polynucleotide encoding LEC2 or BBM.

In an embodiment, the seed comprises a first exogenous polynucleotide encoding a WRI1, a second exogenous polynucleotide encoding a DGAT, preferably a DGAT1, a third exogenous polynucleotide encoding an oleosin, and a fourth exogenous polynucleotide encoding an MGAT, preferably an MGAT2. Preferably, the seed further comprises a fifth exogenous polynucleotide encoding a GPAT.

Where relevant, instead of a polynucleotide encoding an RNA molecule which inhibits expression of a gene encoding a lipase such as a CGi58 polypeptide, the vegetative plant part, the non-human organism or part thereof, or the seed has one or more introduced mutations in the lipase gene such as a CGi58 gene which confers

reduced levels of the lipase polypeptide when compared to an isogenic vegetative plant part, non-human organism or part thereof, or seed lacking the mutation.

In a preferred embodiment, the exogenous polynucleotides encoding the DGAT and oleosin are operably linked to a constitutive promoter, or a promoter active in green tissues of a plant at least before and up until flowering, which is capable of directing expression of the polynucleotides in the vegetative plant part, the non-human organism or part thereof, or the seed. In a further preferred embodiment, the exogenous polynucleotide encoding WRI1, and RNA molecule which inhibits expression of a gene encoding a lipase such as a CGi58 polypeptide, is operably linked to a constitutive promoter, a promoter active in green tissues of a plant at least before and up until flowering, or an inducible promoter, which is capable of directing expression of the polynucleotides in the vegetative plant part, the non-human organism or part thereof, or the seed. In yet a further preferred embodiment, the exogenous polynucleotides encoding LEC2, BBM and/or MGAT2 are operably linked to an inducible promoter which is capable of directing expression of the polynucleotides in the vegetative plant part, the non-human organism or part thereof, or the seed.

In each of the above embodiments, the total non-polar lipid content of the vegetative plant part, or non-human organism or part thereof, or the seed, preferably a plant leaf or part thereof, stem or tuber, is at least about 3%, more preferably at least about 5%, preferably at least about 7%, more preferably at least about 10%, more preferably at least about 11%, more preferably at least about 12%, more preferably at least about 13%, more preferably at least about 14%, or more preferably at least about 15% (w/w dry weight or seed weight). In a further preferred embodiment, the total non-polar lipid content is between 5% and 25%, between 7% and 25%, between 10% and 25%, between 12% and 25%, between 15% and 25%, between 7% and 20%, between 10% and 20%, between 10% and 15%, between 15% and 20%, between 20% and 25%, about 10%, about 11%, about 12%, about 13%, about 14%, about 15%, about 16%, about 17%, about 18%, about 20%, or about 22%, each as a percentage of dry weight or seed weight. In a particularly preferred embodiment, the vegetative plant part is a leaf (or leaves) or a portion thereof. In a more preferred embodiment, the vegetative plant part is a leaf portion having a surface area of at least 1 cm².

Furthermore, in each of the above embodiments, the total TAG content of the vegetative plant part, or non-human organism or part thereof, or the seed, preferably a plant leaf or part thereof, stem or tuber, is at least about 3%, more preferably at least about 5%, preferably at least about 7%, more preferably at least about 10%, more preferably at least about 11%, more preferably at least about 12%, more preferably at least about 13%, more preferably at least about 14%, more preferably at least about

15%, or more preferably at least about 17% (w/w dry weight or seed weight). In a further preferred embodiment, the total TAG content is between 5% and 30%, between 7% and 30%, between 10% and 30%, between 12% and 30%, between 15% and 30%, between 7% and 30%, between 10% and 30%, between 20% and 28%,
 5 between 18% and 25%, between 22% and 30%, about 10%, about 11%, about 12%, about 13%, about 14%, about 15%, about 16%, about 17%, about 18%, about 20%, or about 22%, each as a percentage of dry weight or seed weight. In a particularly preferred embodiment, the vegetative plant part is a leaf (or leaves) or a portion thereof. In a more preferred embodiment, the vegetative plant part is a leaf portion
 10 having a surface area of at least 1 cm².

Furthermore, in each of the above embodiments, the total lipid content of the vegetative plant part, or non-human organism or part thereof, or the seed, preferably a plant leaf or part thereof, stem or tuber, is at least about 3%, more preferably at least about 5%, preferably at least about 7%, more preferably at least about 10%, more
 15 preferably at least about 11%, more preferably at least about 12%, more preferably at least about 13%, more preferably at least about 14%, more preferably at least about 15%, more preferably at least about 17% (w/w dry weight or seed weight), more preferably at least about 20%, more preferably at least about 25%. In a further preferred embodiment, the total lipid content is between 5% and 35%, between 7%
 20 and 35%, between 10% and 35%, between 12% and 35%, between 15% and 35%, between 7% and 35%, between 10% and 20%, between 18% and 28%, between 20% and 28%, between 22% and 28%, about 10%, about 11%, about 12%, about 13%, about 14%, about 15%, about 16%, about 17%, about 18%, about 20%, about 22%, or about 25%, each as a percentage of dry weight or seed weight. Typically, the total
 25 lipid content of the vegetative plant part, or non-human organism or part thereof is about 2-3% higher than the non-polar lipid content. In a particularly preferred embodiment, the vegetative plant part is a leaf (or leaves) or a portion thereof. In a more preferred embodiment, the vegetative plant part is a leaf portion having a surface area of at least 1 cm².

30 In an embodiment, the vegetative plant part, the non-human organism or part thereof, or the seed, preferably the vegetative plant part, comprises a first exogenous polynucleotide encoding a WRI1, a second exogenous polynucleotide encoding a DGAT, preferably a DGAT1, a third exogenous polynucleotide encoding an MGAT, preferably an MGAT2, and a fourth exogenous polynucleotide encoding an oleosin,
 35 wherein the vegetative plant part, non-human organism or part thereof, or seed has one or more or all of the following features:

i) a total lipid content of at least 8%, at least 10%, at least 12%, at least 14%, or at least 15.5% (% weight of dry weight or seed weight),

ii) at least a 3 fold, at least a 5 fold, at least a 7 fold, at least an 8 fold, or least a 10 fold, at higher total lipid content in the vegetative plant part or non-human organism relative to a corresponding vegetative plant part or non-human organism lacking the exogenous polynucleotides,

5 iii) a total TAG content of at least 5%, at least 6%, at least 6.5% or at least 7% (% weight of dry weight or seed weight),

iv) at least a 40 fold, at least a 50 fold, at least a 60 fold, or at least a 70 fold, or at least a 100 fold, higher total TAG content relative to a corresponding vegetative plant part or non-human organism lacking the exogenous polynucleotides,

10 v) oleic acid comprises at least 15%, at least 19% or at least 22% (% weight) of the fatty acids in TAG,

vi) at least a 10 fold, at least a 15 fold or at least a 17 fold higher level of oleic acid in TAG relative to a corresponding vegetative plant part or non-human organism lacking the exogenous polynucleotides,

15 vii) palmitic acid comprises at least 20%, at least 25%, at least 30% or at least 33% (% weight) of the fatty acids in TAG,

viii) at least a 1.5 fold higher level of palmitic acid in TAG relative to a corresponding vegetative plant part on non-human organism lacking the exogenous polynucleotides,

20 ix) linoleic acid comprises at least 22%, at least 25%, at least 30% or at least 34% (% weight) of the fatty acids in TAG,

x) α -linolenic acid comprises less than 20%, less than 15%, less than 11% or less than 8% (% weight) of the fatty acids in TAG, and

25 xi) at least a 5 fold, or at least an 8 fold, lower level of α -linolenic acid in TAG relative to a corresponding vegetative plant part or non-human organism lacking the exogenous polynucleotides. In this embodiment, preferably the vegetative plant part at least has feature(s), i), ii) iii), iv), i) and ii), i) and iii), i) and iv), i) to iii), i), iii) and iv), i) to iv), ii) and iii), ii) and iv), ii) to iv), or iii) and iv). In an embodiment, % dry weight is % leaf dry weight.

30 In a further embodiment, the vegetative plant part, the non-human organism or part thereof, or the seed, preferably the vegetative plant part, comprises a first exogenous polynucleotide encoding a WRI1, a second exogenous polynucleotide encoding a DGAT, preferably a DGAT1, a third exogenous polynucleotide encoding an oleosin, wherein the vegetative plant part, non-human organism or part thereof, or seed has one or more or all of the following features:

35 i) a total TAG content of at least 10%, at least 12.5%, at least 15% or at least 17% (% weight of dry weight or seed weight),

ii) least a 40 fold, at least a 50 fold, at least a 60 fold, or at least a 70 fold, or at least a 100 fold, higher total TAG content in the vegetative plant part or non-human organism relative to a corresponding vegetative plant part or non-human organism lacking the exogenous polynucleotides,

5 iii) oleic acid comprises at least 19%, at least 22%, or at least 25% (% weight) of the fatty acids in TAG,

iv) at least a 10 fold, at least a 15 fold, at least a 17 fold, or at least a 19 fold, higher level of oleic acid in TAG in the vegetative plant part or non-human organism relative to a corresponding vegetative plant part or non-human organism lacking the exogenous polynucleotides,

10 v) palmitic acid comprises at least 20%, at least 25%, or at least 28% (% weight) of the fatty acids in TAG,

vi) at least a 1.25 fold higher level of palmitic acid in TAG in the vegetative plant part or non-human organism relative to a corresponding vegetative plant part or non-human organism lacking the exogenous polynucleotides,

15 vii) linoleic acid comprises at least 15%, or at least 20%, (% weight) of the fatty acids in TAG,

viii) α -linolenic acid comprises less than 15%, less than 11% or less than 8% (% weight) of the fatty acids in TAG, and

20 ix) at least a 5 fold, or at least an 8 fold, lower level of α -linolenic acid in TAG in the vegetative plant part or non-human organism relative to a corresponding vegetative plant part or non-human organism lacking the exogenous polynucleotides. In this embodiment, preferably the vegetative plant part at least has feature(s), i), ii), or i) and ii). In an embodiment, % dry weight is % leaf dry weight.

25 Preferably, the defined features for the two above embodiments are as at the flowering stage of the plant.

In a fifth aspect, the invention provides a non-human organism or a part thereof comprising one or more exogenous polynucleotide(s) and an increased level of one or more non-polar lipid(s) relative to a corresponding non-human organism or a part thereof lacking the one or more exogenous polynucleotide(s), wherein each of the one or more exogenous polynucleotides is operably linked to a promoter which is capable of directing expression of the polynucleotide in a non-human organism or part thereof, and wherein:

35 (i) the one or more exogenous polynucleotide(s) comprise a first exogenous polynucleotide which encodes a Wrinkled 1 (WRI1) transcription factor polypeptide, and a second exogenous polynucleotide which encodes a fatty acid acyltransferase, and

(ii) the first and second exogenous polynucleotides have an effect on non-polar lipid accumulation in the non-human organism that is larger than an additive effect of the individual effects of each polynucleotide alone.

In a sixth aspect, the invention provides a plant seed capable of growing into a plant of the invention, or obtained from a plant of the invention, for example a non-human organism of the invention which is a plant. In an embodiment, the seed comprises one or more exogenous polynucleotides as defined herein.

In a seventh aspect, the invention provides a process for obtaining a cell with enhanced ability to produce one or more non-polar lipids, the process comprising the steps of:

- a) introducing into a cell one or more exogenous polynucleotides,
- b) expressing the one or more exogenous polynucleotides in the cell or a progeny cell thereof,
- c) analysing the lipid content of the cell or progeny cell, and
- d) selecting a cell or progeny cell having an increased level of one or more non-polar lipids relative to a corresponding cell or progeny cell lacking the exogenous polynucleotides,

wherein the one or more exogenous polynucleotides encode

- i) a Wrinkled 1 (WRI1) transcription factor and a DGAT,
- ii) a WRI1 transcription factor and a DGAT and an Oleosin,
- iii) a WRI1 transcription factor, a DGAT, a MGAT and an Oleosin,
- iv) a monoacylglycerol acyltransferase (MGAT),
- v) a diacylglycerol acyltransferase 2 (DGAT2),
- vi) a MGAT and a glycerol-3-phosphate acyltransferase (GPAT),
- vii) a MGAT and a DGAT,
- viii) a MGAT, a GPAT and a DGAT,
- ix) a WRI1 transcription factor and a MGAT,
- x) a WRI1 transcription factor, a DGAT and a MGAT,
- xi) a WRI1 transcription factor, a DGAT, a MGAT, an Oleosin and a GPAT,
- xii) a DGAT and an Oleosin, or
- xiii) a MGAT and an Oleosin, and
- xiv) optionally, a silencing suppressor polypeptide,

wherein each exogenous polynucleotide is operably linked to a promoter that is capable of directing expression of the exogenous polynucleotide in the cell or progeny cell.

In an embodiment, the selected cell or progeny cell comprises:

- i) a first exogenous polynucleotide encoding a WRI1 and a second exogenous polynucleotide encoding a DGAT, preferably a DGAT1,

ii) a first exogenous polynucleotide encoding a WRI1, a second exogenous polynucleotide encoding a DGAT, preferably a DGAT1, and a third exogenous polynucleotide encoding an oleosin,

5 iii) a first exogenous polynucleotide encoding a WRI1, a second exogenous polynucleotide encoding a DGAT, preferably a DGAT1, a third exogenous polynucleotide encoding an oleosin, and a fourth exogenous polynucleotide encoding an MGAT, preferably an MGAT2,

10 iv) a first exogenous polynucleotide encoding a WRI1, a second exogenous polynucleotide encoding a DGAT, preferably a DGAT1, a third exogenous polynucleotide encoding an oleosin, and a fourth exogenous polynucleotide encoding LEC2 or BBM,

15 v) a first exogenous polynucleotide encoding a WRI1, a second exogenous polynucleotide encoding a DGAT, preferably a DGAT1, a third exogenous polynucleotide encoding an oleosin, a fourth exogenous polynucleotide encoding an MGAT, preferably an MGAT2, and a fifth exogenous polynucleotide encoding LEC2 or BBM,

20 vi) a first exogenous polynucleotide encoding a WRI1, a second exogenous polynucleotide encoding a DGAT, preferably a DGAT1, a third exogenous polynucleotide encoding an oleosin, and a fourth exogenous polynucleotide encoding an RNA molecule which inhibits expression of a gene encoding a lipase such as a CGi58 polypeptide,

25 vii) a first exogenous polynucleotide encoding a WRI1, a second exogenous polynucleotide encoding a DGAT, preferably a DGAT1, a third exogenous polynucleotide encoding an oleosin, a fourth exogenous polynucleotide encoding an RNA molecule which inhibits expression of a gene encoding a lipase such as a CGi58 polypeptide, and a fifth exogenous polynucleotide encoding LEC2 or BBM,

30 viii) a first exogenous polynucleotide encoding a WRI1, a second exogenous polynucleotide encoding a DGAT, preferably a DGAT1, a third exogenous polynucleotide encoding an oleosin, a fourth exogenous polynucleotide encoding an RNA molecule which inhibits expression of a gene encoding a lipase such as a CGi58 polypeptide, and a fifth exogenous polynucleotide encoding an MGAT, preferably an MGAT2, or

35 ix) a first exogenous polynucleotide encoding a WRI1, a second exogenous polynucleotide encoding a DGAT, preferably a DGAT1, a third exogenous polynucleotide encoding an oleosin, a fourth exogenous polynucleotide encoding an RNA molecule which inhibits expression of a gene encoding a lipase such as a CGi58 polypeptide, a fifth exogenous polynucleotide encoding an MGAT, preferably an MGAT2, and a sixth exogenous polynucleotide encoding LEC2 or BBM.

In a further embodiment, the selected cell or progeny cell is a cell of a plant seed and comprises a first exogenous polynucleotide encoding a WRI1, a second exogenous polynucleotide encoding a DGAT, preferably a DGAT1, a third exogenous polynucleotide encoding an oleosin, and a fourth exogenous polynucleotide encoding an MGAT, preferably an MGAT2. Preferably, the seed further comprises a fifth exogenous polynucleotide encoding a GPAT.

In a preferred embodiment, the one or more exogenous polynucleotides are stably integrated into the genome of the cell or progeny cell.

In a preferred embodiment, the process further comprises a step of regenerating a transgenic plant from the cell or progeny cell comprising the one or more exogenous polynucleotides. The step of regenerating a transgenic plant may be performed prior to the step of expressing the one or more exogenous polynucleotides in the cell or a progeny cell thereof, and/or prior to the step of analysing the lipid content of the cell or progeny cell, and/or prior to the step of selecting the cell or progeny cell having an increased level of one or more non-polar lipids. The process may further comprise a step of obtaining seed or a progeny plant from the transgenic plant, wherein the seed or progeny plant comprises the one or more exogenous polynucleotides.

The process of the seventh aspect may be used as a screening assay to determine whether a polypeptide encoded by an exogenous polynucleotide has a desired function. The one or more exogenous polynucleotides in this aspect may comprise a sequence as defined above. Further, the one or more exogenous polynucleotides may not be known prior to the process to encode a WRI1 transcription factor and a DGAT, a WRI1 transcription factor and a MGAT, a WRI1 transcription factor, a DGAT and a MGAT, a WRI1 transcription factor, a DGAT, a MGAT and an Oleosin, a WRI1 transcription factor, a DGAT, a MGAT, an Oleosin and a GPAT, a WRI1 transcription factor, a DGAT and an oleosin, a DGAT and an Oleosin, or a MGAT and an Oleosin, but rather may be candidates therefor. The process therefore may be used as an assay to identify or select polynucleotides encoding a WRI1 transcription factor and a DGAT, a WRI1 transcription factor and a MGAT, a WRI1 transcription factor, a DGAT and a MGAT, a WRI1 transcription factor, a DGAT, a MGAT and an Oleosin, a WRI1 transcription factor, a DGAT, a MGAT, an Oleosin and a GPAT, a WRI1 transcription factor, a DGAT and an oleosin, a DGAT and an Oleosin, or a MGAT and an Oleosin. The candidate polynucleotides are introduced into a cell and the products analysed to determine whether the candidates have the desired function.

In an eighth aspect, the invention provides a transgenic cell or transgenic plant obtained using a process of the invention, or a vegetative plant part or seed obtained therefrom which comprises the one or more exogenous polynucleotides.

In a ninth aspect, the invention provides a use of one or more polynucleotides encoding, or a genetic construct comprising polynucleotides encoding:

- i) a Wrinkled 1 (WRI1) transcription factor and a DGAT,
- ii) a WRI1 transcription factor and a DGAT and an Oleosin,
- iii) a WRI1 transcription factor, a DGAT, a MGAT and an Oleosin,
- iv) a monoacylglycerol acyltransferase (MGAT),
- v) a diacylglycerol acyltransferase 2 (DGAT2),
- vi) a MGAT and a glycerol-3-phosphate acyltransferase (GPAT),
- vii) a MGAT and a DGAT,
- viii) a MGAT, a GPAT and a DGAT,
- ix) a WRI1 transcription factor and a MGAT,
- x) a WRI1 transcription factor, a DGAT and a MGAT,
- xi) a WRI1 transcription factor, a DGAT, a MGAT, an Oleosin and a GPAT,
- xii) a DGAT and an Oleosin, or
- xiii) a MGAT and an Oleosin, and
- xiv) optionally, a silencing suppressor polypeptide,

for producing a transgenic cell, a transgenic non-human organism or a part thereof or a transgenic seed having an enhanced ability to produce one or more non-polar lipids relative to a corresponding cell, non-human organism or part thereof, or seed lacking the one or more polynucleotides, wherein each of the one or more polynucleotides is exogenous to the cell, non-human organism or part thereof, or seed and is operably linked to a promoter which is capable of directing expression of the polynucleotide in a cell, a non-human organism or a part thereof or a seed, respectively.

In an embodiment, the invention provides a use of a first polynucleotide encoding an RNA or transcription factor polypeptide that increases the expression of one or more glycolytic or fatty acid biosynthetic genes in a cell, a non-human organism or a part thereof, or a seed, together with a second polynucleotide that encodes an RNA or polypeptide involved in biosynthesis of one or more non-polar lipids, for producing a transgenic cell, a transgenic non-human organism or part thereof, or a transgenic seed having an enhanced ability to produce one or more non-polar lipids relative to a corresponding cell, non-human organism or part thereof, or seed lacking the first and second polynucleotides, wherein the first and second polynucleotides are each exogenous to the cell, non-human organism or part thereof, or seed and are each operably linked to a promoter which is capable of directing

expression of the polynucleotide in the transgenic cell, transgenic non-human organism or part thereof, or transgenic seed, respectively.

In a further embodiment, the invention provides a use of one or more polynucleotides for producing a transgenic cell, a transgenic non-human organism or part thereof, or a transgenic seed having an enhanced ability to produce one or more non-polar lipid(s) relative to a corresponding cell, non-human organism or part thereof, or seed lacking the one or more exogenous polynucleotides, wherein each of the one or more polynucleotides is exogenous to the cell, non-human organism or part thereof, or seed and is operably linked to a promoter which is capable of directing expression of the polynucleotide in a cell, a non-human organism or a part thereof, or a seed, respectively, and wherein the non-polar lipid(s) comprise a fatty acid which comprises a hydroxyl group, an epoxy group, a cyclopropane group, a double carbon-carbon bond, a triple carbon-carbon bond, conjugated double bonds, a branched chain such as a methylated or hydroxylated branched chain, or a combination of two or more thereof, or any of two, three, four, five or six of the aforementioned groups, bonds or branched chains. Such uses also have utility as screening assays.

In a tenth aspect, the invention provides the use of a first polynucleotide encoding a Wrinkled 1 (WRI1) transcription factor polypeptide, together with a second polynucleotide that encodes a fatty acid acyltransferase, for producing a transgenic non-human organism or part thereof having an enhanced ability to produce one or more non-polar lipids relative to a corresponding non-human organism or part thereof lacking the first and second polynucleotides, wherein the first and second polynucleotides are each exogenous to the non-human organism or part thereof and are each operably linked to a promoter which is capable of directing expression of the polynucleotide in the transgenic non-human organism or part thereof, wherein the first and second exogenous polynucleotides have an effect on non-polar lipid accumulation in the non-human organism that is larger than an additive effect of the individual effects of each polynucleotide alone.

In an eleventh aspect, the invention provides a process for producing seed, the process comprising:

i) growing a plant, multiple plants, or non-human organism according to the invention, and

ii) harvesting seed from the plant, plants, or non-human organism.

In a preferred embodiment, the process comprises growing a population of at least about 1000 such plants in a field, and harvesting seed from the population of plants.

The harvested seed may be placed in a container and transported away from the field, for example exported out of the country, or stored prior to use.

In a twelfth aspect, the invention provides a fermentation process comprising the steps of:

5 i) providing a vessel containing a liquid composition comprising a non-human organism of the invention which is suitable for fermentation, and constituents required for fermentation and fatty acid biosynthesis, and

ii) providing conditions conducive to the fermentation of the liquid composition contained in said vessel.

10 In a thirteenth aspect, the invention provides a recovered or extracted lipid obtainable by a process of the invention, or obtainable from a vegetative plant part, non-human organism or part thereof, cell or progeny cell, transgenic plant, or seed of the invention. The recovered or extracted lipid, preferably oil such as seedoil, may have an enhanced TAG content, DAG content, TAG and DAG content, MAG content, 15 PUFA content, specific PUFA content, or a specific fatty acid content, and/or total non-polar lipid content. In a preferred embodiment, the MAG is 2-MAG. The extent of the increased TAG content, DAG content, TAG and DAG content, MAG content, PUFA content, specific PUFA content, specific fatty acid content and/or total non-polar lipid content may be as defined in Feature (i).

20 In a fourteenth aspect, the invention provides an industrial product produced by a process of the invention, preferably which is a hydrocarbon product such as fatty acid esters, preferably fatty acid methyl esters and/or a fatty acid ethyl esters, an alkane such as methane, ethane or a longer-chain alkane, a mixture of longer chain alkanes, an alkene, a biofuel, carbon monoxide and/or hydrogen gas, a bioalcohol 25 such as ethanol, propanol, or butanol, biochar, or a combination of carbon monoxide, hydrogen and biochar.

In a fifteenth aspect, the invention provides a use of a plant, vegetative plant part, non-human organism or a part thereof, cell or progeny cell, transgenic plant produced by a process of the invention, or a seed or a recovered or extracted lipid of 30 the invention for the manufacture of an industrial product. The industrial product may be as defined above.

In a sixteenth aspect, the invention provides the use of a non-human organism or a part thereof as described herein for the manufacture of an industrial product, preferably a hydrocarbon product such as fatty acid esters, preferably fatty acid 35 methyl esters and/or a fatty acid ethyl esters, an alkane such as methane, ethane or a longer-chain alkane, a mixture of longer chain alkanes, an alkene, a biofuel, or carbon monoxide and/or hydrogen gas, a bioalcohol such as ethanol, propanol, or butanol, biochar, or a combination of carbon monoxide, hydrogen and biochar.

In a seventeenth aspect, the invention provides a process for producing fuel, the process comprising:

i) reacting a lipid of the invention with an alcohol, optionally in the presence of a catalyst, to produce alkyl esters, and

ii) optionally, blending the alkyl esters with petroleum based fuel. The alkyl esters are preferably methyl esters. The fuel produced by the process may comprise a minimum level of the lipid of the invention or a hydrocarbon product produced therefrom such as at least 10%, at least 20%, or at least 30% by volume.

In an eighteenth aspect, the invention provides a process for producing a synthetic diesel fuel, the process comprising:

i) converting lipid in a vegetative plant, non-human organism or part thereof of the invention to a syngas by gasification, and

ii) converting the syngas to a biofuel using a metal catalyst or a microbial catalyst.

In a nineteenth aspect, the invention provides a process for producing a biofuel, the process comprising converting lipid in a vegetative plant part, non-human organism or part thereof of the invention to bio-oil by pyrolysis, a bioalcohol by fermentation, or a biogas by gasification or anaerobic digestion.

In a twentieth aspect, the invention provides a process for producing a feedstuff, the process comprising admixing a plant, vegetative plant part thereof, non-human organism or part thereof, cell or progeny cell, transgenic plant produced by a process of the invention, seed, recovered or extracted lipid, or an extract or portion thereof, with at least one other food ingredient.

In a further aspect, the invention provides feedstuffs, cosmetics or chemicals comprising a plant, vegetative part thereof, non-human organism or part thereof, cell or progeny cell, transgenic plant produced by a process of the invention, seed, or a recovered or extracted lipid of the invention, or an extract or portion thereof.

Natutally, when vegetative material of a plant is to be harvested because of its oil content it is desirable to harvest the material when lipid levels are as high as possible. The present inventors have noted an association between the glossiness of the vegetative tissue of the plants of the invention and oil content, with high levels of lipid being associated with high gloss. Thus, the glossiness of the vegetative material can be used as marker to assist in determining when to harvest the material.

In a further aspect, the invention provides a recombinant cell comprising one or more exogenous polynucleotide(s) and an increased level of one or more non-polar lipid(s) relative to a corresponding cell lacking the one or more exogenous polynucleotide(s),

wherein each of the one or more exogenous polynucleotides is operably linked to a promoter which is capable of directing expression of the polynucleotide in a cell, and wherein one or more or all of the following features apply:

(a) the one or more exogenous polynucleotide(s) comprise a first exogenous polynucleotide which encodes an RNA or transcription factor polypeptide that increases the expression of one or more glycolytic or fatty acid biosynthetic genes in a non-human organism or a part thereof, and a second exogenous polynucleotide which encodes an RNA or polypeptide involved in biosynthesis of one or more non-polar lipids,

(b) if the cell is a cell of a vegetative part of a plant, the cell has a total non-polar lipid content of at least about 3%, more preferably at least about 5%, preferably at least about 7%, more preferably at least about 10%, more preferably at least about 11%, more preferably at least about 12%, more preferably at least about 13%, more preferably at least about 14%, or more preferably at least about 15% (w/w),

(c) the cell is an alga selected from the group consisting of diatoms (bacillariophytes), green algae (chlorophytes), blue-green algae (cyanophytes), golden-brown algae (chrysophytes), haptophytes, brown algae and heterokont algae,

(d) the one or more non-polar lipid(s) comprise a fatty acid which comprises a hydroxyl group, an epoxy group, a cyclopropane group, a double carbon-carbon bond, a triple carbon-carbon bond, conjugated double bonds, a branched chain such as a methylated or hydroxylated branched chain, or a combination of two or more thereof, or any of two, three, four, five or six of the aforementioned groups, bonds or branched chains,

(e) the total fatty acid content in the non-polar lipid(s) comprises at least 2% more oleic acid and/or at least 2% less palmitic acid than the non-polar lipid(s) in the corresponding cell lacking the one or more exogenous polynucleotides,

(f) the non-polar lipid(s) comprise a modified level of total sterols, preferably free (non-esterified) sterols, steroyl esters, steroyl glycosides, relative to the non-polar lipid(s) in the corresponding cell lacking the one or more exogenous polynucleotides,

(g) the non-polar lipid(s) comprise waxes and/or wax esters, and

(h) the cell is one member of a population or collection of at least about 1000 such cells.

In an embodiment, the one or more exogenous polynucleotide(s) comprise the first exogenous polynucleotide and the second exogenous polynucleotide, and wherein one or more or all of the features (b) to (h) apply.

In a further aspect, the present invention provides a method of determining when to harvest a plant to optimize the amount of lipid in the vegetative tissue of the plant at harvest, the method comprising

- i) measuring the gloss of the vegetative tissue,
- 5 ii) comparing the measurement with a pre-determined minimum glossiness level, and
- iii) optionally harvesting the plant.

In another aspect, the present invention provides a method of predicting the quantity of lipid in vegetative tissue of a plant, the method comprising measuring the gloss of the vegetative tissue.

In a preferred embodiment of the two above aspects the vegetative tissue is a leaf(leaves) or a portion thereof.

In a further aspect, the present invention provides a method of trading a plant or a part thereof, comprising obtaining the plant or part comprising a cell of the invention, and trading the obtained plant or plant part for pecuniary gain.

In an embodiment, the method further comprises one or more or all of:

- i) cultivating the plant,
- ii) harvesting the plant part from the plant,
- iii) storing the plant or part thereof, or
- 20 iv) transporting the plant or part thereof to a different location.

In a further aspect, the present invention provides a process for producing bins of plant parts comprising:

- a) harvesting plant parts comprising a cell of the invention by collecting the plant parts from the plants, or by separating the plant parts from other parts of the plants,
- 25 b) optionally, sifting and/or sorting the harvested plant parts, and
- c) loading the plant parts of a) or the sifted and/or sorted plant parts of b) into bins, thereby producing bins of the plant parts.

Any embodiment herein shall be taken to apply *mutatis mutandis* to any other embodiment unless specifically stated otherwise.

In a further aspect the invention provides a process for feeding an animal comprising providing an animal with the feedstuff as described herein.

The present invention is not to be limited in scope by the specific embodiments described herein, which are intended for the purpose of exemplification only. Functionally-equivalent products, compositions and methods are clearly within the scope of the invention, as described herein.

Throughout this specification, unless specifically stated otherwise or the context requires otherwise, reference to a single step, composition of matter, group of steps or group of compositions of matter shall be taken to encompass one and a plurality (i.e. one or more) of those steps, compositions of matter, groups of steps or group of compositions of matter.

The invention is hereinafter described by way of the following non-limiting Examples and with reference to the accompanying figures.

BRIEF DESCRIPTION OF THE ACCOMPANYING DRAWINGS

Figure 1. A representation of various lipid synthesis pathways, most of which converge at DAG, a central molecule in lipid synthesis. This model includes one possible route to the formation of *sn*-2 MAG which could be used by a bi-functional MGAT/DGAT for DAG formation from glycerol-3-phosphate (G-3-P). Abbreviations are as follows:

G-3-P; glycerol-3-phosphate

LysoPA; lysophosphatidic acid

PA; phosphatidic acid

MAG; monoacylglycerol

DAG; diacylglycerol

TAG; triacylglycerol

Acyl-CoA and FA-CoA; acyl-coenzyme A and fatty acyl-coenzyme A

PC; phosphatidylcholine

GPAT; glycerol-3-phosphate acyltransferase; glycerol-3-phosphate *O*-acyltransferase; acyl-CoA:*sn*-glycerol-3-phosphate 1-O-acyltransferase; EC 2.3.1.15

GPAT4; glycerol-3-phosphate acyltransferase 4

GPAT6; glycerol-3-phosphate acyltransferase 6

LPAAT; 1-acyl-glycerol-3-phosphate acyltransferase; 1-acylglycerol-3-phosphate *O*-acyltransferase; acyl-CoA:1-acyl-*sn*-glycerol-3-phosphate 2-*O*-acyltransferase; EC 2.3.1.51

PAP; phosphatidic acid phosphatase; phosphatidate phosphatase; phosphatic
5 acid phosphohydrolase; phosphatidic acid phosphatase; EC 3.1.3.4

MGAT; an acyltransferase having monoacylglycerol acyltransferase (MGAT; 2-acylglycerol *O*-acyltransferase acyl-CoA:2-acylglycerol *O*-acyltransferase; EC 2.3.1.22) activity

M/DGAT; an acyltransferase having monoacylglycerol acyltransferase
10 (MGAT; 2-acylglycerol *O*-acyltransferase; acyl-CoA:2-acylglycerol *O*-acyltransferase; EC 2.3.1.22) and/or diacylglycerol acyltransferase (DGAT; diacylglycerol *O*-acyltransferase; acyl-CoA:1,2-diacyl-*sn*-glycerol *O*-acyltransferase; EC 2.3.1.20) activity

LPCAT; acyl-CoA:lysophosphatidylcholine acyltransferase; 1-
15 acylglycerophosphocholine *O*-acyltransferase; acyl-CoA:1-acyl-*sn*-glycero-3-phosphocholine *O*-acyltransferase; EC 2.3.1.23

PLD-Z; Phospholipase D zeta; choline phosphatase; lecithinase D; lipophosphodiesterase II; EC 3.1.4.4

CPT; CDP-choline:diacylglycerol cholinephosphotransferase; 1-alkyl-2-
20 acetyl-glycerol cholinephosphotransferase; alkylacylglycerol cholinephosphotransferase; cholinephosphotransferase; phosphorylcholine-glyceride transferase; EC 2.7.8.2

PDCT; phosphatidylcholine:diacylglycerol cholinephosphotransferase

PLC; phospholipase C; EC 3.1.4.3

PDAT; phospholipid:diacylglycerol acyltransferase; phospholipid:1,2-diacyl-
25 *sn*-glycerol *O*-acyltransferase; EC 2.3.1.158

Pi; inorganic phosphate

Figure 2. Relative DAG and TAG increases in *Nicotiana benthamiana* leaf tissue
30 transformed with constructs encoding p19 (negative control), *Arabidopsis thaliana* DGAT1, *Mus musculus* MGAT1 and a combination of DGAT1 and MGAT1, each expressed from the 35S promoter. The MGAT1 enzyme was far more active than the DGAT1 enzyme in promoting both DAG and TAG accumulation in leaf tissue. Expression of the MGAT1 gene resulted in twice as much DAG and TAG
35 accumulation in leaf tissue compared to expression of DGAT1 alone.

Figure 3. Relative TAG increases in *N. benthamiana* leaf transformed with constructs encoding p19 (negative control), *A. thaliana* DGAT1, *M. musculus*

MGAT2 and a combination of MGAT2 and DGAT1. Error bars denote standard error of triplicate samples.

Figure 4. Radioactivity (DPM) in MAG, DAG and TAG fractions isolated from transiently-transformed *N. benthamiana* leaf lysates fed with *sn*-2-MAG [^{14}C] and unlabelled oleic acid over a time-course. The constructs used were as for Figure 3.

Figure 5. As for Figure 4 but fed [^{14}C]G-3-P and unlabelled oleic acid.

Figure 6. Autoradiogram of TLC plate showing TAG-formation by *A. thaliana* DGAT1 and *M. musculus* MGAT1 but not *M. musculus* MGAT2 in yeast assays. The assay is described in Example 5.

Figure 7. TAG levels in *Arabidopsis thaliana* T2 and T3 seeds transformed with a chimeric DNA expressing MGAT2 relative to parental (untransformed) control. Seeds were harvested at maturity (dessicated). SW: desiccated seed weight. TAG levels are given as μg TAG per 100 μg seed weight.

Figure 8. Total fatty acid content in seed of transformed *Arabidopsis thaliana* plants transformed with constructs encoding MGAT1 or MGAT2.

Figure 9. Relative TAG level in transiently-transformed *N. benthamiana* leaf tissue compared to *Arabidopsis thaliana* DGAT1 overexpression.

Figure 10. TAG conversion from *sn*-1,2-DAG in DGAT assay from microsomes of *N. benthamiana* leaf tissues expressing P19 control, *Arabidopsis thaliana* DGAT1 and *Arabidopsis thaliana* DGAT2

Figure 11. Total FAME quantification in *A. thaliana* seeds transformed with pJP3382 and pJP3383.

Figure 12. Maximum TAG levels obtained for different gene combinations transiently expressed in *N. benthamiana* leaves. The V2 negative control represents the average TAG level based on 15 independent repeats.

Figure 13. Co-expression of the genes coding for the *Arabidopsis thaliana* DGAT1 acyltransferase and *A. thaliana* WRI1 transcription factor resulted in a synergistic effect on TAG levels in *Nicotiana benthamiana* leaves. Data shown are averages and standard deviations of five independent infiltrations.

Figure 14 TAG levels in stably-transformed *N. benthamiana* aerial seedling tissue. Total lipids were extracted from aerial tissues of *N. benthamiana* seedlings and analysed by TLC-FID using an internal DAGE standard to allow accurate comparison between samples.

Figure 15. Total fatty acid levels of *A. thaliana* T2 seed populations transformed with control vector (pORE04), *M. musculus* MGAT1 (35S:MGAT1) or *M. musculus* MGAT2 (35S:MGAT2).

10

Figure 16. Map of the insertion region between the left and right borders of pJP3502. TER Glyma-Lectin denotes the *Glycine max* lectin terminator; Arath-WRI1, *Arabidopsis thaliana* WRI1 transcription factor coding region; PRO Arath-Rubisco SSU, *A. thaliana* rubisco small subunit promoter; Sesin-Oleosin, *Sesame indicum* oleosin coding region; PRO CaMV35S-Ex2, cauliflower mosaic virus 35S promoter having a duplicated enhancer region; Arath-DGAT1, *A. thaliana* DGAT1 acyltransferase coding region; TER Agrtu-NOS, *Agrobacterium tumefaciens* nopaline synthase terminator.

Figure 17. Schematic representation of the construct pJP3503 including the insertion region between the left and right borders of pJP3503. TER Agrtu-NOS denotes the *Agrobacterium tumefaciens* nopaline synthase terminator; Musmu-MGAT2, *Mus Musculus* MGAT2 acyltransferase; PRO CaMV24S-Ex2, cauliflower mosaic virus 35S duplicated enhancer region; TER Glyma-Lectin, *Glycine max* lectin terminator; Arath-WRI1, *Arabidopsis thaliana* WRI1 transcription factor; PRO Arath-Rubisco SSU, *A. thaliana* rubisco small subunit promoter; Sesin-Oleosin, *Sesame indicum* oleosin; Arath-DGAT1, *A. thaliana* DGAT1 acyltransferase

Figure 18. TAG yields in different aged leaves of three wild type tobacco plants (wt1-3) and three pJP3503 primary transformants (4, 29, 21). Leaf stages are indicated by 'G', green; 'YG', yellow-green; 'Y', yellow. Plant stages during sampling were budding, wild type 1; first flowers appearing, wild type 2; flowering, wild type 3; producing seed pods (pJP3503 transformants).

Figure 19A. DNA insert containing expression cassettes for the *Umbelopsis ramanniana* DGAT2A expressed by the *Glycine max* alpha' subunit beta-conglycinin promoter, *Arabidopsis thaliana* WRI1 expressed by the *Glycine max* kunitz trypsin inhibitor 3 promoter and the *Mus musculus* MGAT2 expressed by the *Glycine max*

alpha' subunit beta-conglycinin promoter. Gene coding regions and expression cassettes are excisable by restriction digestion.

Figure 19B. DNA insert containing expression cassettes for the *Arabidopsis thaliana*

- 5 LEC2 and WRI1 transcription factor genes expressed by inducible *Aspergillus* alcA promoters, the *Arabidopsis thaliana* DGAT1 expressed by the constitutive CaMV-35S promoter and the *Aspergillus* alcR gene expressed by the constitutive CsVMV promoter. Expressed of the LEC2 and WRI1 transcription factors is induced by ethanol or an analagous compound.

10

Figure 20. pJP3507 map

Figure 21. pJP3569 map

15

KEY TO THE SEQUENCE LISTING

- SEQ ID NO:1 *Mus musculus* codon optimised MGAT1
 SEQ ID NO:2 *Mus musculus* codon optimised MGAT2
 SEQ ID NO:3 *Ciona intestinalis* codon optimised MGAT1
 20 SEQ ID NO:4 *Tribolium castaneum* codon optimised MGAT1
 SEQ ID NO:5 *Danio rerio* codon optimised MGAT1
 SEQ ID NO:6 *Danio rerio* codon optimised MGAT2
 SEQ ID NO:7 *Homo sapiens* MGAT1 polynucleotide (AF384163)
 SEQ ID NO:8 *Mus musculus* MGAT1 polynucleotide (AF384162)
 25 SEQ ID NO:9 *Pan troglodytes* MGAT1 polynucleotide transcript variant (XM_001166055)
 SEQ ID NO:10 *Pan troglodytes* MGAT1 polynucleotide transcript variant 2 (XM_0526044.2)
 SEQ ID NO:11 *Canis familiaris* MGAT1 polynucleotide (XM_545667.2)
 30 SEQ ID NO:12 *Bos taurus* MGAT1 polynucleotide (NM_001001153.2)
 SEQ ID NO:13 *Rattus norvegicus* MGAT1 polynucleotide (NM_001108803.1)
 SEQ ID NO:14 *Danio rerio* MGAT1 polynucleotide (NM_001122623.1)
 SEQ ID NO:15 *Caenorhabditis elegans* MGAT1 polynucleotide (NM_073012.4)
 SEQ ID NO:16 *Caenorhabditis elegans* MGAT1 polynucleotide (NM_182380.5)
 35 SEQ ID NO:17 *Caenorhabditis elegans* MGAT1 polynucleotide (NM_065258.3)
 SEQ ID NO:18 *Caenorhabditis elegans* MGAT1 polynucleotide (NM_075068.3)
 SEQ ID NO:19 *Caenorhabditis elegans* MGAT1 polynucleotide (NM_072248.3)
 SEQ ID NO:20 *Kluyveromyces lactis* MGAT1 polynucleotide (XM_455588.1)

- SEQ ID NO:21 *Ashbya gossypii* MGAT1 polynucleotide (NM_208895.1)
 SEQ ID NO:22 *Magnaporthe oryzae* MGAT1 polynucleotide (XM_368741.1)
 SEQ ID NO:23 *Ciona intestinalis* MGAT1 polynucleotide (XM_002120843.1)
 SEQ ID NO:24 *Homo sapiens* MGAT2 polynucleotide (AY157608)
 5 SEQ ID NO:25 *Mus musculus* MGAT2 polynucleotide (AY157609)
 SEQ ID NO:26 *Pan troglodytes* MGAT2 polynucleotide (XM_522112.2)
 SEQ ID NO:27 *Canis familiaris* MGAT2 polynucleotide (XM_542304.1)
 SEQ ID NO:28 *Bos taurus* MGAT2 polynucleotide (NM_001099136.1)
 SEQ ID NO:29 *Rattus norvegicus* MGAT2 polynucleotide (NM_001109436.2)
 10 SEQ ID NO:30 *Gallus gallus* MGAT2 polynucleotide (XM_424082.2)
 SEQ ID NO:31 *Danio rerio* MGAT2 polynucleotide (NM_001006083.1)
 SEQ ID NO:32 *Drosophila melanogaster* MGAT2 polynucleotide (NM_136474.2)
 SEQ ID NO:33 *Drosophila melanogaster* MGAT2 polynucleotide (NM_136473.2)
 SEQ ID NO:34 *Drosophila melanogaster* MGAT2 polynucleotide (NM_136475.2)
 15 SEQ ID NO:35 *Anopheles gambiae* MGAT2 polynucleotide (XM_001688709.1)
 SEQ ID NO:36 *Anopheles gambiae* MGAT2 polynucleotide (XM_315985)
 SEQ ID NO:37 *Tribolium castaneum* MGAT2 polynucleotide (XM_970053.1)
 SEQ ID NO:38 *Homo sapiens* MGAT3 polynucleotide (AY229854)
 SEQ ID NO:39 *Pan troglodytes* MGAT3 polynucleotide transcript variant 1
 20 (XM_001154107.1)
 SEQ ID NO:40 *Pan troglodytes* MGAT3 polynucleotide transcript variant 2
 (XM_001154171.1)
 SEQ ID NO:41 *Pan troglodytes* MGAT3 polynucleotide transcript variant 3
 (XM_527842.2)
 25 SEQ ID NO:42 *Canis familiaris* MGAT3 polynucleotide (XM_845212.1)
 SEQ ID NO:43 *Bos taurus* MGAT3 polynucleotide (XM_870406.4)
 SEQ ID NO:44 *Danio rerio* MGAT3 polynucleotide (XM_688413.4)
 SEQ ID NO:45 *Homo sapiens* MGAT1 polypeptide (AAK84178.1)
 SEQ ID NO:46 *Mus musculus* MGAT1 polypeptide (AAK84177.1)
 30 SEQ ID NO:47 *Pan troglodytes* MGAT1 polypeptide isoform 1 (XP_001166055.1)
 SEQ ID NO:48 *Pan troglodytes* MGAT1 polypeptide isoform 2 (XP_526044.2)
 SEQ ID NO:49 *Canis familiaris* MGAT1 polypeptide (XP_545667.2)
 SEQ ID NO:50 *Bos taurus* MGAT1 polypeptide (NP_001001153.1)
 SEQ ID NO:51 *Rattus norvegicus* MGAT1 polypeptide (NP_001102273.1)
 35 SEQ ID NO:52 *Danio rerio* MGAT1 polypeptide (NP_001116095.1)
 SEQ ID NO:53 *Caenorhabditis elegans* MGAT1 polypeptide (NP_505413.1)
 SEQ ID NO:54 *Caenorhabditis elegans* MGAT1 polypeptide (NP_872180.1)
 SEQ ID NO:55 *Caenorhabditis elegans* MGAT1 polypeptide (NP_497659.1)

- SEQ ID NO:56 *Caenorhabditis elegans* MGAT1 polypeptide (NP_507469.1)
 SEQ ID NO:57 *Caenorhabditis elegans* MGAT1 polypeptide (NP_504649.1)
 SEQ ID NO:58 *Kluyveromyces lactis* MGAT1 polypeptide (XP_455588.1)
 SEQ ID NO:59 *Ashbya gossypii* MGAT1 polypeptide (NP_983542.1)
 5 SEQ ID NO:60 *Magnaporthe oryzae* MGAT1 polypeptide (XP_368741.1)
 SEQ ID NO:61 *Ciona intestinalis* MGAT1 polypeptide (XP_002120879)
 SEQ ID NO:62 *Homo sapiens* MGAT2 polypeptide (AA023672.1)
 SEQ ID NO:63 *Mus musculus* MGAT2 polypeptide (AA023673.1)
 SEQ ID NO:64 *Pan troglodytes* MGAT2 polypeptide (XP_522112.2)
 10 SEQ ID NO:65 *Canis familiaris* MGAT2 polypeptide (XP_542304.1)
 SEQ ID NO:66 *Bos taurus* MGAT2 polypeptide (NP_001092606.1)
 SEQ ID NO:67 *Rattus norvegicus* MGAT2 polypeptide (NP_001102906.2)
 SEQ ID NO:68 *Gallus gallus* MGAT2 polypeptide (XP_424082.2)
 SEQ ID NO:69 *Danio rerio* MGAT2 polypeptide (NP_001006083.1)
 15 SEQ ID NO:70 *Drosophila melanogaster* MGAT2 polypeptide (NP_610318.1)
 SEQ ID NO:71 *Drosophila melanogaster* MGAT2 polypeptide (NP_610317.1)
 SEQ ID NO:72 *Drosophila melanogaster* MGAT2 polypeptide (NP_610319.2)
 SEQ ID NO:73 *Anopheles gambiae* MGAT2 polypeptide (XP_001688761)
 SEQ ID NO:74 *Anopheles gambiae* MGAT2 polypeptide (XP_315985.3)
 20 SEQ ID NO:75 *Tribolium castaneum* MGAT2 polypeptide (XP_975146)
 SEQ ID NO:76 *Homo sapiens* MGAT3 polypeptide (AA063579.1)
 SEQ ID NO:77 *Pan troglodytes* MGAT3 polypeptide isoform 1 (XP_001154107.1)
 SEQ ID NO:78 *Pan troglodytes* MGAT3 polypeptide isoform 2 (XP_001154171.1)
 SEQ ID NO:79 *Pan troglodytes* MGAT3 isoform 3 (XP_527842.2)
 25 SEQ ID NO:80 *Canis familiaris* MGAT3 polypeptide (XP_850305.1)
 SEQ ID NO:81 *Bos taurus* MGAT3 polypeptide (XP_875499.3)
 SEQ ID NO:82 *Danio rerio* MGAT3 polypeptide (XP_693505.1)
 SEQ ID NO:83 *Arabidopsis thaliana* DGAT1 polypeptide (CAB44774.1)
 SEQ ID NO:84 *Arabidopsis thaliana* GPAT4 polynucleotide (NM_100043.4)
 30 SEQ ID NO:85 *Arabidopsis thaliana* GPAT6 polynucleotide (NM_129367.3)
 SEQ ID NO:86 *Arabidopsis thaliana* GPAT polynucleotide (AF195115.1)
 SEQ ID NO:87 *Arabidopsis thaliana* GPAT polynucleotide (AY062466.1)
 SEQ ID NO:88 *Oryza sativa* GPAT polynucleotide (AC118133.4)
 SEQ ID NO:89 *Picea sitchensis* GPAT polynucleotide (EF086095.1)
 35 SEQ ID NO:90 *Zea mays* GPAT polynucleotide (BT067649.1)
 SEQ ID NO:91 *Arabidopsis thaliana* GPAT polynucleotide (AK228870.1)
 SEQ ID NO:92 *Oryza sativa* GPAT polynucleotide (AK241033.1)
 SEQ ID NO:93 *Oryza sativa* GPAT polynucleotide (CM000127.1)

- SEQ ID NO:94 *Oryza sativa* GPAT polynucleotide (CM000130.1)
 SEQ ID NO:95 *Oryza sativa* GPAT polynucleotide (CM000139.1)
 SEQ ID NO:96 *Oryza sativa* GPAT polynucleotide (CM000126.1)
 SEQ ID NO:97 *Oryza sativa* GPAT polynucleotide (CM000128.1)
 5 SEQ ID NO:98 *Oryza sativa* GPAT polynucleotide (CM000140.1)
 SEQ ID NO:99 *Selaginella moellendorffii* GPAT polynucleotide (GL377667.1)
 SEQ ID NO:100 *Selaginella moellendorffii* GPAT polynucleotide (GL377667.1)
 SEQ ID NO:101 *Selaginella moellendorffii* GPAT polynucleotide (GL377648.1)
 SEQ ID NO:102 *Selaginella moellendorffii* GPAT polynucleotide (GL377622.1)
 10 SEQ ID NO:103 *Selaginella moellendorffii* GPAT polynucleotide (GL377590.1)
 SEQ ID NO:104 *Selaginella moellendorffii* GPAT polynucleotide (GL377576.1)
 SEQ ID NO:105 *Selaginella moellendorffii* GPAT polynucleotide (GL377576.1)
 SEQ ID NO:106 *Oryza sativa* GPAT polynucleotide (NM_001051374.2)
 SEQ ID NO:107 *Oryza sativa* GPAT polynucleotide (NM_001052203.1)
 15 SEQ ID NO:108: *Zea mays* GPAT8 polynucleotide (NM_001153970.1)
 SEQ ID NO:109: *Zea mays* GPAT polynucleotide (NM_001155835.1)
 SEQ ID NO:110: *Zea mays* GPAT polynucleotide (NM_001174880.1)
 SEQ ID NO:111 *Brassica napus* GPAT4 polynucleotide (JQ666202.1)
 SEQ ID NO:112 *Arabidopsis thaliana* GPAT8 polynucleotide (NM_116264.5)
 20 SEQ ID NO:113 *Physcomitrella patens* GPAT polynucleotide (XM_001764949.1)
 SEQ ID NO:114 *Physcomitrella patens* GPAT polynucleotide (XM_001769619.1)
 SEQ ID NO:115 *Physcomitrella patens* GPAT polynucleotide (XM_001769672.1)
 SEQ ID NO:116 *Physcomitrella patens* GPAT polynucleotide (XM_001771134.1)
 SEQ ID NO:117 *Physcomitrella patens* GPAT polynucleotide (XM_001780481.1)
 25 SEQ ID NO:118 *Vitis vinifera* GPAT polynucleotide (XM_002268477.1)
 SEQ ID NO:119 *Vitis vinifera* GPAT polynucleotide (XM_002275312.1)
 SEQ ID NO:120 *Vitis vinifera* GPAT polynucleotide (XM_002275996.1)
 SEQ ID NO:121 *Vitis vinifera* GPAT polynucleotide (XM_002279055.1)
 SEQ ID NO:122 *Populus trichocarpa* GPAT polynucleotide (XM_002309088.1)
 30 SEQ ID NO:123 *Populus trichocarpa* GPAT polynucleotide (XM_002309240.1)
 SEQ ID NO:124 *Populus trichocarpa* GPAT polynucleotide (XM_002322716.1)
 SEQ ID NO:125 *Populus trichocarpa* GPAT polynucleotide (XM_002323527.1)
 SEQ ID NO:126 *Sorghum bicolor* GPAT polynucleotide (XM_002439842.1)
 SEQ ID NO:127 *Sorghum bicolor* GPAT polynucleotide (XM_002458741.1)
 35 SEQ ID NO:128 *Sorghum bicolor* GPAT polynucleotide (XM_002463871.1)
 SEQ ID NO:129 *Sorghum bicolor* GPAT polynucleotide (XM_002464585.1)
 SEQ ID NO:130 *Ricinus communis* GPAT polynucleotide (XM_002511827.1)
 SEQ ID NO:131 *Ricinus communis* GPAT polynucleotide (XM_002517392.1)

- SEQ ID NO:132 *Ricinus communis* GPAT polynucleotide (XM_002520125.1)
 SEQ ID NO:133 *Arabidopsis lyrata* GPAT polynucleotide (XM_002872909.1)
 SEQ ID NO:134 *Arabidopsis lyrata* GPAT6 polynucleotide (XM_002881518.1)
 SEQ ID NO 135 *Vernicia fordii* putative GPAT8 polynucleotide (FJ479753.1)
- 5 SEQ ID NO 136 *Oryza sativa* GPAT polynucleotide (NM_001057724.1)
 SEQ ID NO:137 *Brassica napus* GPAT4 polynucleotide (JQ666203.1)SEQ ID NO.
 138 *Populus trichocarpa* GPAT polynucleotide (XM_002320102.1)
 SEQ ID NO:139 *Sorghum bicolor* GPAT polynucleotide (XM_002451332.1)
 SEQ ID NO:140 *Ricinus communis* GPAT polynucleotide (XM_002531304.1)
- 10 SEQ ID NO:141 *Arabidopsis lyrata* GPAT4 polynucleotide (XM_002889315.1)
 SEQ ID NO:142 *Arabidopsis thaliana* GPAT1 polynucleotide (NM_100531.2)
 SEQ ID NO 143 *Arabidopsis thaliana* GPAT3 polynucleotide (NM_116426.2)
 SEQ ID NO:144 *Arabidopsis thaliana* GPAT4 polypeptide (NP_171667.1)
 SEQ ID NO:145 *Arabidopsis thaliana* GPAT6 polypeptide (NP_181346.1)
- 15 SEQ ID NO:146 *Arabidopsis thaliana* GPAT polypeptide (AAF02784.1)
 SEQ ID NO:147 *Arabidopsis thaliana* GPAT polypeptide (AAL32544.1)
 SEQ ID NO:148 *Oryza sativa* GPAT polypeptide (AAP03413.1)
 SEQ ID NO:149 *Picea sitchensis* GPAT polypeptide (ABK25381.1)
 SEQ ID NO:150 *Zea mays* GPAT polypeptide (ACN34546.1)
- 20 SEQ NO ID:151 *Arabidopsis thaliana* GPAT polypeptide (BAF00762.1)
 SEQ ID NO:152 *Oryza sativa* GPAT polypeptide (BAH00933.1)
 SEQ ID NO:153 *Oryza sativa* GPAT polypeptide (EAY84189.1)
 SEQ ID NO:154 *Oryza sativa* GPAT polypeptide (EAY98245.1)
 SEQ ID NO:155 *Oryza sativa* GPAT polypeptide (EAZ21484.1)
- 25 SEQ ID NO:156 *Oryza sativa* GPAT polypeptide (EEC71826.1)
 SEQ ID NO:157 *Oryza sativa* GPAT polypeptide (EEC76137.1)
 SEQ ID NO:158 *Oryza sativa* GPAT polypeptide (EEE59882.1)
 SEQ ID NO:159 *Selaginella moellendorffii* GPAT polypeptide (EFJ08963.1)
 SEQ ID NO:160 *Selaginella moellendorffii* GPAT polypeptide (EFJ08964.1)
- 30 SEQ ID NO:161 *Selaginella moellendorffii* GPAT polypeptide (EFJ11200.1)
 SEQ ID NO:162 *Selaginella moellendorffii* GPAT polypeptide (EFJ15664.1)
 SEQ ID NO:163 *Selaginella moellendorffii* GPAT polypeptide (EFJ24086.1)
 SEQ ID NO:164 *Selaginella moellendorffii* GPAT polypeptide (EFJ29816.1)
 SEQ ID NO:165 *Selaginella moellendorffii* GPAT polypeptide (EFJ29817.1)
- 35 SEQ ID NO:166 *Oryza sativa* GPAT polypeptide (NP_001044839.1)
 SEQ ID NO:167 *Oryza sativa* GPAT polypeptide (NP_001045668.1)
 SEQ ID NO:168 *Zea mays* GPAT 8 polypeptide (NP_001147442.1)
 SEQ ID NO:169 *Zea mays* GPAT polypeptide (NP_001149307.1)

- SEQ ID NO:170 *Zea mays* protein GPAT polypeptide (NP_001168351.1)
 SEQ ID NO:171 *Brassica napus* GPAT4 polypeptide (AFH02724.1)
 SEQ ID NO:172 *Arabidopsis thaliana* GPAT8 polypeptide (NP_191950.2)
 SEQ ID NO:173 *Physcomitrella patens* GPAT polypeptide (XP_001765001.1)
 5 SEQ ID NO:174 *Physcomitrella patens* GPAT polypeptide (XP_001769671.1)
 SEQ ID NO:175 *Physcomitrella patens* GPAT polypeptide (XP_001769724.1)
 SEQ ID NO:176 *Physcomitrella patens* GPAT polypeptide (XP_001771186.1)
 SEQ ID NO:177 *Physcomitrella patens* GPAT polypeptide (XP_001780533.1)
 SEQ ID NO:178 *Vitis vinifera* GPAT polypeptide (XP_002268513.1)
 10 SEQ ID NO:179 *Vitis vinifera* GPAT polypeptide (XP_002275348.1)
 SEQ ID NO:180 *Vitis vinifera* GPAT polypeptide (XP_002276032.1)
 SEQ ID NO:181 *Vitis vinifera* GPAT polypeptide (XP_002279091.1)
 SEQ ID NO:182 *Populus trichocarpa* GPAT polypeptide (XP_002309124.1)
 SEQ ID NO:183 *Populus trichocarpa* GPAT polypeptide (XP_002309276.1)
 15 SEQ ID NO:184 *Populus trichocarpa* GPAT polypeptide (XP_002322752.1)
 SEQ ID NO:185 *Populus trichocarpa* GPAT polypeptide (XP_002323563.1)
 SEQ ID NO:186 *Sorghum bicolor* GPAT polypeptide (XP_002439887.1)
 SEQ ID NO:187 *Sorghum bicolor* GPAT polypeptide (XP_002458786.1)
 SEQ ID NO:188 *Sorghum bicolor* GPAT polypeptide (XP_002463916.1)
 20 SEQ ID NO:189 *Sorghum bicolor* GPAT polypeptide (XP_002464630.1)
 SEQ ID NO:190 *Ricinus communis* GPAT polypeptide (XP_002511873.1)
 SEQ ID NO:191 *Ricinus communis* GPAT polypeptide (XP_002517438.1)
 SEQ ID NO:192 *Ricinus communis* GPAT polypeptide (XP_002520171.1)
 SEQ ID NO:193 *Arabidopsis lyrata* GPAT polypeptide (XP_002872955.1)
 25 SEQ ID NO:194 *Arabidopsis lyrata* GPAT6 polypeptide (XP_002881564.1)
 SEQ ID NO:195 *Vernicia fordii* GPAT polypeptide (ACT32032.1)
 SEQ ID NO:196 *Oryza sativa* GPAT polypeptide (NP_001051189.1)
 SEQ ID NO:197 *Brassica napus* GPAT4 polypeptide (AFH02725.1)
 SEQ ID NO:198 *Populus trichocarpa* GPAT polypeptide (XP_002320138.1)
 30 SEQ ID NO:199 *Sorghum bicolor* GPAT polypeptide (XP_002451377.1)
 SEQ ID NO:200 *Ricinus communis* GPAT polypeptide (XP_002531350.1)
 SEQ ID NO:201 *Arabidopsis lyrata* GPAT4 polypeptide (XP_002889361.1)
 SEQ ID NO:202 *Arabidopsis thaliana* GPAT1 polypeptide (NP_563768.1)
 SEQ ID NO:203 *Arabidopsis thaliana* GPAT3 polypeptide (NP_192104.1)
 35 SEQ ID NO:204 *Arabidopsis thaliana* DGAT2 polynucleotide (NM_115011.3)
 SEQ ID NO:205 *Ricinus communis* DGAT2 polynucleotide (AY916129.1)
 SEQ ID NO:206 *Vernicia fordii* DGAT2 polynucleotide (DQ356682.1)
 SEQ ID NO:207 *Mortierella ramanniana* DGAT2 polynucleotide (AF391089.1)

- SEQ ID NO:208 *Homo sapiens* DGAT2 polynucleotide (NM_032564.1)
 SEQ ID NO:209 *Homo sapiens* DGAT2 polynucleotide (NM_001013579.2)
 SEQ ID NO:210 *Bos taurus* DGAT2 polynucleotide (NM_205793.2)
 SEQ ID NO:211 *Mus musculus* DGAT2 polynucleotide (AF384160.1)
 5 SEQ ID NO:212 *Arabidopsis thaliana* DGAT2 polypeptide (NP_566952.1)
 SEQ ID NO:213 *Ricinus communis* DGAT2 polypeptide (AAY16324.1)
 SEQ ID NO:214 *Vernicia fordii* DGAT2 polypeptide (ABC94474.1)
 SEQ ID NO:215 *Mortierella ramanniana* DGAT2 polypeptide (AAK84179.1)
 SEQ ID NO:216 *Homo sapiens* DGAT2 polypeptide (Q96PD7.2)
 10 SEQ ID NO:217 *Homo sapiens* DGAT2 polypeptide (Q58HT5.1)
 SEQ ID NO:218 *Bos taurus* DGAT2 polypeptide (Q70VZ8.1)
 SEQ ID NO:219 *Mus musculus* DGAT2 polypeptide (AAK84175.1)
 SEQ ID NO:220 YFP tripeptide – conserved DGAT2 and/or MGAT1/2 sequence motif
 15 SEQ ID NO:221 HPHG tetrapeptide – conserved DGAT2 and/or MGAT1/2 sequence motif
 SEQ ID NO:222 EPHS tetrapeptide – conserved plant DGAT2 sequence motif
 SEQ ID NO:223 RXGFX(K/R)XAXXXGXXX(L/V)VPXXXFG(E/Q) – long conserved sequence motif of DGAT2 which is part of the putative glycerol
 20 phospholipid domain
 SEQ ID NO:224 FLXLXXXN – conserved sequence motif of mouse DGAT2 and MGAT1/2 which is a putative neutral lipid binding domain
 SEQ ID NO:225 plsC acyltransferase domain (PF01553) of GPAT
 SEQ ID NO:226 HAD-like hydrolase (PF12710) superfamily domain of GPAT
 25 SEQ ID NO:227 Phosphoserine phosphatase domain (PF00702). GPAT4-8 contain a N-terminal region homologous to this domain
 SEQ ID NO:228 Conserved GPAT amino acid sequence GDLVICPEGTTCTREP
 SEQ ID NO:229 Conserved GPAT/phosphatase amino acid sequence (Motif I)
 SEQ ID NO:230 Conserved GPAT/phosphatase amino acid sequence (Motif III)
 30 SEQ ID NO:231 *Arabidopsis thaliana* WRI1 polynucleotide (NM_202701.2)
 SEQ ID NO:232 *Arabidopsis thaliana* WRI1 polynucleotide (NM_001035780.2)
 SEQ ID NO:233 *Arabidopsis thaliana* WRI1 polynucleotide (NM_115292.4)
 SEQ ID NO:234 *Arabidopsis lyrata subsp. lyrata* polynucleotide (XM_002876205.1)
 SEQ ID NO:235 *Brassica napus* WRI1 polynucleotide (DQ370141.1)
 35 SEQ ID NO:236 *Brassica napus* WRI1 polynucleotide (HM370542.1)
 SEQ ID NO:237 *Glycine max* WRI1 polynucleotide (XM_003530322.1)
 SEQ ID NO:238 *Jatropha curcas* WRI1 polynucleotide (JF703666.1)
 SEQ ID NO:239 *Ricinus communis* WRI1 polynucleotide (XM_002525259.1)

- SEQ ID NO:240 *Populus trichocarpa* WRI1 polynucleotide (XM_002316423.1)
 SEQ ID NO:241 *Brachypodium distachyon* WRI1 polynucleotide (XM_003578949.1)
 SEQ ID NO:242 *Hordeum vulgare subsp. vulgare* WRI1 polynucleotide (AK355408.1)
- 5 SEQ ID NO:243 *Sorghum bicolor* WRI1 polynucleotide (XM_002450149.1)
 SEQ ID NO:244 *Zea mays* WRI1 polynucleotide (EU960249.1)
 SEQ ID NO:245 *Brachypodium distachyon* WRI1 polynucleotide (XM_003561141.1)
 SEQ ID NO:246 *Sorghum bicolor* WRI1 polynucleotide (XM_002437774.1)
- 10 SEQ ID NO:247 *Sorghum bicolor* WRI1 polynucleotide (XM_002441399.1)
 SEQ ID NO:248 *Glycine max* WRI1 polynucleotide (XM_003530638.1)
 SEQ ID NO:249 *Glycine max* WRI1 polynucleotide (XM_003553155.1)
 SEQ ID NO:250 *Populus trichocarpa* WRI1 polynucleotide (XM_002315758.1)
 SEQ ID NO:251 *Vitis vinifera* WRI1 polynucleotide (XM_002270113.1)
- 15 SEQ ID NO:252 *Glycine max* WRI1 polynucleotide (XM_003533500.1)
 SEQ ID NO:253 *Glycine max* WRI1 polynucleotide (XM_003551675.1)
 SEQ ID NO:254 *Medicago truncatula* WRI1 polynucleotide (XM_003621069.1)
 SEQ ID NO:255 *Populus trichocarpa* WRI1 polynucleotide (XM_002323800.1)
 SEQ ID NO:256 *Ricinus communis* WRI1 polynucleotide (XM_002517428.1)
- 20 SEQ ID NO:257 *Brachypodium distachyon* WRI1 polynucleotide (XM_003572188.1)
 SEQ ID NO:258 *Sorghum bicolor* WRI1 polynucleotide (XM_002444384.1)
 SEQ ID NO:259 *Zea mays* WRI1 polynucleotide (NM_001176888.1)
 SEQ ID NO:260 *Arabidopsis lyrata subsp. lyrata* WRI1 polynucleotide (XM_002889219.1)
- 25 SEQ ID NO:261 *Arabidopsis thaliana* WRI1 polynucleotide (NM_106619.3)
 SEQ ID NO:262 *Arabidopsis lyrata subsp. lyrata* WRI1 polynucleotide (XM_002890099.1)
 SEQ ID NO:263 *Thellungiella halophila* WRI1 polynucleotide (AK352786.1)
- 30 SEQ ID NO:264 *Arabidopsis thaliana* WRI1 polynucleotide (NM_101474.2)
 SEQ ID NO:265 *Glycine max* WRI1 polynucleotide (XM_003530302.1)
 SEQ ID NO:266 *Brachypodium distachyon* WRI1 polynucleotide (XM_003578094.1)
 SEQ ID NO:267 *Sorghum bicolor* WRI1 polynucleotide (XM_002460191.1)
- 35 SEQ ID NO:268 *Zea mays* WRI1 polynucleotide (NM_001152866.1)
 SEQ ID NO:269 *Glycine max* WRI1 polynucleotide (XM_003519119.1)
 SEQ ID NO:270 *Glycine max* WRI1 polynucleotide (XM_003550628.1)
 SEQ ID NO:271 *Medicago truncatula* WRI1 polynucleotide (XM_003610213.1)

- SEQ ID NO:272 *Glycine max* WRI1 polynucleotide (XM_003523982.1)
 SEQ ID NO:273 *Glycine max* WRI1 polynucleotide (XM_003525901.1)
 SEQ ID NO:274 *Populus trichocarpa* WRI1 polynucleotide (XM_002325075.1)
 SEQ ID NO:275 *Vitis vinifera* WRI1 polynucleotide (XM_002273010.2)
 5 SEQ ID NO:276 *Populus trichocarpa* WRI1 polynucleotide (XM_002303830.1)
 SEQ ID NO:277 *Lupinus angustifolius* WRI1 polynucleotide, partial sequence (NA-080818_Plate14f06.b1)
 SEQ ID NO:278 *Lupinus angustifolius* WRI1 polynucleotide
 SEQ ID NO:279 *Arabidopsis thaliana* WRI1 polypeptide (A8MS57)
 10 SEQ ID NO:280 *Arabidopsis thaliana* WRI1 polypeptide (Q6X5Y6)
 SEQ ID NO:281 *Arabidopsis lyrata* subsp. *lyrata* WRI1 polypeptide (XP_002876251.1)
 SEQ ID NO:282 *Brassica napus* WRI1 polypeptide (ABD16282.1)
 SEQ ID NO:283 *Brassica napus* WRI1 polypeptide (ADO16346.1)
 15 SEQ ID NO:284 *Glycine max* WRI1 polypeptide (XP_003530370.1)
 SEQ ID NO:285 *Jatropha curcas* WRI1 polypeptide (AEO22131.1)
 SEQ ID NO:286 *Ricinus communis* WRI1 polypeptide (XP_002525305.1)
 SEQ ID NO:287 *Populus trichocarpa* WRI1 polypeptide (XP_002316459.1)
 SEQ ID NO:288 *Vitis vinifera* WRI1 polypeptide (CBI29147.3)
 20 SEQ ID NO:289 *Brachypodium distachyon* WRI1 polypeptide (XP_003578997.1)
 SEQ ID NO:290 *Hordeum vulgare* subsp. *vulgare* WRI1 polypeptide (BAJ86627.1)
 SEQ ID NO:291 *Oryza sativa* WRI1 polypeptide (EAY79792.1)
 SEQ ID NO:292 *Sorghum bicolor* WRI1 polypeptide (XP_002450194.1)
 SEQ ID NO:293 *Zea mays* WRI1 polypeptide (ACG32367.1)
 25 SEQ ID NO:294 *Brachypodium distachyon* WRI1 polypeptide (XP_003561189.1)
 SEQ ID NO:295 *Brachypodium sylvaticum* WRI1 polypeptide (ABL85061.1)
 SEQ ID NO:296 *Oryza sativa* WRI1 polypeptide (BAD68417.1)
 SEQ ID NO:297 *Sorghum bicolor* WRI1 polypeptide (XP_002437819.1)
 SEQ ID NO:298 *Sorghum bicolor* WRI1 polypeptide (XP_002441444.1)
 30 SEQ ID NO:299 *Glycine max* WRI1 polypeptide (XP_003530686.1)
 SEQ ID NO:300 *Glycine max* WRI1 polypeptide (XP_003553203.1)
 SEQ ID NO:301 *Populus trichocarpa* WRI1 polypeptide (XP_002315794.1)
 SEQ ID NO:302 *Vitis vinifera* WRI1 polypeptide (XP_002270149.1)
 SEQ ID NO:303 *Glycine max* WRI1 polypeptide (XP_003533548.1)
 35 SEQ ID NO:304 *Glycine max* WRI1 polypeptide (XP_003551723.1)
 SEQ ID NO:305 *Medicago truncatula* WRI1 polypeptide (XP_003621117.1)
 SEQ ID NO:306 *Populus trichocarpa* WRI1 polypeptide (XP_002323836.1)
 SEQ ID NO:307 *Ricinus communis* WRI1 polypeptide (XP_002517474.1)

- SEQ ID NO:308 *Vitis vinifera* WRI1 polypeptide (CAN79925.1)
 SEQ ID NO:309 *Brachypodium distachyon* WRI1 polypeptide (XP_003572236.1)
 SEQ ID NO:310 *Oryza sativa* WRI1 polypeptide (BAD10030.1)
 SEQ ID NO:311 *Sorghum bicolor* WRI1 polypeptide (XP_002444429.1)
 5 SEQ ID NO:312 *Zea mays* WRI1 polypeptide (NP_001170359.1)
 SEQ ID NO:313 *Arabidopsis lyrata subsp. lyrata* WRI1 polypeptide (XP_002889265.1)
 SEQ ID NO:314 *Arabidopsis thaliana* WRI1 polypeptide (AAF68121.1)
 SEQ ID NO:315 *Arabidopsis thaliana* WRI1 polypeptide (NP_178088.2)
 10 SEQ ID NO:316 *Arabidopsis lyrata subsp. lyrata* WRI1 polypeptide (XP_002890145.1)
 SEQ ID NO:317 *Thellungiella halophila* WRI1 polypeptide (BAJ33872.1)
 SEQ ID NO:318 *Arabidopsis thaliana* WRI1 polypeptide (NP_563990.1)
 SEQ ID NO:319 *Glycine max* WRI1 polypeptide (XP_003530350.1)
 15 SEQ ID NO:320 *Brachypodium distachyon* WRI1 polypeptide (XP_003578142.1)
 SEQ ID NO:321 *Oryza sativa* WRI1 polypeptide (EAO9147.1)
 SEQ ID NO:322 *Sorghum bicolor* WRI1 polypeptide (XP_002460236.1)
 SEQ ID NO:323 *Zea mays* WRI1 polypeptide (NP_001146338.1)
 SEQ ID NO:324 *Glycine max* WRI1 polypeptide (XP_003519167.1)
 20 SEQ ID NO:325 *Glycine max* WRI1 polypeptide (XP_003550676.1)
 SEQ ID NO:326 *Medicago truncatula* WRI1 polypeptide (XP_003610261.1)
 SEQ ID NO:327 *Glycine max* WRI1 polypeptide (XP_003524030.1)
 SEQ ID NO:328 *Glycine max* WRI1 polypeptide (XP_003525949.1)
 SEQ ID NO:329 *Populus trichocarpa* WRI1 polypeptide (XP_002325111.1)
 25 SEQ ID NO:330 *Vitis vinifera* WRI1 polypeptide (CBI36586.3)
 SEQ ID NO:331 *Vitis vinifera* WRI1 polypeptide (XP_002273046.2)
 SEQ ID NO:332 *Populus trichocarpa* WRI1 polypeptide (XP_002303866.1)
 SEQ ID NO:333 *Vitis vinifera* WRI1 polypeptide (CBI25261.3)
 SEQ ID NO:334 Sorbi-WRL1
 30 SEQ ID NO: 335 Lupan-WRL1
 SEQ ID NO:336 Ricco-WRL1
 SEQ ID NO:337 *Lupin angustifolius* WRI1 polypeptide
 SEQ ID NO:338 *Aspergillus fumigatus* DGAT polynucleotide (XM_750079.1)
 SEQ ID NO:339 *Ricinus communis* DGAT polynucleotide (AY366496.1)
 35 SEQ ID NO:340 *Vernicia fordii* DGAT1 polynucleotide (DQ356680.1)
 SEQ ID NO:341 *Vernonia galamensis* DGAT1 polynucleotide (EF653276.1)
 SEQ ID NO:342 *Vernonia galamensis* DGAT1 polynucleotide (EF653277.1)
 SEQ ID NO:343 *Euonymus alatus* DGAT1 polynucleotide (AY751297.1)

- SEQ ID NO:344 *Caenorhabditis elegans* DGAT1 polynucleotide (AF221132.1)
 SEQ ID NO:345 *Rattus norvegicus* DGAT1 polynucleotide (NM_053437.1)
 SEQ ID NO:346 *Homo sapiens* DGAT1 polynucleotide (NM_012079.4)
 SEQ ID NO:347 *Aspergillus fumigatus* DGAT1 polypeptide (XP_755172.1)
 5 SEQ ID NO:348 *Ricinus communis* DGAT1 polypeptide (AAR11479.1)
 SEQ ID NO:349 *Vernicia fordii* DGAT1 polypeptide (ABC94472.1)
 SEQ ID NO:350 *Vernonia galamensis* DGAT1 polypeptide (ABV21945.1)
 SEQ ID NO:351 *Vernonia galamensis* DGAT1 polypeptide (ABV21946.1)
 SEQ ID NO:352 *Euonymus alatus* DGAT1 polypeptide (AAV31083.1)
 10 SEQ ID NO:353 *Caenorhabditis elegans* DGAT1 polypeptide (AAF82410.1)
 SEQ ID NO:354 *Rattus norvegicus* DGAT1 polypeptide (NP_445889.1)
 SEQ ID NO:355 *Homo sapiens* DGAT1 polypeptide (NP_036211.2)
 SEQ ID NO:356 WRI1 motif (R G V T/S R H R W T G R)
 SEQ ID NO:357 WRI1 motif (F/Y E A H L W D K)
 15 SEQ ID NO:358 WRI1 motif (D L A A L K Y W G)
 SEQ ID NO:359 WRI1 motif (S X G F S/A R G X)
 SEQ ID NO:360 WRI1 motif (H H H/Q N G R/K W E A R I G R/K V)
 SEQ ID NO:361 WRI1 motif (Q E E A A X Y D)
 SEQ ID NO:362 *Brassica napus* oleosin polypeptide (CAA57545.1)
 20 SEQ ID NO:363 *Brassica napus* oleosin S1-1 polypeptide (ACG69504.1)
 SEQ ID NO:364 *Brassica napus* oleosin S2-1 polypeptide (ACG69503.1)
 SEQ ID NO:365 *Brassica napus* oleosin S3-1 polypeptide (ACG69513.1)
 SEQ ID NO:366 *Brassica napus* oleosin S4-1 polypeptide (ACG69507.1)
 SEQ ID NO:367 *Brassica napus* oleosin S5-1 polypeptide (ACG69511.1)
 25 SEQ ID NO:368 *Arachis hypogaea* oleosin 1 polypeptide (AAZ20276.1)
 SEQ ID NO:369 *Arachis hypogaea* oleosin 2 polypeptide (AAU21500.1)
 SEQ ID NO:370 *Arachis hypogaea* oleosin 3 polypeptide (AAU21501.1)
 SEQ ID NO:371 *Arachis hypogaea* oleosin 5 polypeptide (ABC96763.1)
 SEQ ID NO:372 *Ricinus communis* oleosin 1 polypeptide (EEF40948.1)
 30 SEQ ID NO:373 *Ricinus communis* oleosin 2 polypeptide (EEF51616.1)
 SEQ ID NO:374 *Glycine max* oleosin isoform a polypeptide (P29530.2)
 SEQ ID NO:375 *Glycine max* oleosin isoform b polypeptide (P29531.1)
 SEQ ID NO:376 *Linum usitatissimum* oleosin low molecular weight isoform polypeptide (ABB01622.1)
 35 SEQ ID NO:377 amino acid sequence of *Linum usitatissimum* oleosin high molecular weight isoform polypeptide (ABB01624.1)
 SEQ ID NO:378 *Helianthus annuus* oleosin polypeptide (CAA44224.1)
 SEQ ID NO:379 *Zea mays* oleosin polypeptide (NP_001105338.1)

- SEQ ID NO:380 *Brassica napus* steroleosin polypeptide (ABM30178.1)
 SEQ ID NO:381 *Brassica napus* steroleosin SLO1-1 polypeptide (ACG69522.1)
 SEQ ID NO:382 *Brassica napus* steroleosin SLO2-1 polypeptide (ACG69525.1)
 SEQ ID NO:383 *Sesamum indicum* steroleosin polypeptide (AAL13315.1)
 5 SEQ ID NO:384 *Zea mays* steroleosin polypeptide (NP_001152614.1)
 SEQ ID NO:385 *Brassica napus* caleosin CLO-1 polypeptide (ACG69529.1)
 SEQ ID NO:386 *Brassica napus* caleosin CLO-3 polypeptide (ACG69527.1)
 SEQ ID NO:387 *Sesamum indicum* caleosin polypeptide (AAF13743.1)
 SEQ ID NO:388 *Zea mays* caleosin polypeptide (NP_001151906.1)
 10 SEQ ID NO:389 *Brassica napus* oleosin polynucleotide (X82020.1)
 SEQ ID NO:390 *Brassica napus* oleosin S1-1 polynucleotide (EU678256.1)
 SEQ ID NO:391 *Brassica napus* oleosin S2-1 polynucleotide (EU678255.1)
 SEQ ID NO:392 *Brassica napus* oleosin S3-1 polynucleotide (EU678265.1)
 SEQ ID NO:393 *Brassica napus* oleosin S4-1 polynucleotide (EU678259.1)
 15 SEQ ID NO:394 *Brassica napus* oleosin S5-1 polynucleotide (EU678263.1)
 SEQ ID NO:395 *Arachis hypogaea* oleosin 1 polynucleotide (DQ097716.1)
 SEQ ID NO:396 *Arachis hypogaea* oleosin 2 polynucleotide (AY722695.1)
 SEQ ID NO:397 *Arachis hypogaea* oleosin 3 polynucleotide (AY722696.1)
 SEQ ID NO:398 *Arachis hypogaea* oleosin 5 polynucleotide (DQ368496.1)
 20 SEQ ID NO:399 *Helianthus annuus* oleosin polynucleotide (X62352.1)
 SEQ ID NO:400 *Zea mays* oleosin polynucleotide (NM_001111868.1)
 SEQ ID NO:401 *Brassica napus* steroleosin polynucleotide (EF143915.1)
 SEQ ID NO:402 *Brassica napus* steroleosin SLO1-1 polynucleotide (EU678274.1)
 SEQ ID NO:403 *Brassica napus* steroleosin SLO2-1 polynucleotide (EU678277.1)
 25 SEQ ID NO:404 *Zea mays* steroleosin polynucleotide (NM_001159142.1)
 SEQ ID NO:405 *Brassica napus* caleosin CLO-1 polynucleotide (EU678281.1)
 SEQ ID NO:406 *Brassica napus* caleosin CLO-3 polynucleotide (EU678279.1)
 SEQ ID NO:407 *Sesamum indicum* caleosin polynucleotide (AF109921.1)
 SEQ ID NO:408 *Zea mays* caleosin polynucleotide (NM_001158434.1)
 30 SEQ ID NO:409 pJP3502 entire vector sequence (three-gene)
 SEQ ID NO:410 pJP3503 entire vector sequence (four-gene)
 SEQ ID NO:411 pJP3502 TDNA (inserted into genome) sequence
 SEQ ID NO:412 pJP3503 TDNA (inserted into genome) sequence
 SEQ ID NO:413 pJP3507 vector sequence
 35 SEQ ID NO:414 Linker sequence
 SEQ ID NO:415 Soybean Synergy
 SEQ ID NO:416 12ABFJYC_pJP3569_insert

- SEQ ID NO:417 Partial *N. benthamiana* CGI-58 sequence selected for hpRNAi silencing (pTV46)
- SEQ ID NO:418 Partial *N. tabacum* AGPase sequence selected for hpRNAi silencing (pTV35)
- 5 SEQ ID NO:419 GX SXG lipase motif
- SEQ ID NO:420 HX(4)D acyltransferase motif
- SEQ ID NO:421 VX(3)HGF probable lipid binding motif
- SEQ ID NO:422 *Arabidopsis thaliana* CGI58 polynucleotide (NM_118548.1)
- SEQ ID NO:423: *Brachypodium distachyon* CGI58 polynucleotide
- 10 (XM_003578402.1)
- SEQ ID NO:424 *Glycine max* CGI58 polynucleotide (XM_003523590.1)
- SEQ ID NO:425 *Zea mays* CGI58 polynucleotide (NM_001155541.1)
- SEQ ID NO:426 *Sorghum bicolor* CGI58 polynucleotide (XM_002460493.1)
- SEQ ID NO:427 *Ricinus communis* CGI58 polynucleotide (XM_002510439.1)
- 15 SEQ ID NO:428 *Medicago truncatula* CGI58 polynucleotide (XM_003603685.1)
- SEQ ID NO:429 *Arabidopsis thaliana* CGI58 polypeptide (NP_194147.2)
- SEQ ID NO:430 *Brachypodium distachyon* CGI58 polypeptide (XP_003578450.1)
- SEQ ID NO:431 *Glycine max* CGI58 polypeptide (XP_003523638.1)
- SEQ ID NO:432 *Zea Mays* CGI58 polypeptide (NP_001149013.1)
- 20 SEQ ID NO:433 *Sorghum bicolor* CGI58 polypeptide (XP_002460538.1)
- SEQ ID NO:434 *Ricinus communis* CGI58 polypeptide (XP_002510485.1)
- SEQ ID NO:435 *Medicago truncatula* CGI58 polypeptide (XP_003603733.1)
- SEQ ID NO:436 *Oryza sativa* CGI58 polypeptide (EAZ09782.1)
- SEQ ID NO:437 *Arabidopsis thaliana* LEC2 polynucleotide (NM_102595.2)
- 25 SEQ ID NO:438 *Medicago truncatula* LEC2 polynucleotide (X60387.1)
- SEQ ID NO:439 *Brassica napus* LEC2 polynucleotide (HM370539.1)
- SEQ ID NO:440 *Arabidopsis thaliana* BBM polynucleotide (NM_121749.2)
- SEQ ID NO:441 *Medicago truncatula* BBM polynucleotide (AY899909.1)
- SEQ ID NO:442 *Arabidopsis thaliana* LEC2 polypeptide (NP_564304.1)
- 30 SEQ ID NO:443 *Medicago truncatula* LEC2 polypeptide (CAA42938.1)
- SEQ ID NO:444 *Brassica napus* LEC2 polypeptide (ADO16343.1)
- SEQ ID NO:445 *Arabidopsis thaliana* BBM polypeptide (NP_197245.2)
- SEQ ID NO:446 *Medicago truncatula* BBM polypeptide (AAW82334.1)
- SEQ ID NO:447 Inducible *Aspergillus niger* alcA promoter
- 35 SEQ ID NO:448 AlcR inducer that activates the AlcA promoter in the presence of ethanol

DETAILED DESCRIPTION OF THE INVENTION**General Techniques and Definitions**

Unless specifically defined otherwise, all technical and scientific terms used herein shall be taken to have the same meaning as commonly understood by one of
5 ordinary skill in the art (e.g., in cell culture, molecular genetics, immunology, immunohistochemistry, protein chemistry, lipid and fatty acid chemistry, biofuel production, and biochemistry).

Unless otherwise indicated, the recombinant protein, cell culture, and immunological techniques utilized in the present invention are standard procedures,
10 well known to those skilled in the art. Such techniques are described and explained throughout the literature in sources such as, J. Perbal, A Practical Guide to Molecular Cloning, John Wiley and Sons (1984), J. Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbour Laboratory Press (1989), T.A. Brown (editor), Essential Molecular Biology: A Practical Approach, Volumes 1 and 2, IRL
15 Press (1991), D.M. Glover and B.D. Hames (editors), DNA Cloning: A Practical Approach, Volumes 1-4, IRL Press (1995 and 1996), F.M. Ausubel et al. (editors), Current Protocols in Molecular Biology, Greene Pub. Associates and Wiley-Interscience (1988, including all updates until present), Ed Harlow and David Lane (editors) Antibodies: A Laboratory Manual, Cold Spring Harbour Laboratory, (1988),
20 and J.E. Coligan et al. (editors) Current Protocols in Immunology, John Wiley & Sons (including all updates until present).

Selected Definitions

The term "transgenic non-human organism" refers to, for example, a whole
25 plant, alga, non-human animal, or an organism suitable for fermentation such as a yeast or fungus, comprising an exogenous polynucleotide (transgene) or an exogenous polypeptide. In an embodiment, the transgenic non-human organism is not an animal or part thereof. In one embodiment, the transgenic non-human organism is a phototrophic organism (for example, a plant or alga) capable of obtaining energy from
30 sunlight to synthesize organic compounds for nutrition. In another embodiment, the transgenic non-human organism is a photosynthetic bacterium.

The term "exogenous" in the context of a polynucleotide or polypeptide refers to the polynucleotide or polypeptide when present in a cell which does not naturally comprise the polynucleotide or polypeptide. Such a cell is referred to herein as a
35 "recombinant cell" or a "transgenic cell". In an embodiment, the exogenous polynucleotide or polypeptide is from a different genus to the cell comprising the exogenous polynucleotide or polypeptide. In another embodiment, the exogenous polynucleotide or polypeptide is from a different species. In one embodiment the

exogenous polynucleotide or polypeptide is expressed in a host plant or plant cell and the exogenous polynucleotide or polypeptide is from a different species or genus. The exogenous polynucleotide or polypeptide may be non-naturally occurring, such as for example, a synthetic DNA molecule which has been produced by recombinant DNA methods. The DNA molecule may, often preferably, include a protein coding region which has been codon-optimised for expression in the cell, thereby producing a polypeptide which has the same amino acid sequence as a naturally occurring polypeptide, even though the nucleotide sequence of the protein coding region is non-naturally occurring. The exogenous polynucleotide may encode, or the exogenous polypeptide may be: a diacylglycerol acyltransferase (DGAT) such as a DGAT1 or a DGAT2, a glycerol-3-phosphate acyltransferase (GPAT) such as a GPAT which is capable of synthesising MAG, a Wrinkled 1 (WRI1) transcription factor, an Oleosin, or a silencing suppressor polypeptide. In one embodiment, the exogenous polypeptide is an exogenous MGAT such as an MGAT1 or an MGAT2.

As used herein, the term "extracted lipid" refers to a composition extracted from a transgenic organism or part thereof which comprises at least 60% (w/w) lipid.

As used herein, the term "non-polar lipid" refers to fatty acids and derivatives thereof which are soluble in organic solvents but insoluble in water. The fatty acids may be free fatty acids and/or in an esterified form. Examples of esterified forms include, but are not limited to, triacylglycerol (TAG), diacylglycerol (DAG), monoacylglycerol (MAG). Non-polar lipids also include sterols, sterol esters and wax esters. Non-polar lipids are also known as "neutral lipids". Non-polar lipid is typically a liquid at room temperature. Preferably, the non-polar lipid predominantly (>50%) comprises fatty acids that are at least 16 carbons in length. More preferably, at least 50% of the total fatty acids in the non-polar lipid are C18 fatty acids for example, oleic acid. In an embodiment, at least 50%, more preferably at least 70%, more preferably at least 80%, more preferably at least 90%, more preferably at least 91%, more preferably at least 92%, more preferably at least 93%, more preferably at least 94%, more preferably at least 95%, more preferably at least 96%, more preferably at least 97%, more preferably at least 98%, more preferably at least 99% of the fatty acids in non-polar lipid of the invention can be found as TAG. The non-polar lipid may be further purified or treated, for example by hydrolysis with a strong base to release the free fatty acid, or by fractionation, distillation, or the like. Non-polar lipid may be present in or obtained from plant parts such as seed, leaves or fruit, from recombinant cells or from non-human organisms such as yeast. Non-polar lipid of the invention may form part of "seedoil" if it is obtained from seed. The free and esterified sterol (for example, sitosterol, campesterol, stigmasterol, brassicasterol, Δ^5 -avenasterol, sitostanol, campestanol, and cholesterol) concentrations in the extracted

lipid may be as described in Phillips et al., 2002. Sterols in plant oils are present as free alcohols, esters with fatty acids (esterified sterols), glycosides and acylated glycosides of sterols. Sterol concentrations in naturally occurring vegetable oils (seedoils) ranges up to a maximum of about 1100mg/100g. Hydrogenated palm oil
5 has one of the lowest concentrations of naturally occurring vegetable oils at about 60mg/100g. The recovered or extracted seedoils of the invention preferably have between about 100 and about 1000mg total sterol/100g of oil. For use as food or feed, it is preferred that sterols are present primarily as free or esterified forms rather than glycosylated forms. In the seedoils of the present invention, preferably at least 50%
10 of the sterols in the oils are present as esterified sterols, except for soybean seedoil which has about 25% of the sterols esterified. The canola seedoil and rapeseed oil of the invention preferably have between about 500 and about 800 mg total sterol/100g, with sitosterol the main sterol and campesterol the next most abundant. The corn seedoil of the invention preferably has between about 600 and about 800 mg total
15 sterol/100g, with sitosterol the main sterol. The soybean seedoil of the invention preferably has between about 150 and about 350 mg total sterol/100g, with sitosterol the main sterol and stigmasterol the next most abundant, and with more free sterol than esterified sterol. The cottonseed oil of the invention preferably has between about 200 and about 350 mg total sterol/100g, with sitosterol the main sterol. The
20 coconut oil and palm oil of the invention preferably have between about 50 and about 100mg total sterol/100g, with sitosterol the main sterol. The safflower seedoil of the invention preferably has between about 150 and about 250mg total sterol/100g, with sitosterol the main sterol. The peanut seedoil of the invention preferably has between about 100 and about 200mg total sterol/100g, with sitosterol the main sterol. The
25 sesame seedoil of the invention preferably has between about 400 and about 600mg total sterol/100g, with sitosterol the main sterol. The sunflower seedoil of the invention preferably has between about 200 and 400mg total sterol/100g, with sitosterol the main sterol. Oils obtained from vegetative plant parts according to the invention preferably have less than 200mg total sterol/100g, more preferably less than
30 100mg total sterol/100 g, and most preferably less than 50mg total sterols/100g, with the majority of the sterols being free sterols.

As used herein, the term "seedoil" refers to a composition obtained from the seed/grain of a plant which comprises at least 60% (w/w) lipid, or obtainable from the seed/grain if the seedoil is still present in the seed/grain. That is, seedoil of the
35 invention includes seedoil which is present in the seed/grain or portion thereof, as well as seedoil which has been extracted from the seed/grain. The seedoil is preferably extracted seedoil. Seedoil is typically a liquid at room temperature. Preferably, the total fatty acid (TFA) content in the seedoil predominantly (>50%)

comprises fatty acids that are at least 16 carbons in length. More preferably, at least 50% of the total fatty acids in the seedoil are C18 fatty acids for example, oleic acid. The fatty acids are typically in an esterified form such as for example, TAG, DAG, acyl-CoA or phospholipid. The fatty acids may be free fatty acids and/or in an esterified form. In an embodiment, at least 50%, more preferably at least 70%, more preferably at least 80%, more preferably at least 90%, more preferably at least 91%, more preferably at least 92%, more preferably at least 93%, more preferably at least 94%, more preferably at least 95%, more preferably at least 96%, more preferably at least 97%, more preferably at least 98%, more preferably at least 99% of the fatty acids in seedoil of the invention can be found as TAG. In an embodiment, seedoil of the invention is "substantially purified" or "purified" oil that has been separated from one or more other lipids, nucleic acids, polypeptides, or other contaminating molecules with which it is associated in the seed or in a crude extract. It is preferred that the substantially purified seedoil is at least 60% free, more preferably at least 75% free, and more preferably, at least 90% free from other components with which it is associated in the seed or extract. Seedoil of the invention may further comprise non-fatty acid molecules such as, but not limited to, sterols. In an embodiment, the seedoil is canola oil (*Brassica* sp. such as *Brassica carinata*, *Brassica juncea*, *Brassica napobrassica*, *Brassica napus*) mustard oil (*Brassica juncea*), other *Brassica* oil (e.g., *Brassica napobrassica*, *Brassica camelina*), sunflower oil (*Helianthus* sp. such as *Helianthus annuus*), linseed oil (*Linum usitatissimum*), soybean oil (*Glycine max*), safflower oil (*Carthamus tinctorius*), corn oil (*Zea mays*), tobacco oil (*Nicotiana* sp. such as *Nicotiana tabacum* or *Nicotiana benthamiana*), peanut oil (*Arachis hypogaea*), palm oil (*Elaeis guineensis*), cottonseed oil (*Gossypium hirsutum*), coconut oil (*Cocos nucifera*), avocado oil (*Persea americana*), olive oil (*Olea europaea*), cashew oil (*Anacardium occidentale*), macadamia oil (*Macadamia intergrifolia*), almond oil (*Prunus amygdalus*), oat seed oil (*Avena sativa*), rice oil (*Oryza* sp. such as *Oryza sativa* and *Oryza glaberrima*), *Arabidopsis* seed oil (*Arabidopsis thaliana*), or oil from the seed of *Acrocomia aculeata* (macauba palm), *Aracinis hypogaea* (peanut), *Astrocaryum murumuru* (murumuru), *Astrocaryum vulgare* (tucumã), *Attalea geraensis* (Indaiá-rateiro), *Attalea humilis* (American oil palm), *Attalea oleifera* (andaiá), *Attalea phalerata* (uricuri), *Attalea speciosa* (babassu), *Beta vulgaris* (sugar beet), *Camelina sativa* (false flax), *Caryocar brasiliense* (pequi), *Crambe abyssinica* (Abyssinian kale), *Cucumis melo* (melon), *Hordeum vulgare* (barley), *Jatropha curcas* (physic nut), *Joannesia princeps* (arara nut-tree), *Licania rigida* (oiticica), *Lupinus angustifolius* (lupin), *Mauritia flexuosa* (buriti palm), *Maximiliana maripa* (inaja palm), *Miscanthus* sp. such as *Miscanthus x giganteus* and *Miscanthus sinensis*, *Oenocarpus bacaba* (bacaba-do-azeite),

Oenocarpus bataua (patauã), *Oenocarpus distichus* (bacaba-de-leque), *Panicum virgatum* (switchgrass), *Paraqueiba paraensis* (mari), *Persea amencana* (avocado), *Pongamia pinnata* (Indian beech), *Populus trichocarpa*, *Ricinus communis* (castor), *Saccharum sp.* (sugarcane), *Sesamum indicum* (sesame), *Solanum tuberosum* (potato),

5 *Sorghum sp.* such as *Sorghum bicolor*, *Sorghum vulgare*, *Theobroma grandiflorum* (cupuassu), *Trifolium sp.*, *Trithrinax brasiliensis* (Brazilian needle palm) and *Triticum sp.* (wheat) such as *Triticum aestivum*. Seedoil may be extracted from seed/grain by any method known in the art. This typically involves extraction with nonpolar solvents such as diethyl ether, petroleum ether, chloroform/methanol or butanol

10 mixtures, generally associated with first crushing of the seeds. Lipids associated with the starch in the grain may be extracted with water-saturated butanol. The seedoil may be "de-gummed" by methods known in the art to remove polysaccharides or treated in other ways to remove contaminants or improve purity, stability, or colour. The TAGs and other esters in the seedoil may be hydrolysed to release free fatty

15 acids, or the seedoil hydrogenated, treated chemically, or enzymatically as known in the art.

As used herein, the term "fatty acid" refers to a carboxylic acid with a long aliphatic tail of at least 8 carbon atoms in length, either saturated or unsaturated. Typically, fatty acids have a carbon-carbon bonded chain of at least 12 carbons in

20 length. Most naturally occurring fatty acids have an even number of carbon atoms because their biosynthesis involves acetate which has two carbon atoms. The fatty acids may be in a free state (non-esterified) or in an esterified form such as part of a TAG, DAG, MAG, acyl-CoA (thio-ester) bound, or other covalently bound form. When covalently bound in an esterified form, the fatty acid is referred to herein as an

25 "acyl" group. The fatty acid may be esterified as a phospholipid such as a phosphatidylcholine (PC), phosphatidylethanolamine, phosphatidylserine, phosphatidylglycerol, phosphatidylinositol, or diphosphatidylglycerol. Saturated fatty acids do not contain any double bonds or other functional groups along the chain. The term "saturated" refers to hydrogen, in that all carbons (apart from the carboxylic

30 acid [-COOH] group) contain as many hydrogens as possible. In other words, the omega (ω) end contains 3 hydrogens (CH₃-) and each carbon within the chain contains 2 hydrogens (-CH₂-). Unsaturated fatty acids are of similar form to saturated fatty acids, except that one or more alkene functional groups exist along the chain, with each alkene substituting a singly-bonded "-CH₂-CH₂-" part of the chain with a

35 doubly-bonded "-CH=CH-" portion (that is, a carbon double bonded to another carbon). The two next carbon atoms in the chain that are bound to either side of the double bond can occur in a *cis* or *trans* configuration.

As used herein, the terms "polyunsaturated fatty acid" or "PUFA" refer to a fatty acid which comprises at least 12 carbon atoms in its carbon chain and at least two alkene groups (carbon-carbon double bonds). The PUFA content of the vegetative plant part, or the non-human organism or part thereof of the invention may be increased or decreased depending on the combination of exogenous polynucleotides expressed in the vegetative plant part, or non-human organism or part thereof, or seed of the invention. For example, when an MGAT is expressed the PUFA level typically increases, whereas when DGAT1 is expressed alone or in combination with WRI1, the PUFA level is typically decreased due to an increase in the level of oleic acid. Furthermore, if $\Delta 12$ desaturase activity is reduced, for example by silencing an endogenous $\Delta 12$ desaturase, PUFA content is unlikely to increase in the absence of an exogenous polynucleotide encoding a different $\Delta 12$ desaturase.

"Monoacylglyceride" or "MAG" is glyceride in which the glycerol is esterified with one fatty acid. As used herein, MAG comprises a hydroxyl group at an *sn*-1/3 (also referred to herein as *sn*-1 MAG or 1-MAG or 1/3-MAG) or *sn*-2 position (also referred to herein as 2-MAG), and therefore MAG does not include phosphorylated molecules such as PA or PC. MAG is thus a component of neutral lipids in a cell.

"Diacylglyceride" or "DAG" is glyceride in which the glycerol is esterified with two fatty acids which may be the same or, preferably, different. As used herein, DAG comprises a hydroxyl group at a *sn*-1,3 or *sn*-2 position, and therefore DAG does not include phosphorylated molecules such as PA or PC. DAG is thus a component of neutral lipids in a cell. In the Kennedy pathway of DAG synthesis (Figure 1), the precursor *sn*-glycerol-3-phosphate (G-3-P) is esterified to two acyl groups, each coming from a fatty acid coenzyme A ester, in a first reaction catalysed by a glycerol-3-phosphate acyltransferase (GPAT) at position *sn*-1 to form LysoPA, followed by a second acylation at position *sn*-2 catalysed by a lysophosphatidic acid acyltransferase (LPAAT) to form phosphatidic acid (PA). This intermediate is then de-phosphorylated to form DAG. In an alternative anabolic pathway (Figure 1), DAG may be formed by the acylation of either *sn*-1 MAG or preferably *sn*-2 MAG, catalysed by MGAT. DAG may also be formed from TAG by removal of an acyl group by a lipase, or from PC essentially by removal of a choline headgroup by any of the enzymes CPT, PDCT or PLC (Figure 1).

"Triacylglyceride" or "TAG" is glyceride in which the glycerol is esterified with three fatty acids which may be the same (e.g. as in tri-olein) or, more commonly, different. In the Kennedy pathway of TAG synthesis, DAG is formed as described above, and then a third acyl group is esterified to the glycerol backbone by the activity of DGAT. Alternative pathways for formation of TAG include one catalysed by the enzyme PDAT and the MGAT pathway described herein.

As used herein, the term "acyltransferase" refers to a protein which is capable of transferring an acyl group from acyl-CoA onto a substrate and includes MGATs, GPATs and DGATs.

As used herein, the term "Wrinkled 1" or "WRI1" or "WRL1" refers to a transcription factor of the AP2/ERWEBP class which regulates the expression of several enzymes involved in glycolysis and *de novo* fatty acid biosynthesis. WRI1 has two plant-specific (AP2/EREB) DNA-binding domains. WRI1 in at least *Arabidopsis* also regulates the breakdown of sucrose via glycolysis thereby regulating the supply of precursors for fatty acid biosynthesis. In other words, it controls the carbon flow from the photosynthate to storage lipids. *wri1* mutants have wrinkled seed phenotype, due to a defect in the incorporation of sucrose and glucose into TAGs.

Examples of genes which are transcribed by WRI1 include, but are not limited to, one or more, preferably all, of pyruvate kinase (At5g52920, At3g22960), pyruvate dehydrogenase (PDH) E1alpha subunit (At1g01090), acetyl-CoA carboxylase (ACCase), BCCP2 subunit (At5g15530), enoyl-ACP reductase (At2g05990; EAR), phosphoglycerate mutase (At1g22170), cytosolic fructokinase, and cytosolic phosphoglycerate mutase, sucrose synthase (SuSy) (see, for example, Liu et al., 2010b; Baud et al., 2007; Ruuska et al., 2002).

WRL1 contains the conserved domain AP2 (cd00018). AP2 is a DNA-binding domain found in transcription regulators in plants such as APETALA2 and EREBP (ethylene responsive element binding protein). In EREBPs the domain specifically binds to the 11bp GCC box of the ethylene response element (ERE), a promoter element essential for ethylene responsiveness. EREBPs and the C-repeat binding factor CBF1, which is involved in stress response, contain a single copy of the AP2 domain. APETALA2-like proteins, which play a role in plant development contain two copies.

Other sequence motifs in WRI1 and its functional homologs include:

1. R G V T/S R H R W T G R (SEQ ID NO:356).
2. F/Y E A H L W D K (SEQ ID NO:357).
3. D L A A L K Y W G (SEQ ID NO:358).
4. S X G F S/A R G X (SEQ ID NO:359).
5. H H H/Q N G R/K W E A R I G R/K V (SEQ ID NO:360).
6. Q E E A A A X Y D (SEQ ID NO:361).

As used herein, the term "Wrinkled 1" or "WRI1" also includes "Wrinkled 1-like" or "WRI1-like" proteins. Examples of WRI1 proteins include Accession Nos: Q6X5Y6, (*Arabidopsis thaliana*; SEQ ID NO:280), XP_002876251.1 (*Arabidopsis lyrata subsp. Lyrata*; SEQ ID NO:281), ABD16282.1 (*Brassica napus*; SEQ ID

NO:282), ADO16346.1 (*Brassica napus*; SEQ ID NO:283), XP_003530370.1 (*Glycine max*; SEQ ID NO:284), AEO22131.1 (*Jatropha curcas*; SEQ ID NO:285), XP_002525305.1 (*Ricinus communis*; SEQ ID NO:286), XP_002316459.1 (*Populus trichocarpa*; SEQ ID NO:287), CBI29147.3 (*Vitis vinifera*; SEQ ID NO:288),

5 XP_003578997.1 (*Brachypodium distachyon*; SEQ ID NO:289), BAJ86627.1 (*Hordeum vulgare* subsp. *vulgare*; SEQ ID NO:290), EAY79792.1 (*Oryza sativa*; SEQ ID NO:291), XP_002450194.1 (*Sorghum bicolor*; SEQ ID NO:292), ACG32367.1 (*Zea mays*; SEQ ID NO:293), XP_003561189.1 (*Brachypodium distachyon*; SEQ ID NO:294), ABL85061.1 (*Brachypodium sylvaticum*; SEQ ID

10 NO:295), BAD68417.1 (*Oryza sativa*; SEQ ID NO:296), XP_002437819.1 (*Sorghum bicolor*; SEQ ID NO:297), XP_002441444.1 (*Sorghum bicolor*; SEQ ID NO:298), XP_003530686.1 (*Glycine max*; SEQ ID NO:299), XP_003553203.1 (*Glycine max*; SEQ ID NO:300), XP_002315794.1 (*Populus trichocarpa*; SEQ ID NO:301), XP_002270149.1 (*Vitis vinifera*; SEQ ID NO:302), XP_003533548.1 (*Glycine max*;

15 SEQ ID NO:303), XP_003551723.1 (*Glycine max*; SEQ ID NO:304), XP_003621117.1 (*Medicago truncatula*; SEQ ID NO:305), XP_002323836.1 (*Populus trichocarpa*; SEQ ID NO:306), XP_002517474.1 (*Ricinus communis*; SEQ ID NO:307), CAN79925.1 (*Vitis vinifera*; SEQ ID NO:308), XP_003572236.1 (*Brachypodium distachyon*; SEQ ID NO:309), BAD10030.1 (*Oryza sativa*; SEQ ID

20 NO:310), XP_002444429.1 (*Sorghum bicolor*; SEQ ID NO:311), NP_001170359.1 (*Zea mays*; SEQ ID NO:312), XP_002889265.1 (*Arabidopsis lyrata* subsp. *lyrata*; SEQ ID NO:313), AAF68121.1 (*Arabidopsis thaliana*; SEQ ID NO:314), NP_178088.2 (*Arabidopsis thaliana*; SEQ ID NO:315), XP_002890145.1 (*Arabidopsis lyrata* subsp. *lyrata*; SEQ ID NO:316), BAJ33872.1 (*Thellungiella*

25 *halophila*; SEQ ID NO:317), NP_563990.1 (*Arabidopsis thaliana*; SEQ ID NO:318), XP_003530350.1 (*Glycine max*; SEQ ID NO:319), XP_003578142.1 (*Brachypodium distachyon*; SEQ ID NO:320), EAZ09147.1 (*Oryza sativa*; SEQ ID NO:321), XP_002460236.1 (*Sorghum bicolor*; SEQ ID NO:322), NP_001146338.1 (*Zea mays*; SEQ ID NO:323), XP_003519167.1 (*Glycine max*; SEQ ID NO:324),

30 XP_003550676.1 (*Glycine max*; SEQ ID NO:325), XP_003610261.1 (*Medicago truncatula*; SEQ ID NO:326), XP_003524030.1 (*Glycine max*; SEQ ID NO:327), XP_003525949.1 (*Glycine max*; SEQ ID NO:328), XP_002325111.1 (*Populus trichocarpa*; SEQ ID NO:329), CBI36586.3 (*Vitis vinifera*; SEQ ID NO:330), XP_002273046.2 (*Vitis vinifera*; SEQ ID NO:331), XP_002303866.1 (*Populus trichocarpa*; SEQ ID NO:332), and CBI25261.3 (*Vitis vinifera*; SEQ ID NO:333).

35 Further examples include Sorbi-WRL1 (SEQ ID NO:334), Lupan-WRL1 (SEQ ID NO:335), Ricco-WRL1 (SEQ ID NO:336), and *Lupin angustifolius* WRI1 (SEQ ID NO:337).

As used herein, the term "monoacylglycerol acyltransferase" or "MGAT" refers to a protein which transfers a fatty acyl group from acyl-CoA to a MAG substrate to produce DAG. Thus, the term "monoacylglycerol acyltransferase activity" at least refers to the transfer of an acyl group from acyl-CoA to MAG to produce DAG. MGAT is best known for its role in fat absorption in the intestine of mammals, where the fatty acids and *sn*-2 MAG generated from the digestion of dietary fat are resynthesized into TAG in enterocytes for chylomicron synthesis and secretion. MGAT catalyzes the first step of this process, in which the acyl group from fatty acyl-CoA, formed from fatty acids and CoA, and *sn*-2 MAG are covalently joined. The term "MGAT" as used herein includes enzymes that act on *sn*-1/3 MAG and/or *sn*-2 MAG substrates to form *sn*-1,3 DAG and/or *sn*-1,2/2,3-DAG, respectively. In a preferred embodiment, the MGAT has a preference for *sn*-2 MAG substrate relative to *sn*-1 MAG, or substantially uses only *sn*-2 MAG as substrate (examples include MGATs described in Cao et al., 2003 (specificity of mouse MGAT1 for *sn*2-18:1-MAG > *sn*1/3-18:1-MAG (Figure 5)); Yen and Farese, 2003 (general activities of mouse MGAT1 and human MGAT2 are higher on 2-MAG than on 1-MAG acyl-acceptor substrates (Figure 5); and Cheng et al., 2003 (activity of human MGAT3 on 2-MAGs is much higher than on 1/3-MAG substrates (Figure 2D)).

As used herein, MGAT does not include enzymes which transfer an acyl group preferentially to LysoPA relative to MAG, such enzymes are known as LPAATs. That is, a MGAT preferentially uses non-phosphorylated monoacyl substrates, even though they may have low catalytic activity on LysoPA. A preferred MGAT does not have detectable activity in acylating LysoPA. As shown herein, a MGAT (i.e., *M. musculus* MGAT2) may also have DGAT function but predominantly functions as a MGAT, i.e., it has greater catalytic activity as a MGAT than as a DGAT when the enzyme activity is expressed in units of nmoles product/min/mg protein (also see Yen et al., 2002).

There are three known classes of MGAT, referred to as, MGAT1, MGAT2 and MGAT3, respectively. Homologs of the human MGAT1 gene (AF384163; SEQ ID NO:7) are present (i.e. sequences are known) at least in chimpanzee, dog, cow, mouse, rat, zebrafish, *Caenorhabditis elegans*, *Schizosaccharomyces pombe*, *Saccharomyces cerevisiae*, *Kluyveromyces lactis*, *Eremothecium gossypii*, *Magnaporthe grisea*, and *Neurospora crassa*. Homologs of the human MGAT2 gene (AY157608) are present at least in chimpanzee, dog, cow, mouse, rat, chicken, zebrafish, fruit fly, and mosquito. Homologs of the human MGAT3 gene (AY229854) are present at least in chimpanzee, dog, cow, and zebrafish. However,

homologs from other organisms can be readily identified by methods known in the art for identifying homologous sequences.

Examples of MGAT1 polypeptides include proteins encoded by MGAT1 genes from *Homo sapiens* (AF384163; SEQ ID NO:7), *Mus musculus* (AF384162; SEQ ID NO:8), *Pan troglodytes* (XM_001166055 and XM_0526044.2; SEQ ID NO:9 and SEQ ID NO:10, respectively), *Canis familiaris* (XM_545667.2; SEQ ID NO:11), *Bos taurus* (NM_001001153.2; SEQ ID NO:12), *Rattus norvegicus* (NM_001108803.1; SEQ ID NO:13), *Danio rerio* MGAT1 (NM_001122623.1; SEQ ID NO:14), *Caenorhabditis elegans* (NM_073012.4, NM_182380.5, NM_065258.3, NM_075068.3, and NM_072248.3; SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, and SEQ ID NO:19, respectively), *Kluyveromyces lactis* (XM_455588.1; SEQ ID NO:20), *Ashbya gossypii* (NM_208895.1; SEQ ID NO:21), *Magnaporthe oryzae* (XM_368741.1; SEQ ID NO:22), *Ciona intestinalis* predicted (XM_002120843.1 SEQ ID NO:23). Examples of MGAT2 polypeptides include proteins encoded by MGAT2 genes from *Homo sapiens* (AY157608; SEQ ID NO:24), *Mus musculus* (AY157609; SEQ ID NO:25), *Pan troglodytes* (XM_522112.2; SEQ ID NO:26), *Canis familiaris* (XM_542304.1; SEQ ID NO:27), *Bos taurus* (NM_001099136.1; SEQ ID NO:28), *Rattus norvegicus* (NM_001109436.2; SEQ ID NO:29), *Gallus gallus* (XM_424082.2; SEQ ID NO:30), *Danio rerio* (NM_001006083.1 SEQ ID NO:31), *Drosophila melanogaster* (NM_136474.2, NM_136473.2, and NM_136475.2; SEQ ID NO:32, SEQ ID NO:33, and SEQ ID NO:34, respectively), *Anopheles gambiae* (XM_001688709.1 and XM_315985; SEQ ID NO:35 and SEQ ID NO:36, respectively), *Tribolium castaneum* (XM_970053.1; SEQ ID NO:37). Examples of MGAT3 polypeptides include proteins encoded by MGAT3 genes from *Homo sapiens* (AY229854; SEQ ID NO:38), *Pan troglodytes* (XM_001154107.1, XM_001154171.1, and XM_527842.2; SEQ ID NO:39, SEQ ID NO:40, and SEQ ID NO:41), *Canis familiaris* (XM_845212.1; SEQ ID NO:42), *Bos taurus* (XM_870406.4; SEQ ID NO:43), *Danio rerio* (XM_688413.4; SEQ ID NO:44).

As used herein "MGAT pathway" refers to an anabolic pathway, different to the Kennedy pathway for the formation of TAG, in which DAG is formed by the acylation of either *sn*-1 MAG or preferably *sn*-2 MAG, catalysed by MGAT. The DAG may subsequently be used to form TAG or other lipids. The MGAT pathway is exemplified in Figure 1.

As used herein, the term "diacylglycerol acyltransferase" (DGAT) refers to a protein which transfers a fatty acyl group from acyl-CoA to a DAG substrate to produce TAG. Thus, the term "diacylglycerol acyltransferase activity" refers to the transfer of an acyl group from acyl-CoA to DAG to produce TAG. A DGAT may

also have MGAT function but predominantly functions as a DGAT, i.e., it has greater catalytic activity as a DGAT than as a MGAT when the enzyme activity is expressed in units of nmoles product/min/mg protein (see for example, Yen et al., 2005).

There are three known types of DGAT, referred to as DGAT1, DGAT2 and DGAT3, respectively. DGAT1 polypeptides typically have 10 transmembrane domains, DGAT2 polypeptides typically have 2 transmembrane domains, whilst DGAT3 polypeptides typically have none and are thought to be soluble in the cytoplasm, not integrated into membranes. Examples of DGAT1 polypeptides include proteins encoded by DGAT1 genes from *Aspergillus fumigatus* (XP_755172.1; SEQ ID NO:347), *Arabidopsis thaliana* (CAB44774.1; SEQ ID NO:83), *Ricinus communis* (AAR11479.1; SEQ ID NO:348), *Vernicia fordii* (ABC94472.1; SEQ ID NO:349), *Vernonia galamensis* (ABV21945.1 and ABV21946.1; SEQ ID NO:350 and SEQ ID NO:351, respectively), *Euonymus alatus* (AAV31083.1; SEQ ID NO:352), *Caenorhabditis elegans* (AAF82410.1; SEQ ID NO:353), *Rattus norvegicus* (NP_445889.1; SEQ ID NO:354), *Homo sapiens* (NP_036211.2; SEQ ID NO:355), as well as variants and/or mutants thereof. Examples of DGAT2 polypeptides include proteins encoded by DGAT2 genes from *Arabidopsis thaliana* (NP_566952.1; SEQ ID NO:212), *Ricinus communis* (AAY16324.1; SEQ ID NO:213), *Vernicia fordii* (ABC94474.1; SEQ ID NO:214), *Mortierella ramanniana* (AAK84179.1; SEQ ID NO:215), *Homo sapiens* (Q96PD7.2; SEQ ID NO:216) (Q58HT5.1; SEQ ID NO:217), *Bos taurus* (Q70VZ8.1; SEQ ID NO:218), *Mus musculus* (AAK84175.1; SEQ ID NO:219), as well as variants and/or mutants thereof.

Examples of DGAT3 polypeptides include proteins encoded by DGAT3 genes from peanut (*Arachis hypogaea*, Saha, et al., 2006), as well as variants and/or mutants thereof. A DGAT has little or no detectable MGAT activity, for example, less than 300 pmol/min/mg protein, preferably less than 200 pmol/min/mg protein, more preferably less than 100 pmol/min/mg protein.

DGAT2 but not DGAT1 shares high sequence homology with the MGAT enzymes, suggesting that DGAT2 and MGAT genes likely share a common genetic origin. Although multiple isoforms are involved in catalysing the same step in TAG synthesis, they may play distinct functional roles, as suggested by differential tissue distribution and subcellular localization of the DGAT/MGAT family of enzymes. In mammals, MGAT1 is mainly expressed in stomach, kidney, adipose tissue, whilst MGAT2 and MGAT3 show highest expression in the small intestine. In mammals, DGAT1 is ubiquitously expressed in many tissues, with highest expression in small intestine, whilst DGAT2 is most abundant in liver. MGAT3 only exists in higher mammals and humans, but not in rodents from bioinformatic analysis. MGAT3

shares higher sequence homology to DGAT2 than MGAT1 and MGAT3. MGAT3 exhibits significantly higher DGAT activity than MGAT1 and MGAT2 enzymes (MGAT3 > MGAT1 > MGAT2) when either MAGs or DAGs were used as substrates, suggesting MGAT3 functions as a putative TAG synthase.

5 Both MGAT1 and MGAT2 belong to the same class of acyltransferases as DGAT2. Some of the motifs that have been shown to be important for DGAT2 catalytic activity in some DGAT2s are also conserved in MGAT acyltransferases. Of particular interest is a putative neutral lipid-binding domain with the consensus sequence FLXLXXXN (SEQ ID NO:224) where each X is independently any amino
10 acid other than proline, and N is any nonpolar amino acid, located within the N-terminal transmembrane region followed by a putative glycerol/phospholipid acyltransferase domain. The FLXLXXXN motif (SEQ ID NO:224) is found in the mouse DGAT2 (amino acids 81-88) and MGAT1/2 but not in yeast or plant DGAT2s. It is important for activity of the mouse DGAT2. Other DGAT2 and/or MGAT1/2
15 sequence motifs include:

1. A highly conserved YFP tripeptide (SEQ ID NO:220) in most DGAT2 polypeptides and also in MGAT1 and MGAT2, for example, present as amino acids 139-141 in mouse DGAT2. Mutating this motif within the yeast DGAT2 with non-conservative substitutions rendered the enzyme non-functional.

20 2. HPHG tetrapeptide (SEQ ID NO:221), highly conserved in MGATs as well as in DGAT2 sequences from animals and fungi, for example, present as amino acids 161-164 in mouse DGAT2, and important for catalytic activity at least in yeast and mouse DGAT2. Plant DGAT2 acyltransferases have a EPHS (SEQ ID NO:222) conserved sequence instead, so conservative changes to the first and fourth amino
25 acids can be tolerated.

3. A longer conserved motif which is part of the putative glycerol phospholipid domain. An example of this motif is RXGFX(K/R)AXXXGXXX(L/V)VPXXXFG(E/Q) (SEQ ID NO:223), which is present as amino acids 304-327 in mouse DGAT2. This motif is less conserved in
30 amino acid sequence than the others, as would be expected from its length, but homologs can be recognised by motif searching. The spacing may vary between the more conserved amino acids, i.e., there may be additional X amino acids within the motif, or less X amino acids compared to the sequence above.

As used herein, the term "glycerol-3-phosphate acyltransferase" or "GPAT"
35 refers to a protein which acylates glycerol-3-phosphate (G-3-P) to form LysoPA and/or MAG, the latter product forming if the GPAT also has phosphatase activity on LysoPA. The acyl group that is transferred is typically from acyl-CoA. Thus, the term "glycerol-3-phosphate acyltransferase activity" refers to the acylation of G-3-P to

form LysoPA and/or MAG. The term "GPAT" encompasses enzymes that acylate G-3-P to form *sn*-1 LPA and/or *sn*-2 LPA, preferably *sn*-2 LPA. In a preferred embodiment, the GPAT has phosphatase activity. In a most preferred embodiment, the GPAT is a *sn*-2 GPAT having phosphatase activity which produces *sn*-2 MAG.

5 As used herein, the term "*sn*-1 glycerol-3-phosphate acyltransferase" (*sn*-1 GPAT) refers to a protein which acylates *sn*-glycerol-3-phosphate (G-3-P) to preferentially form 1-acyl-*sn*-glycerol-3-phosphate (*sn*-1 LPA). Thus, the term "*sn*-1 glycerol-3-phosphate acyltransferase activity" refers to the acylation of *sn*-glycerol-3-phosphate to form 1-acyl-*sn*-glycerol-3-phosphate (*sn*-1 LPA).

10 As used herein, the term "*sn*-2 glycerol-3-phosphate acyltransferase" (*sn*-2 GPAT) refers to a protein which acylates *sn*-glycerol-3-phosphate (G-3-P) to preferentially form 2-acyl-*sn*-glycerol-3-phosphate (*sn*-2 LPA). Thus, the term "*sn*-2 glycerol-3-phosphate acyltransferase activity" refers to the acylation of *sn*-glycerol-3-phosphate to form 2-acyl-*sn*-glycerol-3-phosphate (*sn*-2 LPA).

15 The GPAT family is large and all known members contain two conserved domains, a plsC acyltransferase domain (PF01553; SEQ ID NO:225) and a HAD-like hydrolase (PF12710; SEQ ID NO:226) superfamily domain. In addition to this, in *Arabidopsis thaliana*, GPAT4-8 all contain a N-terminal region homologous to a phosphoserine phosphatase domain (PF00702; SEQ ID NO:227). GPAT4 and
20 GPAT6 both contain conserved residues that are known to be critical to phosphatase activity, specifically conserved amino acids (shown in bold) in Motif I (DXDX[T/V][L/V]; SEQ ID NO:229) and Motif III (K-[G/S][D/S]XXX[D/N]; SEQ ID NO:330) located at the N-terminus (Yang et al., 2010). Preferably, the GPAT has *sn*-2 preference and phosphatase activity to produce *sn*-2 MAG (also referred to
25 herein as "2-MAG") from glycerol-3-phosphate (G-3-P) (Figure 1), for example, GPAT4 (NP_171667.1; SEQ ID NO:144) and GPAT6 (NP_181346.1; SEQ ID NO:145) from *Arabidopsis*. More preferably, the GPAT uses acyl-CoA as a fatty acid substrate.

Homologues of GPAT4 (NP_171667.1; SEQ ID NO:144) and GPAT6
30 (NP_181346.1; SEQ ID NO:145) include AAF02784.1 (*Arabidopsis thaliana*; SEQ ID NO:146), AAL32544.1 (*Arabidopsis thaliana*; SEQ ID NO:147), AAP03413.1 (*Oryza sativa*; SEQ ID NO:148), ABK25381.1 (*Picea sitchensis*; SEQ ID NO:149), ACN34546.1 (*Zea Mays*; SEQ ID NO:150), BAF00762.1 (*Arabidopsis thaliana*; SEQ ID NO:151), BAH00933.1 (*Oryza sativa*; SEQ ID NO:152), EAY84189.1 (*Oryza sativa*; SEQ ID NO:153), EAY98245.1 (*Oryza sativa*; SEQ ID NO:154), EAZ21484.1
35 (*Oryza sativa*; SEQ ID NO:155), EEC71826.1 (*Oryza sativa*; SEQ ID NO:156), EEC76137.1 (*Oryza sativa*; SEQ ID NO:157), EEE59882.1 (*Oryza sativa*; SEQ ID NO:158), EFJ08963.1 (*Selaginella moellendorffii*; SEQ ID NO:159), EFJ08964.1

(*Selaginella moellendorffii*; SEQ ID NO:160), EFJ11200.1 (*Selaginella moellendorffii*; SEQ ID NO:161), EFJ15664.1 (*Selaginella moellendorffii*; SEQ ID NO:162), EFJ24086.1 (*Selaginella moellendorffii*; SEQ ID NO:163), EFJ29816.1 (*Selaginella moellendorffii*; SEQ ID NO:164), EFJ29817.1 (*Selaginella moellendorffii*; SEQ ID NO:165), NP_001044839.1 (*Oryza sativa*; SEQ ID NO:166), NP_001045668.1 (*Oryza sativa*; SEQ ID NO:167), NP_001147442.1 (*Zea mays*; SEQ ID NO:168), NP_001149307.1 (*Zea mays*; SEQ ID NO:169), NP_001168351.1 (*Zea mays*; SEQ ID NO:170), AFH02724.1 (*Brassica napus*; SEQ ID NO:171), NP_191950.2 (*Arabidopsis thaliana*; SEQ ID NO:172), XP_001765001.1 (*Physcomitrella patens*; SEQ ID NO:173), XP_001769671.1 (*Physcomitrella patens*; SEQ ID NO:174), XP_001769724.1 (*Physcomitrella patens*; SEQ ID NO:175), XP_001771186.1 (*Physcomitrella patens*; SEQ ID NO:176), XP_001780533.1 (*Physcomitrella patens*; SEQ ID NO:177), XP_002268513.1 (*Vitis vinifera*; SEQ ID NO:178), XP_002275348.1 (*Vitis vinifera*; SEQ ID NO:179), XP_002276032.1 (*Vitis vinifera*; SEQ ID NO:180), XP_002279091.1 (*Vitis vinifera*; SEQ ID NO:181), XP_002309124.1 (*Populus trichocarpa*; SEQ ID NO:182), XP_002309276.1 (*Populus trichocarpa*; SEQ ID NO:183), XP_002322752.1 (*Populus trichocarpa*; SEQ ID NO:184), XP_002323563.1 (*Populus trichocarpa*; SEQ ID NO:185), XP_002439887.1 (*Sorghum bicolor*; SEQ ID NO:186), XP_002458786.1 (*Sorghum bicolor*; SEQ ID NO:187), XP_002463916.1 (*Sorghum bicolor*; SEQ ID NO:188), XP_002464630.1 (*Sorghum bicolor*; SEQ ID NO:189), XP_002511873.1 (*Ricinus communis*; SEQ ID NO:190), XP_002517438.1 (*Ricinus communis*; SEQ ID NO:191), XP_002520171.1 (*Ricinus communis*; SEQ ID NO:192), XP_002872955.1 (*Arabidopsis lyrata*; SEQ ID NO:193), XP_002881564.1 (*Arabidopsis lyrata*; SEQ ID NO:194), ACT32032.1 (*Vernicia fordii*; SEQ ID NO:195), NP_001051189.1 (*Oryza sativa*; SEQ ID NO:196), AFH02725.1 (*Brassica napus*; SEQ ID NO:197), XP_002320138.1 (*Populus trichocarpa*; SEQ ID NO:198), XP_002451377.1 (*Sorghum bicolor*; SEQ ID NO:199), XP_002531350.1 (*Ricinus communis*; SEQ ID NO:200), and XP_002889361.1 (*Arabidopsis lyrata*; SEQ ID NO:201).

Conserved motifs and/or residues can be used as a sequence-based diagnostic for the identification of bifunctional GPAT/phosphatase enzymes. Alternatively, a more stringent function-based assay could be utilised. Such an assay involves, for example, feeding labelled glycerol-3-phosphate to cells or microsomes and quantifying the levels of labelled products by thin-layer chromatography or a similar technique. GPAT activity results in the production of labelled LPA whilst GPAT/phosphatase activity results in the production of labelled MAG.

As used herein, the term "Oleosin" refers to an amphipathic protein present in the membrane of oil bodies in the storage tissues of seeds (see, for example, Huang,

1996; Lin et al., 2005; Capuano et al., 2007; Lui et al., 2009; Shimada and Hara-Nishimura, 2010).

Plant seeds accumulate TAG in subcellular structures called oil bodies. These organelles consist of a TAG core surround by a phospholipid monolayer containing
5 several embedded proteins including oleosins (Jolivet et al., 2004). Oleosins represent the most abundant protein in the membrane of oil bodies.

Oleosins are of low M_r (15-26,000). Within each seed species, there are usually two or more oleosins of different M_r . Each oleosin molecule contains a relatively hydrophilic N-terminal domain (for example, about 48 amino acid
10 residues), a central totally hydrophobic domain (for example, of about 70-80 amino acid residues) which is particularly rich in aliphatic amino acids such as alanine, glycine, leucine, isoleucine and valine, and an amphipathic α -helical domain (for example, about of about 33 amino acid residues) at or near the C-terminus. Generally, the central stretch of hydrophobic residues is inserted into the lipid core and the
15 amphipathic N-terminal and/or amphipathic C-terminal are located at the surface of the oil bodies, with positively charged residues embedded in a phospholipid monolayer and the negatively charged ones exposed to the exterior.

As used herein, the term "Oleosin" encompasses polyoleosins which have multiple oleosin polypeptides fused together as a single polypeptide, for example 2x,
20 4x or 6x oleosin peptides, and caleosins which bind calcium (Froissard et al., 2009), and steroleosins which bind sterols. However, generally a large proportion of the oleosins of oil bodies will not be caleosins and/or steroleosins.

A substantial number of oleosin protein sequences, and nucleotide sequences encoding therefor, are known from a large number of different plant species.
25 Examples include, but are not limited to, oleosins from *Arabidopsis*, canola, corn, rice, peanut, castor, soybean, flax, grape, cabbage, cotton, sunflower, sorghum and barley. Examples of oleosins (with their Accession Nos) include *Brassica napus* oleosin (CAA57545.1; SEQ ID NO:362), *Brassica napus* oleosin S1-1 (ACG69504.1; SEQ ID NO:363), *Brassica napus* oleosin S2-1 (ACG69503.1; SEQ ID NO:364),
30 *Brassica napus* oleosin S3-1 (ACG69513.1; SEQ ID NO:365), *Brassica napus* oleosin S4-1 (ACG69507.1; SEQ ID NO:366), *Brassica napus* oleosin S5-1 (ACG69511.1; SEQ ID NO:367), *Arachis hypogaea* oleosin 1 (AAZ20276.1; SEQ ID NO:368), *Arachis hypogaea* oleosin 2 (AAU21500.1; SEQ ID NO:369), *Arachis hypogaea* oleosin 3 (AAU21501.1; SEQ ID NO:370), *Arachis hypogaea* oleosin 5
35 (ABC96763.1; SEQ ID NO:371), *Ricinus communis* oleosin 1 (EEF40948.1; SEQ ID NO:372), *Ricinus communis* oleosin 2 (EEF51616.1; SEQ ID NO:373), *Glycine max* oleosin isoform a (P29530.2; SEQ ID NO:374), *Glycine max* oleosin isoform b (P29531.1; SEQ ID NO:375), *Linum usitatissimum* oleosin low molecular weight

isoform (ABB01622.1; SEQ ID NO:376), *Linum usitatissimum* oleosin high molecular weight isoform (ABB01624.1; SEQ ID NO:377), *Helianthus annuus* oleosin (CAA44224.1; SEQ ID NO:378), *Zea mays* oleosin (NP_001105338.1; SEQ ID NO:379), *Brassica napus* steroleosin (ABM30178.1; SEQ ID NO:380), *Brassica*
 5 *napus* steroleosin SLO1-1 (ACG69522.1; SEQ ID NO:381), *Brassica napus* steroleosin SLO2-1 (ACG69525.1; SEQ ID NO:382), *Sesamum indicum* steroleosin (AAL13315.1; SEQ ID NO:383), *Zea mays* steroleosin (NP_001152614.1; SEQ ID NO:384), *Brassica napus* caleosin CLO-1 (ACG69529.1; SEQ ID NO:385), *Brassica napus* caleosin CLO-3 (ACG69527.1; SEQ ID NO:386), *Sesamum indicum* caleosin
 10 (AAF13743.1; SEQ ID NO:387), *Zea mays* caleosin (NP_001151906.1; SEQ ID NO:388).

As used herein, the term a "polypeptide involved in starch biosynthesis" refers to any polypeptide, the downregulation of which in a cell below normal (wild-type) levels results in a reduction in the level of starch synthesis and an increase in the
 15 levels of starch. An example of such a polypeptide is AGPase.

As used herein, the term "ADP-glucose phosphorylase" or "AGPase" refers to an enzyme which regulates starch biosynthesis, catalysing conversion of glucose-1-phosphate and ATP to ADP-glucose which serves as the building block for starch polymers. The active form of the AGPase enzyme consists of 2 large and 2 small
 20 subunits.

The ADPase enzyme in plants exists primarily as a tetramer which consists of 2 large and 2 small subunits. Although these subunits differ in their catalytic and regulatory roles depending on the species (Kuhn et al., 2009), in plants the small subunit generally displays catalytic activity. The molecular weight of the small
 25 subunit is approximately 50-55 kDa. The molecular weight of the large large subunit is approximately 55-60 kDa. The plant enzyme is strongly activated by 3-phosphoglycerate (PGA), a product of carbon dioxide fixation; in the absence of PGA, the enzyme exhibits only about 3% of its activity. Plant AGPase is also strongly inhibited by inorganic phosphate (Pi). In contrast, bacterial and algal
 30 AGPase exist as homotetramers of 50kDa. The algal enzyme, like its plant counterpart, is activated by PGA and inhibited by Pi, whereas the bacterial enzyme is activated by fructose-1, 6-bisphosphate (FBP) and inhibited by AMP and Pi.

As used herein, the term "polypeptide involved in the degradation of lipid and/or which reduces lipid content" refers to any polypeptide, the downregulation of
 35 which in a cell below normal (wild-type) levels results an increase in the level of oil, such as fatty acids and/or TAGs, in the cell, preferably a cell of vegetative tissue of a plant. Examples of such polypeptides include, but are not limited, lipases, or a lipase

such as CGI58 polypeptide, SUGAR-DEPENDENT1 triacylglycerol lipase (see, for example, Kelly et al., 2012) or a lipase described in WO 2009/027335.

As used herein, the term "lipase" refers to a protein which hydrolyzes fats into glycerol and fatty acids. Thus, the term "lipase activity" refers to the hydrolysis of
5 fats into glycerol and fatty acids.

As used herein, the term "CGI58" refers to a soluble acyl-CoA-dependent lysophosphatidic acid acyltransferase also known in the art as "At4g24160" (in plants) and "Ict1p" (in yeast). The plant gene such as that from *Arabidopsis* gene locus, At4g24160, is expressed as two alternative transcripts: a longer full-length
10 isoform (At4g24160.1) and a smaller isoform (At4g24160.2) missing a portion of the 3' end (see James et al., 2010; Ghosh et al., 2009; US 201000221400). Both mRNAs code for a protein that is homologous to the human CGI-58 protein and other orthologous members of this α/β hydrolase family (ABHD). In an embodiment, the CGI58 (At4g24160) protein contains three motifs that are conserved across plant
15 species: a GX SXG lipase motif (SEQ ID NO:419), a HX(4)D acyltransferase motif (SEQ ID NO:420), and VX(3)HGF, a probable lipid binding motif (SEQ ID NO:421). The human CGI-58 protein has lysophosphatidic acid acyltransferase (LPAAT) activity but not lipase activity. In contrast, the plant and yeast proteins possess a canonical lipase sequence motif GX SXG (SEQ ID NO:419), that is absent from
20 vertebrate (humans, mice, and zebrafish) proteins. Although the plant and yeast CGI58 proteins appear to possess detectable amounts of TAG lipase and phospholipase A activities in addition to LPAAT activity, the human protein does not.

Disruption of the homologous *CGI-58* gene in *Arabidopsis thaliana* results in the accumulation of neutral lipid droplets in mature leaves. Mass spectroscopy of
25 isolated lipid droplets from *cgi-58* loss-of-function mutants showed they contain triacylglycerols with common leaf-specific fatty acids. Leaves of mature *cgi-58* plants exhibit a marked increase in absolute triacylglycerol levels, more than 10-fold higher than in wild-type plants. Lipid levels in the oil-storing seeds of *cgi-58* loss-of-function plants were unchanged, and unlike mutations in β -oxidation, the *cgi-58* seeds
30 germinated and grew normally, requiring no rescue with sucrose (James et al., 2010).

Examples of CGI58 polypeptides include proteins from *Arabidopsis thaliana* (NP_194147.2; SEQ ID NO:429), *Brachypodium distachyon* (XP_003578450.1; SEQ ID NO:430), *Glycine max* (XP_003523638.1; SEQ ID NO:431), *Zea mays* (NP_001149013.1; SEQ ID NO:432), *Sorghum bicolor* (XP_002460538.1; SEQ ID
35 NO:433), *Ricinus communis* (XP_002510485.1; SEQ ID NO:434), *Medicago truncatula* (XP_003603733.1; SEQ ID NO:435), and *Oryza sativa* (EAA09782.1; SEQ ID NO:436).

As used herein, the term "Leafy Cotyledon 2" or "LEC2" refers to a B3 domain transcription factor which participates in zygotic and in somatic embryogenesis. Its ectopic expression facilitates the embryogenesis from vegetative plant tissues (Alemanno et al., 2008). LEC2 also comprises a DNA binding region
 5 found thus far only in plant proteins. Examples of LEC2 polypeptides include proteins from *Arabidopsis thaliana* (NP_564304.1) (SEQ ID NO:442), *Medicago truncatula* (CAA42938.1) (SEQ ID NO:443) and *Brassica napus* (ADO16343.1) (SEQ ID NO:444).

As used herein, the term "BABY BOOM" or "BBM" refers an AP2/ERF
 10 transcription factor that induces regeneration under culture conditions that normally do not support regeneration in wild-type plants. Ectopic expression of *Brassica napus* *BBM* (*BnBBM*) genes in *B. napus* and *Arabidopsis* induces spontaneous somatic embryogenesis and organogenesis from seedlings grown on hormone-free basal medium (Boutilier et al., 2002). In tobacco, ectopic *BBM* expression is sufficient to
 15 induce adventitious shoot and root regeneration on basal medium, but exogenous cytokinin is required for somatic embryo (SE) formation (Srinivasan et al., 2007). Examples of BBM polypeptides include proteins from *Arabidopsis thaliana* (NP_197245.2) (SEQ ID NO:445) and *Medicago truncatula* (AAW82334.1) (SEQ ID NO:446).

20 As used herein, the term "FAD2" refers to a membrane bound delta-12 fatty acid desaturase that desaturates oleic acid (18:1^{Δ9}) to produce linoleic acid (C18:2^{Δ9,12}).

As used herein, the term "epoxygenase" or "fatty acid epoxygenase" refers to an enzyme that introduces an epoxy group into a fatty acid resulting in the production of an epoxy fatty acid. In preferred embodiment, the epoxy group is introduced at the
 25 12th carbon on a fatty acid chain, in which case the epoxygenase is a Δ12-epoxygenase, especially of a C16 or C18 fatty acid chain. The epoxygenase may be a Δ9-epoxygenase, a Δ15 epoxygenase, or act at a different position in the acyl chain as known in the art. The epoxygenase may be of the P450 class. Preferred epoxygenases are of the mono-oxygenase class as described in WO98/46762. Numerous
 30 epoxygenases or presumed epoxygenases have been cloned and are known in the art. Further examples of epoxygenases include proteins comprising an amino acid sequence provided in SEQ ID NO:21 of WO 2009/129582, polypeptides encoded by genes from *Crepis palestina* (CAA76156, Lee et al., 1998), *Stokesia laevis* (AAR23815, Hatanaka et al., 2004) (monooxygenase type), *Euphorbia lagascae* (AAL62063) (P450 type), human CYP2J2 (arachidonic acid epoxygenase, U37143);
 35 (AAL62063) (P450 type), human CYP1A1 (arachidonic acid epoxygenase, K03191), as well as variants and/or mutants thereof.

As used herein, the term, "hydroxylase" or "fatty acid hydroxylase" refers to an enzyme that introduces a hydroxyl group into a fatty acid resulting in the production of a hydroxylated fatty acid. In a preferred embodiment, the hydroxyl group is introduced at the 2nd, 12th and/or 17th carbon on a C18 fatty acid chain. Preferably, the hydroxyl group is introduced at the 12th carbon, in which case the hydroxylase is a Δ 12-hydroxylase. In another preferred embodiment, the hydroxyl group is introduced at the 15th carbon on a C16 fatty acid chain. Hydroxylases may also have enzyme activity as a fatty acid desaturase. Examples of genes encoding Δ 12-hydroxylases include those from *Ricinus communis* (AAC9010, van de Loo 1995); *Physaria lindheimeri*, (ABQ01458, Dauk et al., 2007); *Lesquerella fendleri*, (AAC32755, Broun et al., 1998); *Daucus carota*, (AAK30206); fatty acid hydroxylases which hydroxylate the terminus of fatty acids, for example: *A. thaliana* CYP86A1 (P48422, fatty acid ω -hydroxylase); *Vicia sativa* CYP94A1 (P98188, fatty acid ω -hydroxylase); mouse CYP2E1 (X62595, lauric acid ω -1 hydroxylase); rat CYP4A1 (M57718, fatty acid ω -hydroxylase), as well as as variants and/or mutants thereof.

As used herein, the term "conjugase" or "fatty acid conjugase" refers to an enzyme capable of forming a conjugated bond in the acyl chain of a fatty acid. Examples of conjugases include those encoded by genes from *Calendula officinalis* (AF343064, Qiu et al., 2001); *Vernicia fordii* (AAN87574, Dyer et al., 2002); *Punica granatum* (AY178446, Iwabuchi et al., 2003) and *Trichosanthes kirilowii* (AY178444, Iwabuchi et al., 2003); as well as as variants and/or mutants thereof.

As used herein, the term "acetylenase" or "fatty acid acetylenase" refers to an enzyme that introduces a triple bond into a fatty acid resulting in the production of an acetylenic fatty acid. In a preferred embodiment, the triple bond is introduced at the 2nd, 6th, 12th and/or 17th carbon on a C18 fatty acid chain. Examples acetylenases include those from *Helianthus annuus* (AA038032, ABC59684), as well as as variants and/or mutants thereof.

Examples of such fatty acid modifying genes include proteins according to the following Accession Numbers which are grouped by putative function, and homologues from other species: Δ 12 acetylenases ABC00769, CAA76158, AAO38036, AAO38032; Δ 12 conjugases AAG42259, AAG42260, AAN87574; Δ 12 desaturases P46313, ABS18716, AAS57577, AAL61825, AAF04093, AAF04094; Δ 12 epoxxygenases XP_001840127, CAA76156, AAR23815; Δ 12 hydroxylases ACF37070, AAC32755, ABQ01458, AAC49010; and Δ 12 P450 enzymes such as AF406732.

As used herein, the term "vegetative tissue" or "vegetative plant part" is any plant tissue, organ or part other than organs for sexual reproduction of plants, specifically seed bearing organs, flowers, pollen, fruits and seeds. Vegetative tissues

and parts include at least plant leaves, stems (including bolts and tillers but excluding the heads), tubers and roots, but excludes flowers, pollen, seed including the seed coat, embryo and endosperm, fruit including mesocarp tissue, seed-bearing pods and seed-bearing heads. In one embodiment, the vegetative part of the plant is an aerial
5 plant part. In another or further embodiment, the vegetative plant part is a green part such as a leaf or stem.

As used herein, the term "wild-type" or variations thereof refers to a cell, or non-human organism or part thereof that has not been genetically modified.

The term "corresponding" refers to a vegetative plant part, a cell, or non-
10 human organism or part thereof, or seed that has the same or similar genetic background as a vegetative plant part, a cell, or non-human organism or part thereof, or seed of the invention but that has not been modified as described herein (for example, a vegetative plant part, a cell, or non-human organism or part thereof, or seed lacks an exogenous polynucleotide encoding a MGAT or an exogenous MGAT).
15 In a preferred embodiment, a vegetative plant part, a cell, or non-human organism or part thereof, or seed is at the same developmental stage as a vegetative plant part, a cell, or non-human organism or part thereof, or seed of the invention. For example, if the non-human organism is a flowering plant, then preferably the corresponding plant is also flowering. A corresponding a vegetative plant part, a cell, or non-human
20 organism or part thereof, or seed can be used as a control to compare levels of nucleic acid or protein expression, or the extent and nature of trait modification, for example non-polar lipid production and/or content, with a vegetative plant part, a cell, or non-human organism or part thereof, or seed modified as described herein. A person skilled in the art is readily able to determine an appropriate "corresponding" cell,
25 tissue, organ or organism for such a comparison.

As used herein "compared with" refers to comparing levels of a non-polar lipid or total non-polar lipid content of the transgenic non-human organism or part thereof expressing the one or more exogenous polynucleotides or exogenous polypeptides with a transgenic non-human organism or part thereof lacking the one or more
30 exogenous polynucleotides or polypeptides.

As used herein, "enhanced ability to produce non-polar lipid" is a relative term which refers to the total amount of non-polar lipid being produced by a cell, or non-human organism or part thereof of the invention being increased relative to a corresponding cell, or non-human organism or part thereof. In one embodiment, the
35 TAG and/or polyunsaturated fatty acid content of the non-polar lipid is increased.

As used herein, "germinate at a rate substantially the same as for a corresponding wild-type plant" refers to seed of a plant of the invention being relatively fertile when compared to seed of a wild type plant lacking the defined

exogenous polynucleotide(s). In one embodiment, the number of seeds which germinate, for instance when grown under optimal greenhouse conditions for the plant species, is at least 75%, more preferably at least 90%, of that when compared to corresponding wild-type seed. In another embodiment, the seeds which germinate,
5 for instance when grown under optimal greenhouse conditions for the plant species, grow at a rate which, on average, is at least 75%, more preferably at least 90%, of that when compared to corresponding wild-type plants.

As used herein, the term "an isolated or recombinant polynucleotide which down regulates the production and/or activity of an endogenous enzyme" or variations
10 thereof, refers to a polynucleotide that encodes an RNA molecule that down regulates the production and/or activity (for example, encoding an siRNA, hpRNAi), or itself down regulates the production and/or activity (for example, is an siRNA which can be delivered directly to, for example, a cell) of an endogenous enzyme for example, DGAT, *sn*-1 glycerol-3-phosphate acyltransferase (GPAT), 1-acyl-glycerol-3-
15 phosphate acyltransferase (LPAAT), acyl-CoA:lysophosphatidylcholine acyltransferase (LPCAT), phosphatidic acid phosphatase (PAP), AGPase, or delta-12 fatty acid desaturase (FAD2), or a combination of two or more thereof.

As used herein, the term "on a weight basis" refers to the weight of a substance (for example, TAG, DAG, fatty acid) as a percentage of the weight of the composition
20 comprising the substance (for example, seed, leaf). For example, if a transgenic seed has 25 µg total fatty acid per 120 µg seed weight; the percentage of total fatty acid on a weight basis is 20.8%.

As used herein, the term "on a relative basis" refers to the amount of a substance in a composition comprising the substance in comparison with a
25 corresponding composition, as a percentage.

As used herein, the term "the relative non-lipid content" refers to the expression of the non-polar lipid content of a cell, organism or part thereof, or extracted lipid therefrom, in comparison with a corresponding cell, organism or part thereof, or the lipid extracted from the corresponding cell, organism or part thereof, as
30 a percentage. For example, if a transgenic seed has 25 µg total fatty acid, whilst the corresponding seed had 20 µg total fatty acid; the increase in non-polar lipid content on a relative basis equals 25%.

As used herein, the term "biofuel" refers to any type of fuel, typically as used to power machinery such as automobiles, trucks or petroleum powered motors, whose
35 energy is derived from biological carbon fixation. Biofuels include fuels derived from biomass conversion, as well as solid biomass, liquid fuels and biogases. Examples of biofuels include bioalcohols, biodiesel, synthetic diesel, vegetable oil, bioethers, biogas, syngas, solid biofuels, algae-derived fuel, biohydrogen, biomethanol, 2,5-

Dimethylfuran (DMF), biodimethyl ether (bioDME), Fischer-Tropsch diesel, biohydrogen diesel, mixed alcohols and wood diesel.

As used herein, the term "bioalcohol" refers to biologically produced alcohols, for example, ethanol, propanol and butanol. Bioalcohols are produced by the action
5 of microorganisms and/or enzymes through the fermentation of sugars, hemicellulose or cellulose.

As used herein, the term "biodiesel" refers to a composition comprising fatty acid methyl- or ethyl- esters derived from non-polar lipids by transesterification.

As used herein, the term "synthetic diesel" refers to a form of diesel fuel which
10 is derived from renewable feedstock rather than the fossil feedstock used in most diesel fuels.

As used herein, the term "vegetable oil" includes a pure plant oil (or straight vegetable oil) or a waste vegetable oil (by product of other industries).

As used herein, the term "bioethers" refers to compounds that act as octane
15 rating enhancers.

As used herein, the term "biogas" refers to methane or a flammable mixture of methane and other gases produced by anaerobic digestion of organic material by anaerobes.

As used herein, the term "syngas" refers to a gas mixture that contains varying
20 amounts of carbon monoxide and hydrogen and possibly other hydrocarbons, produced by partial combustion of biomass.

As used herein, the term "solid biofuels" includes wood, sawdust, grass trimming, and non-food energy crops.

As used herein, the term "cellulosic ethanol" refers to ethanol produced from
25 cellulose or hemicellulose.

As used herein, the term "algae fuel" refers to a biofuel made from algae and includes algal biodiesel, biobutanol, biogasoline, methane, ethanol, and the equivalent of vegetable oil made from algae.

As used herein, the term "biohydrogen" refers to hydrogen produced
30 biologically by, for example, algae.

As used herein, the term "biomethanol" refers to methanol produced biologically. Biomethanol may be produced by gasification of organic materials to syngas followed by conventional methanol synthesis.

As used herein, the term "2,5-Dimethylfuran" or "DMF" refers to a
35 heterocyclic compound with the formula $(\text{CH}_3)_2\text{C}_4\text{H}_2\text{O}$. DMF is a derivative of furan that is derivable from cellulose.

As used herein, the term "biodimethyl ether" or "bioDME", also known as methoxymethane, refers to an organic compound with the formula CH_3OCH_3 .

Syngas may be converted into methanol in the presence of catalyst (usually copper-based), with subsequent methanol dehydration in the presence of a different catalyst (for example, silica-alumina) resulting in the production of DME.

As used herein, the term "Fischer-Tropsch" refers to a set of chemical reactions that convert a mixture of carbon monoxide and hydrogen into liquid hydrocarbons. The syngas can first be conditioned using for example, a water gas shift to achieve the required H_2/CO ratio. The conversion takes place in the presence of a catalyst, usually iron or cobalt. The temperature, pressure and catalyst determine whether a light or heavy syncrude is produced. For example at 330°C mostly gasoline and olefins are produced whereas at 180° to 250°C mostly diesel and waxes are produced. The liquids produced from the syngas, which comprise various hydrocarbon fractions, are very clean (sulphur free) straight-chain hydrocarbons. Fischer-Tropsch diesel can be produced directly, but a higher yield is achieved if first Fischer-Tropsch wax is produced, followed by hydrocracking.

As used herein, the term "biochar" refers to charcoal made from biomass, for example, by pyrolysis of the biomass.

As used herein, the term "feedstock" refers to a material, for example, biomass or a conversion product thereof (for example, syngas) when used to produce a product, for example, a biofuel such as biodiesel or a synthetic diesel.

As used herein, the term "industrial product" refers to a hydrocarbon product which is predominantly made of carbon and hydrogen such as fatty acid methyl- and/or ethyl-esters or alkanes such as methane, mixtures of longer chain alkanes which are typically liquids at ambient temperatures, a biofuel, carbon monoxide and/or hydrogen, or a bioalcohol such as ethanol, propanol, or butanol, or biochar. The term "industrial product" is intended to include intermediary products that can be converted to other industrial products, for example, syngas is itself considered to be an industrial product which can be used to synthesize a hydrocarbon product which is also considered to be an industrial product. The term industrial product as used herein includes both pure forms of the above compounds, or more commonly a mixture of various compounds and components, for example the hydrocarbon product may contain a range of carbon chain lengths, as well understood in the art.

As used herein, "gloss" refers to an optical phenomenon caused when evaluating the appearance of a surface. The evaluation of gloss describes the capacity of a surface to reflect directed light.

Throughout this specification the word "comprise", or variations such as "comprises" or "comprising", will be understood to imply the inclusion of a stated element, integer or step, or group of elements, integers or steps, but not the exclusion of any other element, integer or step, or group of elements, integers or steps.

The term "and/or", e.g., "X and/or Y" shall be understood to mean either "X and Y" or "X or Y" and shall be taken to provide explicit support for both meanings or for either meaning.

As used herein, the term about, unless stated to the contrary, refers to +/- 10%,
 5 more preferably +/- 5%, more preferably +/- 2%, more preferably +/- 1%, even more preferably +/- 0.5%, of the designated value.

Production of Diacylglycerols and Triacylglycerols

In one embodiment, the vegetative plant part, transgenic non-human organism
 10 or part thereof of the invention produces higher levels of non-polar lipids such as DAG or TAG, preferably both, than a corresponding vegetative plant part, non-human organism or part thereof. In one example, transgenic plants of the invention produce seeds, leaves, leaf portions of at least 1 cm² in surface area, stems and/or tubers having an increased non-polar lipid content such as DAG or TAG, preferably both, when
 15 compared to corresponding seeds, leaves, leaf portions of at least 1 cm² in surface area, stems or tubers. The non-polar lipid content of the vegetative plant part, non-human organism or part thereof is at 0.5% greater on a weight basis when compared to a corresponding non-human organism or part thereof, or as further defined in Feature (i).

20 In another embodiment, the vegetative plant part, transgenic non-human organism or part thereof, preferably a plant or seed, produce DAGs and/or TAGs that are enriched for one or more particular fatty acids. A wide spectrum of fatty acids can be incorporated into DAGs and/or TAGs, including saturated and unsaturated fatty acids and short-chain and long-chain fatty acids. Some non-limiting examples of fatty
 25 acids that can be incorporated into DAGs and/or TAGs and which may be increased in level include: capric (10:0), lauric (12:0), myristic (14:0), palmitic (16:0), palmitoleic (16:1), stearic (18:0), oleic (18:1), vaccenic (18:1), linoleic (18:2), eleostearic (18:3), γ -linolenic (18:3), α -linolenic (18:3 ω 3), stearidonic (18:4 ω 3), arachidic (20:0), eicosadienoic (20:2), dihomogamma-linoleic (20:3), eicosatrienoic (20:3),
 30 arachidonic (20:4), eicosatetraenoic (20:4), eicosapentaenoic (20:5 ω 3), behenic (22:0), docosapentaenoic (22:5 ω), docosahexaenoic (22:6 ω 3), lignoceric (24:0), nervonic (24:1), cerotic (26:0), and montanic (28:0) fatty acids. In one embodiment of the present invention, the vegetative plant part, transgenic organism or parts thereof is enriched for DAGs and/or TAGs comprising oleic acid, or polyunsaturated fatty
 35 acids.

In one embodiment of the invention, the vegetative plant part, transgenic non-human organism or part thereof, preferably a plant or seed, is transformed with a chimeric DNA which encodes an MGAT which may or may not have DGAT activity.

Expression of the MGAT preferably results in higher levels of non-polar lipids such as DAG or TAG and/or increased non-polar lipid yield in said vegetative plant part, transgenic non-human organism or part thereof. In a preferred embodiment, the transgenic non-human organism is a plant.

5 In a further embodiment, the vegetative plant part, transgenic non-human organism or part thereof is transformed with a chimeric DNA which encodes a GPAT or a DGAT. Preferably, the vegetative plant part or transgenic non-human organism is transformed with both chimeric DNAs, which are preferably covalently linked on one DNA molecule such as, for example, a single T-DNA molecule.

10 Yang et al. (2010) describe two glycerol-3-phosphate acyltransferases (GPAT4 and GPAT6) from *Arabidopsis* with *sn*-2 preference and phosphatase activity that are able to produce *sn*-2 MAG from glycerol-3-phosphate (G-3-P) (Figure 1). These enzymes are proposed to be part of the cutin synthesis pathway. *Arabidopsis* GPAT4 and GPAT6 have been shown to use acyl-CoA as a fatty acid substrate (Zheng et al.,
15 2003).

Combining a bifunctional GPAT/phosphatase with a MGAT yields a novel DAG synthesis pathway using G-3-P as one substrate and two acyl groups derived from acyl-CoA as the other substrates. Similarly, combining such a bifunctional GPAT/phosphatase with a MGAT which has DGAT activity yields a novel TAG
20 synthesis pathway using glycerol-3-phosphate as one substrate and three acyl groups derived from acyl-CoA as other substrates.

Accordingly, in one embodiment of the invention, the vegetative plant part, transgenic non-human organism or part thereof is co-transformed with a bifunctional GPAT/phosphatase and with a MGAT which does not have DGAT activity. This
25 would result in the production of MAG by the bifunctional GPAT/phosphatase which would then be converted to DAG by the MGAT and then TAG by a native DGAT or other activity. Novel DAG production could be confirmed and selected for by, for example, performing such a co-transformation in a yeast strain containing lethal *SLC1* + *SLC4* knockouts such as that described by Benghezal et al. (2007; Figure 2). Figure
30 2 of Benghezal et al. (2007) shows that knocking out the two yeast LPATS (*SLC1* & *SLC4*) is lethal. The *SLC1* + *SLC4* double yeast mutant can only be maintained because of a complementing plasmid which provides one of the *slc* genes (*SLC1* in their case) in trans. Negative selection by adding FOA to the medium results in the loss of this complementing plasmid (counterselection of the Ura selection marker) and
35 renders the cells non viable.

In another embodiment of the invention, the vegetative plant part, transgenic non-human organism or part thereof, preferably a plant or seed, is co-transformed with chimeric DNAs encoding a bifunctional GPAT/phosphatase and a MGAT which

has DGAT activity. This would result in the production of MAG by the bifunctional GPAT/phosphatase which would then be converted to DAG and then TAG by the MGAT.

In a further embodiment, one or more endogenous GPATs with no detectable phosphatase activity are silenced, for example one or more genes encoding GPATs that acylate glycerol-3-phosphate to form LPA in the Kennedy Pathway (for example, *Arabidopsis* GPAT1) is silenced.

In another embodiment, the vegetative plant part, transgenic non-human organism or part thereof, preferably a plant or seed, is transformed with a chimeric DNAs encoding a DGAT1, a DGAT2, a Wrinkled 1 (WRI1) transcription factor, an Oleosin, or a silencing suppressor polypeptide. The chimeric DNAs are preferably covalently linked on one DNA molecule such as, for example, a single T-DNA molecule, and the vegetative plant part, transgenic non-human organism or part thereof is preferably homozygous for the one DNA molecule inserted into its genome.

Substrate preferences could be engineered into the novel DAG and TAG synthesis pathways by, for example, supplying transgenic H1246 yeast strains expressing MGAT variants with a concentration of a particular free fatty acid (for example, DHA) that prevents complementation by the wildtype MGAT gene. Only the variants able to use the supplied free fatty acid would grow. Several cycles of MGAT engineering would result in the production of a MGAT with increased preference for particular fatty acids.

The various Kennedy Pathway complementations and supplementations described above could be performed in any cell type due to the ubiquitous nature of the initial substrate glycerol-3-phosphate. In one embodiment, the use of transgenes results in increased oil yields.

Polynucleotides

The terms "polynucleotide", and "nucleic acid" are used interchangeably. They refer to a polymeric form of nucleotides of any length, either deoxyribonucleotides or ribonucleotides, or analogs thereof. A polynucleotide of the invention may be of genomic, cDNA, semisynthetic, or synthetic origin, double-stranded or single-stranded and by virtue of its origin or manipulation: (1) is not associated with all or a portion of a polynucleotide with which it is associated in nature, (2) is linked to a polynucleotide other than that to which it is linked in nature, or (3) does not occur in nature. The following are non-limiting examples of polynucleotides: coding or non-coding regions of a gene or gene fragment, loci (locus) defined from linkage analysis, exons, introns, messenger RNA (mRNA), transfer RNA (tRNA), ribosomal RNA (rRNA), ribozymes, cDNA, recombinant

polynucleotides, branched polynucleotides, plasmids, vectors, isolated DNA of any sequence, isolated RNA of any sequence, chimeric DNA of any sequence, nucleic acid probes, and primers. A polynucleotide may comprise modified nucleotides such as methylated nucleotides and nucleotide analogs. If present, modifications to the nucleotide structure may be imparted before or after assembly of the polymer. The sequence of nucleotides may be interrupted by non-nucleotide components. A polynucleotide may be further modified after polymerization such as by conjugation with a labeling component.

By "isolated polynucleotide" it is meant a polynucleotide which has generally been separated from the polynucleotide sequences with which it is associated or linked in its native state. Preferably, the isolated polynucleotide is at least 60% free, more preferably at least 75% free, and more preferably at least 90% free from the polynucleotide sequences with which it is naturally associated or linked.

As used herein, the term "gene" is to be taken in its broadest context and includes the deoxyribonucleotide sequences comprising the transcribed region and, if translated, the protein coding region, of a structural gene and including sequences located adjacent to the coding region on both the 5' and 3' ends for a distance of at least about 2 kb on either end and which are involved in expression of the gene. In this regard, the gene includes control signals such as promoters, enhancers, termination and/or polyadenylation signals that are naturally associated with a given gene, or heterologous control signals, in which case, the gene is referred to as a "chimeric gene". The sequences which are located 5' of the protein coding region and which are present on the mRNA are referred to as 5' non-translated sequences. The sequences which are located 3' or downstream of the protein coding region and which are present on the mRNA are referred to as 3' non-translated sequences. The term "gene" encompasses both cDNA and genomic forms of a gene. A genomic form or clone of a gene contains the coding region which may be interrupted with non-coding sequences termed "introns", "intervening regions", or "intervening sequences." Introns are segments of a gene which are transcribed into nuclear RNA (nRNA). Introns may contain regulatory elements such as enhancers. Introns are removed or "spliced out" from the nuclear or primary transcript; introns therefore are absent in the mRNA transcript. The mRNA functions during translation to specify the sequence or order of amino acids in a nascent polypeptide. The term "gene" includes a synthetic or fusion molecule encoding all or part of the proteins of the invention described herein and a complementary nucleotide sequence to any one of the above.

As used herein, "chimeric DNA" refers to any DNA molecule that is not naturally found in nature; also referred to herein as a "DNA construct". Typically, chimeric DNA comprises regulatory and transcribed or protein coding sequences that

are not naturally found together in nature. Accordingly, chimeric DNA may comprise regulatory sequences and coding sequences that are derived from different sources, or regulatory sequences and coding sequences derived from the same source, but arranged in a manner different than that found in nature. The open reading frame may
5 or may not be linked to its natural upstream and downstream regulatory elements. The open reading frame may be incorporated into, for example, the plant genome, in a non-natural location, or in a replicon or vector where it is not naturally found such as a bacterial plasmid or a viral vector. The term "chimeric DNA" is not limited to DNA molecules which are replicable in a host, but includes DNA capable of being ligated
10 into a replicon by, for example, specific adaptor sequences.

A "transgene" is a gene that has been introduced into the genome by a transformation procedure. The term includes a gene in a progeny cell, plant, seed, non-human organism or part thereof which was introduced into the genome of a progenitor cell thereof. Such progeny cells etc may be at least a 3rd or 4th generation
15 progeny from the progenitor cell which was the primary transformed cell. Progeny may be produced by sexual reproduction or vegetatively such as, for example, from tubers in potatoes or ratoons in sugarcane. The term "genetically modified", and variations thereof, is a broader term that includes introducing a gene into a cell by transformation or transduction, mutating a gene in a cell and genetically altering or
20 modulating the regulation of a gene in a cell, or the progeny of any cell modified as described above.

A "genomic region" as used herein refers to a position within the genome where a transgene, or group of transgenes (also referred to herein as a cluster), have been inserted into a cell, or predecessor thereof. Such regions only comprise
25 nucleotides that have been incorporated by the intervention of man such as by methods described herein.

A "recombinant polynucleotide" of the invention refers to a nucleic acid molecule which has been constructed or modified by artificial recombinant methods. The recombinant polynucleotide may be present in a cell in an altered amount or
30 expressed at an altered rate (e.g., in the case of mRNA) compared to its native state. In one embodiment, the polynucleotide is introduced into a cell that does not naturally comprise the polynucleotide. Typically an exogenous DNA is used as a template for transcription of mRNA which is then translated into a continuous sequence of amino acid residues coding for a polypeptide of the invention within the transformed cell. In
35 another embodiment, the polynucleotide is endogenous to the cell and its expression is altered by recombinant means, for example, an exogenous control sequence is introduced upstream of an endogenous gene of interest to enable the transformed cell to express the polypeptide encoded by the gene.

A recombinant polynucleotide of the invention includes polynucleotides which have not been separated from other components of the cell-based or cell-free expression system, in which it is present, and polynucleotides produced in said cell-based or cell-free systems which are subsequently purified away from at least some other components. The polynucleotide can be a contiguous stretch of nucleotides existing in nature, or comprise two or more contiguous stretches of nucleotides from different sources (naturally occurring and/or synthetic) joined to form a single polynucleotide. Typically, such chimeric polynucleotides comprise at least an open reading frame encoding a polypeptide of the invention operably linked to a promoter suitable of driving transcription of the open reading frame in a cell of interest.

With regard to the defined polynucleotides, it will be appreciated that % identity figures higher than those provided above will encompass preferred embodiments. Thus, where applicable, in light of the minimum % identity figures, it is preferred that the polynucleotide comprises a polynucleotide sequence which is at least 60%, more preferably at least 65%, more preferably at least 70%, more preferably at least 75%, more preferably at least 80%, more preferably at least 85%, more preferably at least 90%, more preferably at least 91%, more preferably at least 92%, more preferably at least 93%, more preferably at least 94%, more preferably at least 95%, more preferably at least 96%, more preferably at least 97%, more preferably at least 98%, more preferably at least 99%, more preferably at least 99.1%, more preferably at least 99.2%, more preferably at least 99.3%, more preferably at least 99.4%, more preferably at least 99.5%, more preferably at least 99.6%, more preferably at least 99.7%, more preferably at least 99.8%, and even more preferably at least 99.9% identical to the relevant nominated SEQ ID NO.

A polynucleotide of, or useful for, the present invention, may selectively hybridise, under stringent conditions, to a polynucleotide defined herein. As used herein, stringent conditions are those that: (1) employ during hybridisation a denaturing agent such as formamide, for example, 50% (v/v) formamide with 0.1% (w/v) bovine serum albumin, 0.1% Ficoll, 0.1% polyvinylpyrrolidone, 50 mM sodium phosphate buffer at pH 6.5 with 750 mM NaCl, 75 mM sodium citrate at 42°C; or (2) employ 50% formamide, 5 x SSC (0.75 M NaCl, 0.075 M sodium citrate), 50 mM sodium phosphate (pH 6.8), 0.1% sodium pyrophosphate, 5 x Denhardt's solution, sonicated salmon sperm DNA (50 g/ml), 0.1% SDS and 10% dextran sulfate at 42°C in 0.2 x SSC and 0.1% SDS, and/or (3) employ low ionic strength and high temperature for washing, for example, 0.015 M NaCl/0.0015 M sodium citrate/0.1% SDS at 50°C.

Polynucleotides of the invention may possess, when compared to naturally occurring molecules, one or more mutations which are deletions, insertions, or

substitutions of nucleotide residues. Polynucleotides which have mutations relative to a reference sequence can be either naturally occurring (that is to say, isolated from a natural source) or synthetic (for example, by performing site-directed mutagenesis or DNA shuffling on the nucleic acid as described above).

5

Polynucleotide for Reducing Expression Levels of Endogenous Proteins

RNA Interference

RNA interference (RNAi) is particularly useful for specifically inhibiting the
10 production of a particular protein. Although not wishing to be limited by theory, Waterhouse et al. (1998) have provided a model for the mechanism by which dsRNA (duplex RNA) can be used to reduce protein production. This technology relies on the presence of dsRNA molecules that contain a sequence that is essentially identical to the mRNA of the gene of interest or part thereof. Conveniently, the dsRNA can be
15 produced from a single promoter in a recombinant vector or host cell, where the sense and anti-sense sequences are flanked by an unrelated sequence which enables the sense and anti-sense sequences to hybridize to form the dsRNA molecule with the unrelated sequence forming a loop structure. The design and production of suitable dsRNA molecules is well within the capacity of a person skilled in the art, particularly
20 considering Waterhouse et al. (1998), Smith et al. (2000), WO 99/32619, WO 99/53050, WO 99/49029, and WO 01/34815.

In one example, a DNA is introduced that directs the synthesis of an at least partly double stranded RNA product(s) with homology to the target gene to be inactivated. The DNA therefore comprises both sense and antisense sequences that,
25 when transcribed into RNA, can hybridize to form the double stranded RNA region. In one embodiment of the invention, the sense and antisense sequences are separated by a spacer region that comprises an intron which, when transcribed into RNA, is spliced out. This arrangement has been shown to result in a higher efficiency of gene silencing. The double stranded region may comprise one or two RNA molecules,
30 transcribed from either one DNA region or two. The presence of the double stranded molecule is thought to trigger a response from an endogenous system that destroys both the double stranded RNA and also the homologous RNA transcript from the target gene, efficiently reducing or eliminating the activity of the target gene.

The length of the sense and antisense sequences that hybridize should each be
35 at least 19 contiguous nucleotides. The full-length sequence corresponding to the entire gene transcript may be used. The degree of identity of the sense and antisense sequences to the targeted transcript should be at least 85%, at least 90%, or at least 95-100%. The RNA molecule may of course comprise unrelated sequences which

may function to stabilize the molecule. The RNA molecule may be expressed under the control of a RNA polymerase II or RNA polymerase III promoter. Examples of the latter include tRNA or snRNA promoters.

Preferred small interfering RNA ("siRNA") molecules comprise a nucleotide
 5 sequence that is identical to about 19-21 contiguous nucleotides of the target mRNA. Preferably, the siRNA sequence commences with the dinucleotide AA, comprises a GC-content of about 30-70% (preferably, 30-60%, more preferably 40-60% and more preferably about 45%-55%), and does not have a high percentage identity to any nucleotide sequence other than the target in the genome of the organism in which it is
 10 to be introduced, for example, as determined by standard BLAST search.

microRNA

MicroRNAs (abbreviated miRNAs) are generally 19-25 nucleotides (commonly about 20-24 nucleotides in plants) non-coding RNA molecules that are
 15 derived from larger precursors that form imperfect stem-loop structures.

miRNAs bind to complementary sequences on target messenger RNA transcripts (mRNAs), usually resulting in translational repression or target degradation and gene silencing.

In plant cells, miRNA precursor molecules are believed to be largely processed
 20 in the nucleus. The pri-miRNA (containing one or more local double-stranded or "hairpin" regions as well as the usual 5' "cap" and polyadenylated tail of an mRNA) is processed to a shorter miRNA precursor molecule that also includes a stem-loop or fold-back structure and is termed the "pre-miRNA". In plants, the pre-miRNAs are cleaved by distinct DICER-like (DCL) enzymes, yielding miRNA:miRNA* duplexes.
 25 Prior to transport out of the nucleus, these duplexes are methylated.

In the cytoplasm, the miRNA strand from the miRNA:miRNA duplex is selectively incorporated into an active RNA-induced silencing complex (RISC) for target recognition. The RISC- complexes contain a particular subset of Argonaute proteins that exert sequence-specific gene repression (see, for example, Millar and
 30 Waterhouse, 2005; Pasquinelli et al., 2005; Almeida and Allshire, 2005).

Cosuppression

Genes can suppress the expression of related endogenous genes and/or transgenes already present in the genome, a phenomenon termed homology-dependent
 35 gene silencing. Most of the instances of homology-dependent gene silencing fall into two classes - those that function at the level of transcription of the transgene, and those that operate post-transcriptionally.

Post-transcriptional homology-dependent gene silencing (i.e., cosuppression) describes the loss of expression of a transgene and related endogenous or viral genes in transgenic plants. Cosuppression often, but not always, occurs when transgene transcripts are abundant, and it is generally thought to be triggered at the level of
5 mRNA processing, localization, and/or degradation. Several models exist to explain how cosuppression works (see in Taylor, 1997).

One model, the "quantitative" or "RNA threshold" model, proposes that cells can cope with the accumulation of large amounts of transgene transcripts, but only up to a point. Once that critical threshold has been crossed, the sequence-dependent
10 degradation of both transgene and related endogenous gene transcripts is initiated. It has been proposed that this mode of cosuppression may be triggered following the synthesis of copy RNA (cRNA) molecules by reverse transcription of the excess transgene mRNA, presumably by endogenous RNA-dependent RNA polymerases. These cRNAs may hybridize with transgene and endogenous mRNAs, the unusual
15 hybrids targeting homologous transcripts for degradation. However, this model does not account for reports suggesting that cosuppression can apparently occur in the absence of transgene transcription and/or without the detectable accumulation of transgene transcripts.

To account for these data, a second model, the "qualitative" or "aberrant RNA"
20 model, proposes that interactions between transgene RNA and DNA and/or between endogenous and introduced DNAs lead to the methylation of transcribed regions of the genes. The methylated genes are proposed to produce RNAs that are in some way aberrant, their anomalous features triggering the specific degradation of all related transcripts. Such aberrant RNAs may be produced by complex transgene loci,
25 particularly those that contain inverted repeats.

A third model proposes that intermolecular base pairing between transcripts, rather than cRNA-mRNA hybrids generated through the action of an RNA-dependent RNA polymerase, may trigger cosuppression. Such base pairing may become more common as transcript levels rise, the putative double-stranded regions triggering the
30 targeted degradation of homologous transcripts. A similar model proposes intramolecular base pairing instead of intermolecular base pairing between transcripts.

Cosuppression involves introducing an extra copy of a gene or a fragment thereof into a plant in the sense orientation with respect to a promoter for its expression. A skilled person would appreciate that the size of the sense fragment, its
35 correspondence to target gene regions, and its degree of sequence identity to the target gene can vary. In some instances, the additional copy of the gene sequence interferes with the expression of the target plant gene. Reference is made to WO 97/20936 and EP 0465572 for methods of implementing co-suppression approaches.

Antisense Polynucleotides

The term "antisense polynucleotide" shall be taken to mean a DNA or RNA, or combination thereof, molecule that is complementary to at least a portion of a specific mRNA molecule encoding an endogenous polypeptide and capable of interfering with a post-transcriptional event such as mRNA translation. The use of antisense methods is well known in the art (see for example, G. Hartmann and S. Endres, *Manual of Antisense Methodology*, Kluwer (1999)). The use of antisense techniques in plants has been reviewed by Bourque (1995) and Senior (1998). Bourque (1995) lists a large number of examples of how antisense sequences have been utilized in plant systems as a method of gene inactivation. Bourque also states that attaining 100% inhibition of any enzyme activity may not be necessary as partial inhibition will more than likely result in measurable change in the system. Senior (1998) states that antisense methods are now a very well established technique for manipulating gene expression.

In one embodiment, the antisense polynucleotide hybridises under physiological conditions, that is, the antisense polynucleotide (which is fully or partially single stranded) is at least capable of forming a double stranded polynucleotide with mRNA encoding a protein such as an endogenous enzyme, for example, DGAT, GPAT, LPAA, LPCAT, PAP, AGPase, under normal conditions in a cell.

Antisense molecules may include sequences that correspond to the structural genes or for sequences that effect control over the gene expression or splicing event. For example, the antisense sequence may correspond to the targeted coding region of endogenous gene, or the 5'-untranslated region (UTR) or the 3'-UTR or combination of these. It may be complementary in part to intron sequences, which may be spliced out during or after transcription, preferably only to exon sequences of the target gene. In view of the generally greater divergence of the UTRs, targeting these regions provides greater specificity of gene inhibition.

The length of the antisense sequence should be at least 19 contiguous nucleotides, preferably at least 50 nucleotides, and more preferably at least 100, 200, 500 or 1000 nucleotides. The full-length sequence complementary to the entire gene transcript may be used. The length is most preferably 100-2000 nucleotides. The degree of identity of the antisense sequence to the targeted transcript should be at least 90% and more preferably 95-100%. The antisense RNA molecule may of course comprise unrelated sequences which may function to stabilize the molecule.

Catalytic Polynucleotides

The term "catalytic polynucleotide" refers to a DNA molecule or DNA-containing molecule (also known in the art as a "deoxyribozyme") or an RNA or RNA-containing molecule (also known as a "ribozyme") which specifically
5 recognizes a distinct substrate and catalyses the chemical modification of this substrate. The nucleic acid bases in the catalytic nucleic acid can be bases A, C, G, T (and U for RNA).

Typically, the catalytic nucleic acid contains an antisense sequence for specific recognition of a target nucleic acid, and a nucleic acid cleaving enzymatic activity
10 (also referred to herein as the "catalytic domain"). The types of ribozymes that are particularly useful in this invention are hammerhead ribozymes (Haseloff and Gerlach, 1988; Perriman et al., 1992) and hairpin ribozymes (Zolotukhin et al., 1996; Klein et al., 1998; Shippy et al., 1999).

Ribozymes useful in the invention and DNA encoding the ribozymes can be
15 chemically synthesized using methods well known in the art. The ribozymes can also be prepared from a DNA molecule (that upon transcription, yields an RNA molecule) operably linked to an RNA polymerase promoter, for example, the promoter for T7 RNA polymerase or SP6 RNA polymerase. In a separate embodiment, the DNA can be inserted into an expression cassette or transcription cassette. After synthesis, the
20 RNA molecule can be modified by ligation to a DNA molecule having the ability to stabilize the ribozyme and make it resistant to RNase.

As with antisense oligonucleotides, small interfering RNA and microRNA described herein, catalytic polynucleotides useful in the invention should be capable of "hybridizing" the target nucleic acid molecule under "physiological conditions",
25 namely those conditions within a plant, algal or fungal cell.

Recombinant Vectors

One embodiment of the present invention includes a recombinant vector, which comprises at least one polynucleotide defined herein and is capable of
30 delivering the polynucleotide into a host cell. Recombinant vectors include expression vectors. Recombinant vectors contain heterologous polynucleotide sequences, that is, polynucleotide sequences that are not naturally found adjacent to a polynucleotide defined herein, that preferably, are derived from a different species. The vector can be either RNA or DNA, either prokaryotic or eukaryotic, and typically
35 is a viral vector, derived from a virus, or a plasmid. Plasmid vectors typically include additional nucleic acid sequences that provide for easy selection, amplification, and transformation of the expression cassette in prokaryotic cells, e.g., pUC-derived vectors, pSK-derived vectors, pGEM-derived vectors, pSP-derived vectors, pBS-

derived vectors, or binary vectors containing one or more T-DNA regions. Additional nucleic acid sequences include origins of replication to provide for autonomous replication of the vector, selectable marker genes, preferably encoding antibiotic or herbicide resistance, unique multiple cloning sites providing for multiple sites to
5 insert nucleic acid sequences or genes encoded in the nucleic acid construct, and sequences that enhance transformation of prokaryotic and eukaryotic (especially plant) cells.

"Operably linked" as used herein, refers to a functional relationship between two or more nucleic acid (e.g., DNA) segments. Typically, it refers to the functional
10 relationship of transcriptional regulatory element (promoter) to a transcribed sequence. For example, a promoter is operably linked to a coding sequence of a polynucleotide defined herein, if it stimulates or modulates the transcription of the coding sequence in an appropriate cell. Generally, promoter transcriptional regulatory elements that are operably linked to a transcribed sequence are physically contiguous
15 to the transcribed sequence, i.e., they are *cis*-acting. However, some transcriptional regulatory elements such as enhancers, need not be physically contiguous or located in close proximity to the coding sequences whose transcription they enhance.

When there are multiple promoters present, each promoter may independently be the same or different.

20 Recombinant vectors may also contain: (a) one or more secretory signals which encode signal peptide sequences, to enable an expressed polypeptide defined herein to be secreted from the cell that produces the polypeptide, or which provide for localisation of the expressed polypeptide, for example, for retention of the polypeptide in the endoplasmic reticulum (ER) in the cell, or transfer into a plastid, and/or (b)
25 contain fusion sequences which lead to the expression of nucleic acid molecules as fusion proteins. Examples of suitable signal segments include any signal segment capable of directing the secretion or localisation of a polypeptide defined herein. Preferred signal segments include, but are not limited to, *Nicotiana nectarin* signal peptide (US 5,939,288), tobacco extensin signal, or the soy oleosin oil body binding
30 protein signal. Recombinant vectors may also include intervening and/or untranslated sequences surrounding and/or within the nucleic acid sequence of a polynucleotide defined herein.

To facilitate identification of transformants, the recombinant vector desirably comprises a selectable or screenable marker gene as, or in addition to, the nucleic acid
35 sequence of a polynucleotide defined herein. By "marker gene" is meant a gene that imparts a distinct phenotype to cells expressing the marker gene and thus, allows such transformed cells to be distinguished from cells that do not have the marker. A selectable marker gene confers a trait for which one can "select" based on resistance

to a selective agent (e.g., a herbicide, antibiotic, radiation, heat, or other treatment damaging to untransformed cells). A screenable marker gene (or reporter gene) confers a trait that one can identify through observation or testing, that is, by "screening" (e.g., β -glucuronidase, luciferase, GFP or other enzyme activity not
5 present in untransformed cells). The marker gene and the nucleotide sequence of interest do not have to be linked, since co-transformation of unlinked genes as for example, described in US 4,399,216, is also an efficient process in for example, plant transformation. The actual choice of a marker is not crucial as long as it is functional (i.e., selective) in combination with the cells of choice such as a plant cell.

10 Examples of bacterial selectable markers are markers that confer antibiotic resistance such as ampicillin, erythromycin, chloramphenicol, or tetracycline resistance, preferably kanamycin resistance. Exemplary selectable markers for selection of plant transformants include, but are not limited to, a *hyg* gene which encodes hygromycin B resistance; a neomycin phosphotransferase (*nptII*) gene
15 conferring resistance to kanamycin, paromomycin, G418; a glutathione-S-transferase gene from rat liver conferring resistance to glutathione derived herbicides as for example, described in EP 256223; a glutamine synthetase gene conferring, upon overexpression, resistance to glutamine synthetase inhibitors such as phosphinothricin as for example, described in WO 87/05327; an acetyltransferase gene from
20 *Streptomyces viridochromogenes* conferring resistance to the selective agent phosphinothricin as for example, described in EP 275957; a gene encoding a 5-enolshikimate-3-phosphate synthase (EPSPS) conferring tolerance to N-phosphonomethylglycine as for example, described by Hinchee et al. (1988); a *bar* gene conferring resistance against bialaphos as for example, described in
25 WO91/02071; a nitrilase gene such as *bxn* from *Klebsiella ozaenae* which confers resistance to bromoxynil (Stalker et al., 1988); a dihydrofolate reductase (DHFR) gene conferring resistance to methotrexate (Thillet et al., 1988); a mutant acetolactate synthase gene (ALS) which confers resistance to imidazolinone, sulfonylurea, or other ALS-inhibiting chemicals (EP 154,204); a mutated anthranilate synthase gene that
30 confers resistance to 5-methyl tryptophan; or a dalapon dehalogenase gene that confers resistance to the herbicide.

Preferred screenable markers include, but are not limited to, a *uidA* gene encoding a β -glucuronidase (GUS) enzyme for which various chromogenic substrates are known; a β -galactosidase gene encoding an enzyme for which chromogenic
35 substrates are known; an aequorin gene (Prasher et al., 1985) which may be employed in calcium-sensitive bioluminescence detection; a green fluorescent protein gene (Niedz et al., 1995) or derivatives thereof; or a luciferase (*luc*) gene (Ow et al., 1986) which allows for bioluminescence detection. By "reporter molecule" it is meant a

molecule that, by its chemical nature, provides an analytically identifiable signal that facilitates determination of promoter activity by reference to protein product.

Preferably, the recombinant vector is stably incorporated into the genome of the cell such as the plant cell. Accordingly, the recombinant vector may comprise
5 appropriate elements which allow the vector to be incorporated into the genome, or into a chromosome of the cell.

Expression Vector

As used herein, an "expression vector" is a DNA or RNA vector that is capable
10 of transforming a host cell and of effecting expression of one or more specified polynucleotides. Preferably, the expression vector is also capable of replicating within the host cell. Expression vectors can be either prokaryotic or eukaryotic, and are typically viruses or plasmids. Expression vectors of the present invention include any vectors that function (i.e., direct gene expression) in host cells of the present
15 invention, including in bacterial, fungal, endoparasite, arthropod, animal, algal, and plant cells. Particularly preferred expression vectors of the present invention can direct gene expression in yeast, algae and/or plant cells.

Expression vectors of the present invention contain regulatory sequences such as transcription control sequences, translation control sequences, origins of
20 replication, and other regulatory sequences that are compatible with the host cell and that control the expression of polynucleotides of the present invention. In particular, expression vectors of the present invention include transcription control sequences. Transcription control sequences are sequences which control the initiation, elongation, and termination of transcription. Particularly important transcription control
25 sequences are those which control transcription initiation such as promoter, enhancer, operator and repressor sequences. Suitable transcription control sequences include any transcription control sequence that can function in at least one of the recombinant cells of the present invention. The choice of the regulatory sequences used depends on the target organism such as a plant and/or target organ or tissue of interest. Such
30 regulatory sequences may be obtained from any eukaryotic organism such as plants or plant viruses, or may be chemically synthesized. A variety of such transcription control sequences are known to those skilled in the art. Particularly preferred transcription control sequences are promoters active in directing transcription in plants, either constitutively or stage and/or tissue specific, depending on the use of the
35 plant or part(s) thereof.

A number of vectors suitable for stable transfection of plant cells or for the establishment of transgenic plants have been described in for example, Pouwels et al., Cloning Vectors: A Laboratory Manual, 1985, supp. 1987, Weissbach and Weissbach,

Methods for Plant Molecular Biology, Academic Press, 1989, and Gelvin et al., Plant Molecular Biology Manual, Kluwer Academic Publishers, 1990. Typically, plant expression vectors include for example, one or more cloned plant genes under the transcriptional control of 5' and 3' regulatory sequences and a dominant selectable
 5 marker. Such plant expression vectors also can contain a promoter regulatory region (e.g., a regulatory region controlling inducible or constitutive, environmentally- or developmentally-regulated, or cell- or tissue-specific expression), a transcription initiation start site, a ribosome binding site, an RNA processing signal, a transcription termination site, and/or a polyadenylation signal.

10 A number of constitutive promoters that are active in plant cells have been described. Suitable promoters for constitutive expression in plants include, but are not limited to, the cauliflower mosaic virus (CaMV) 35S promoter, the Figwort mosaic virus (FMV) 35S, the sugarcane bacilliform virus promoter, the commelina yellow mottle virus promoter, the light-inducible promoter from the small subunit of
 15 the ribulose-1,5-bis-phosphate carboxylase, the rice cytosolic triosephosphate isomerase promoter, the adenine phosphoribosyltransferase promoter of *Arabidopsis*, the rice actin 1 gene promoter, the mannopine synthase and octopine synthase promoters, the Adh promoter, the sucrose synthase promoter, the R gene complex promoter, and the chlorophyll α/β binding protein gene promoter. These promoters
 20 have been used to create DNA vectors that have been expressed in plants, see for example, WO 84/02913. All of these promoters have been used to create various types of plant-expressible recombinant DNA vectors.

For the purpose of expression in source tissues of the plant such as the leaf, seed, root or stem, it is preferred that the promoters utilized in the present invention
 25 have relatively high expression in these specific tissues. For this purpose, one may choose from a number of promoters for genes with tissue- or cell-specific, or -enhanced expression. Examples of such promoters reported in the literature include, the chloroplast glutamine synthetase GS2 promoter from pea, the chloroplast fructose-1,6-biphosphatase promoter from wheat, the nuclear photosynthetic ST-LS1 promoter
 30 from potato, the serine/threonine kinase promoter and the glucoamylase (CHS) promoter from *Arabidopsis thaliana*. Also reported to be active in photosynthetically active tissues are the ribulose-1,5-bisphosphate carboxylase promoter from eastern larch (*Larix laricina*), the promoter for the Cab gene, Cab6, from pine, the promoter for the Cab-1 gene from wheat, the promoter for the Cab-1 gene from spinach, the
 35 promoter for the Cab 1R gene from rice, the pyruvate, orthophosphate dikinase (PPDK) promoter from *Zea mays*, the promoter for the tobacco Lhcb1*2 gene, the *Arabidopsis thaliana* Suc2 sucrose-H³⁰ symporter promoter, and the promoter for the thylakoid membrane protein genes from spinach (PsaD, PsaF, PsaE, PC, FNR, AtpC,

AtpD, Cab, RbcS). Other promoters for the chlorophyll α/β -binding proteins may also be utilized in the present invention such as the promoters for LhcB gene and PsbP gene from white mustard (*Sinapis alba*).

A variety of plant gene promoters that are regulated in response to environmental, hormonal, chemical, and/or developmental signals, also can be used for expression of RNA-binding protein genes in plant cells, including promoters regulated by (1) heat, (2) light (e.g., pea RbcS-3A promoter, maize RbcS promoter), (3) hormones such as abscisic acid, (4) wounding (e.g., WunI), or (5) chemicals such as methyl jasmonate, salicylic acid, steroid hormones, alcohol, Safeners (WO 97/06269), or it may also be advantageous to employ (6) organ-specific promoters.

As used herein, the term "plant storage organ specific promoter" refers to a promoter that preferentially, when compared to other plant tissues, directs gene transcription in a storage organ of a plant. Preferably, the promoter only directs expression of a gene of interest in the storage organ, and/or expression of the gene of interest in other parts of the plant such as leaves is not detectable by Northern blot analysis and/or RT-PCR. Typically, the promoter drives expression of genes during growth and development of the storage organ, in particular during the phase of synthesis and accumulation of storage compounds in the storage organ. Such promoters may drive gene expression in the entire plant storage organ or only part thereof such as the seedcoat, embryo or cotyledon(s) in seeds of dicotyledonous plants or the endosperm or aleurone layer of seeds of monocotyledonous plants.

For the purpose of expression in sink tissues of the plant such as the tuber of the potato plant, the fruit of tomato, or the seed of soybean, canola, cotton, *Zea mays*, wheat, rice, and barley, it is preferred that the promoters utilized in the present invention have relatively high expression in these specific tissues. A number of promoters for genes with tuber-specific or -enhanced expression are known, including the class I patatin promoter, the promoter for the potato tuber ADPGPP genes, both the large and small subunits, the sucrose synthase promoter, the promoter for the major tuber proteins, including the 22 kD protein complexes and proteinase inhibitors, the promoter for the granule bound starch synthase gene (GBSS), and other class I and II patatins promoters. Other promoters can also be used to express a protein in specific tissues such as seeds or fruits. The promoter for β -conglycinin or other seed-specific promoters such as the napin, zein, linin and phaseolin promoters, can be used. Root specific promoters may also be used. An example of such a promoter is the promoter for the acid chitinase gene. Expression in root tissue could also be accomplished by utilizing the root specific subdomains of the CaMV 35S promoter that have been identified.

In a particularly preferred embodiment, the promoter directs expression in tissues and organs in which lipid biosynthesis take place. Such promoters may act in seed development at a suitable time for modifying lipid composition in seeds.

In an embodiment, the promoter is a plant storage organ specific promoter. In one embodiment, the plant storage organ specific promoter is a seed specific promoter. In a more preferred embodiment, the promoter preferentially directs expression in the cotyledons of a dicotyledonous plant or in the endosperm of a monocotyledonous plant, relative to expression in the embryo of the seed or relative to other organs in the plant such as leaves. Preferred promoters for seed-specific expression include: 1) promoters from genes encoding enzymes involved in lipid biosynthesis and accumulation in seeds such as desaturases and elongases, 2) promoters from genes encoding seed storage proteins, and 3) promoters from genes encoding enzymes involved in carbohydrate biosynthesis and accumulation in seeds. Seed specific promoters which are suitable are, the oilseed rape napin gene promoter (US 5,608,152), the *Vicia faba* USP promoter (Baumlein et al., 1991), the *Arabidopsis* oleosin promoter (WO 98/45461), the *Phaseolus vulgaris* phaseolin promoter (US 5,504,200), the *Brassica* Bce4 promoter (WO 91/13980), or the legumin B4 promoter (Baumlein et al., 1992), and promoters which lead to the seed-specific expression in monocots such as maize, barley, wheat, rye, rice and the like. Notable promoters which are suitable are the barley lpt2 or lpt1 gene promoter (WO 95/15389 and WO 95/23230), or the promoters described in WO 99/16890 (promoters from the barley hordein gene, the rice glutelin gene, the rice oryzin gene, the rice prolamin gene, the wheat gliadin gene, the wheat glutelin gene, the maize zein gene, the oat glutelin gene, the sorghum kasirin gene, the rye secalin gene). Other promoters include those described by Broun et al. (1998), Potenza et al. (2004), US 20070192902 and US 20030159173. In an embodiment, the seed specific promoter is preferentially expressed in defined parts of the seed such as the cotyledon(s) or the endosperm. Examples of cotyledon specific promoters include, but are not limited to, the FP1 promoter (Ellerstrom et al., 1996), the pea legumin promoter (Perrin et al., 2000), and the bean phytohemagglutinin promoter (Perrin et al., 2000). Examples of endosperm specific promoters include, but are not limited to, the maize zein-1 promoter (Chikwamba et al., 2003), the rice glutelin-1 promoter (Yang et al., 2003), the barley D-hordein promoter (Horvath et al., 2000) and wheat HMW glutenin promoters (Alvarez et al., 2000). In a further embodiment, the seed specific promoter is not expressed, or is only expressed at a low level, in the embryo and/or after the seed germinates.

In another embodiment, the plant storage organ specific promoter is a tuber specific promoter. Examples include, but are not limited to, the potato patatin B33,

PAT21 and GBSS promoters, as well as the sweet potato sporamin promoter (for review, see Potenza et al., 2004). In a preferred embodiment, the promoter directs expression preferentially in the pith of the tuber, relative to the outer layers (skin, bark) or the embryo of the tuber.

- 5 In another embodiment, the plant storage organ specific promoter is a fruit specific promoter. Examples include, but are not limited to, the tomato polygalacturonase, E8 and Pds promoters, as well as the apple ACC oxidase promoter (for review, see Potenza et al., 2004). In a preferred embodiment, the promoter preferentially directs expression in the edible parts of the fruit, for example the pith of
10 the fruit, relative to the skin of the fruit or the seeds within the fruit.

- In an embodiment, the inducible promoter is the *Aspergillus nidulans* alc system. Examples of inducible expression systems which can be used instead of the *Aspergillus nidulans* alc system are described in a review by Padidam (2003) and Corrado and Karali (2009). These include tetracycline repressor (TetR)-based and
15 tetracycline inducible systems (Gatz, 1997), tetracycline repressor-based and tetracycline-inactivatable systems (Weinmann et al., 1994), glucocorticoid receptor-based (Picard, 1994), estrogen receptor-based and other steroid-inducible systems systems (Bruce et al., 2000), glucocorticoid receptor-, tetracycline repressor-based dual control systems (Bohner et al., 1999), ecdysone receptor-based, insecticide-
20 inducible systems (Martinez et al., 1999, Padidam et al., 2003, Unger et al, 2002, Riddiford et al., 2000, Dhadialla et al., 1998, Martinez and Jepson, 1999), AlcR-based, ethanol-inducible systems (Felenbok, 1991) and ACEI-based, copper-inducible systems (Mett et al., 1993).

- In another embodiment, the inducible promoter is a safener inducible promoter
25 such as, for example, the maize *ln2-1* or *ln2-2* promoter (Hershey and Stoner, 1991), the safener inducible promoter is the maize GST-27 promoter (Jepson et al., 1994), or the soybean GH2/4 promoter (Ulmasov et al., 1995).

- Safeners are a group of structurally diverse chemicals used to increase the plant's tolerance to the toxic effects of an herbicidal compound. Examples of
30 these compounds include naphthalic anhydride and N,N-diallyl-2,2-dichloroacetamide (DDCA), which protect maize and sorghum against thiocarbamate herbicides; cyometrinil, which protects sorghum against metochlor; triapenthenol, which protects soybeans against metribuzin; and substituted benzenesulfonamides, which improve the tolerance of several cereal crop species to sulfonylurea herbicides.

- 35 In another embodiment, the inducible promoter is a senescence inducible promoter such as, for example, senescence-inducible promoter SAG (senescence associated gene) 12 and SAG 13 from *Arabidopsis* (Gan, 1995; Gan and Amasino, 1995) and LSC54 from *Brassica napus* (Buchanan-Wollaston, 1994).

For expression in vegetative tissue leaf-specific promoters, such as the ribulose biphosphate carboxylase (RBCS) promoters, can be used. For example, the tomato RBCS1, RBCS2 and RBCS3A genes are expressed in leaves and light grown seedlings (Meier et al., 1997). A ribulose biphosphate carboxylase promoters
5 expressed almost exclusively in mesophyll cells in leaf blades and leaf sheaths at high levels, described by Matsuoka et al. (1994), can be used. Another leaf-specific promoter is the light harvesting chlorophyll a/b binding protein gene promoter (see, Shiina et al., 1997). The *Arabidopsis thaliana* myb-related gene promoter (Atmyb5) described by Li et al. (1996), is leaf-specific. The Atmyb5 promoter is expressed in
10 developing leaf trichomes, stipules, and epidermal cells on the margins of young rosette and cauline leaves, and in immature seeds. A leaf promoter identified in maize by Busk et al. (1997), can also be used.

In some instances, for example when LEC2 or BBM is recombinantly expressed, it may be desirable that the transgene is not expressed at high levels. An
15 example of a promoter which can be used in such circumstances is a truncated napin A promoter which retains the seed-specific expression pattern but with a reduced expression level (Tan et al., 2011).

The 5' non-translated leader sequence can be derived from the promoter selected to express the heterologous gene sequence of the polynucleotide of the
20 present invention, or may be heterologous with respect to the coding region of the enzyme to be produced, and can be specifically modified if desired so as to increase translation of mRNA. For a review of optimizing expression of transgenes, see Kozziel et al. (1996). The 5' non-translated regions can also be obtained from plant viral RNAs (Tobacco mosaic virus, Tobacco etch virus, Maize dwarf mosaic virus,
25 Alfalfa mosaic virus, among others) from suitable eukaryotic genes, plant genes (wheat and maize chlorophyll a/b binding protein gene leader), or from a synthetic gene sequence. The present invention is not limited to constructs wherein the non-translated region is derived from the 5' non-translated sequence that accompanies the promoter sequence. The leader sequence could also be derived from an unrelated
30 promoter or coding sequence. Leader sequences useful in context of the present invention comprise the maize Hsp70 leader (US 5,362,865 and US 5,859,347), and the TMV omega element.

The termination of transcription is accomplished by a 3' non-translated DNA sequence operably linked in the expression vector to the polynucleotide of interest.
35 The 3' non-translated region of a recombinant DNA molecule contains a polyadenylation signal that functions in plants to cause the addition of adenylate nucleotides to the 3' end of the RNA. The 3' non-translated region can be obtained from various genes that are expressed in plant cells. The nopaline synthase 3'

untranslated region, the 3' untranslated region from pea small subunit Rubisco gene, the 3' untranslated region from soybean 7S seed storage protein gene are commonly used in this capacity. The 3' transcribed, non-translated regions containing the polyadenylate signal of *Agrobacterium* tumor-inducing (Ti) plasmid genes are also
5 suitable.

Recombinant DNA technologies can be used to improve expression of a transformed polynucleotide by manipulating for example, the number of copies of the polynucleotide within a host cell, the efficiency with which those polynucleotide are transcribed, the efficiency with which the resultant transcripts are translated, and the
10 efficiency of post-translational modifications. Recombinant techniques useful for increasing the expression of polynucleotides defined herein include, but are not limited to, operatively linking the polynucleotide to a high-copy number plasmid, integration of the polynucleotide molecule into one or more host cell chromosomes, addition of vector stability sequences to the plasmid, substitutions or modifications of
15 transcription control signals (e.g., promoters, operators, enhancers), substitutions or modifications of translational control signals (e.g., ribosome binding sites, Shine-Dalgarno sequences), modification of the polynucleotide to correspond to the codon usage of the host cell, and the deletion of sequences that destabilize transcripts.

20 Transfer Nucleic Acids

Transfer nucleic acids can be used to deliver an exogenous polynucleotide to a cell and comprise one, preferably two, border sequences and a polynucleotide of interest. The transfer nucleic acid may or may not encode a selectable marker. Preferably, the transfer nucleic acid forms part of a binary vector in a bacterium,
25 where the binary vector further comprises elements which allow replication of the vector in the bacterium, selection, or maintenance of bacterial cells containing the binary vector. Upon transfer to a eukaryotic cell, the transfer nucleic acid component of the binary vector is capable of integration into the genome of the eukaryotic cell.

As used herein, the term "extrachromosomal transfer nucleic acid" refers to a
30 nucleic acid molecule that is capable of being transferred from a bacterium such as *Agrobacterium sp.*, to a eukaryotic cell such as a plant leaf cell. An extrachromosomal transfer nucleic acid is a genetic element that is well-known as an element capable of being transferred, with the subsequent integration of a nucleotide sequence contained within its borders into the genome of the recipient cell. In this
35 respect, a transfer nucleic acid is flanked, typically, by two "border" sequences, although in some instances a single border at one end can be used and the second end of the transferred nucleic acid is generated randomly in the transfer process. A polynucleotide of interest is typically positioned between the left border-like sequence

and the right border-like sequence of a transfer nucleic acid. The polynucleotide contained within the transfer nucleic acid may be operably linked to a variety of different promoter and terminator regulatory elements that facilitate its expression, that is, transcription and/or translation of the polynucleotide. Transfer DNAs (T-DNAs) from *Agrobacterium sp.* such as *Agrobacterium tumefaciens* or *Agrobacterium rhizogenes*, and man made variants/mutants thereof are probably the best characterized examples of transfer nucleic acids. Another example is P-DNA ("plant-DNA") which comprises T-DNA border-like sequences from plants.

As used herein, "T-DNA" refers to for example, T-DNA of an *Agrobacterium tumefaciens* Ti plasmid or from an *Agrobacterium rhizogenes* Ri plasmid, or man made variants thereof which function as T-DNA. The T-DNA may comprise an entire T-DNA including both right and left border sequences, but need only comprise the minimal sequences required in *cis* for transfer, that is, the right and T-DNA border sequence. The T-DNAs of the invention have inserted into them, anywhere between the right and left border sequences (if present), the polynucleotide of interest flanked by target sites for a site-specific recombinase. The sequences encoding factors required in *trans* for transfer of the T-DNA into a plant cell such as vir genes, may be inserted into the T-DNA, or may be present on the same replicon as the T-DNA, or preferably are in *trans* on a compatible replicon in the *Agrobacterium* host. Such "binary vector systems" are well known in the art.

As used herein, "P-DNA" refers to a transfer nucleic acid isolated from a plant genome, or man made variants/mutants thereof, and comprises at each end, or at only one end, a T-DNA border-like sequence. The border-like sequence preferably shares at least 50%, at least 60%, at least 70%, at least 75%, at least 80%, at least 90% or at least 95%, but less than 100% sequence identity, with a T-DNA border sequence from an *Agrobacterium sp.* such as *Agrobacterium tumefaciens* or *Agrobacterium rhizogenes*. Thus, P-DNAs can be used instead of T-DNAs to transfer a nucleotide sequence contained within the P-DNA from, for example *Agrobacterium*, to another cell. The P-DNA, before insertion of the exogenous polynucleotide which is to be transferred, may be modified to facilitate cloning and should preferably not encode any proteins. The P-DNA is characterized in that it contains, at least a right border sequence and preferably also a left border sequence.

As used herein, a "border" sequence of a transfer nucleic acid can be isolated from a selected organism such as a plant or bacterium, or be a man made variant/mutant thereof. The border sequence promotes and facilitates the transfer of the polynucleotide to which it is linked and may facilitate its integration in the recipient cell genome. In an embodiment, a border-sequence is between 5-100 base pairs (bp) in length, 10-80 bp in length, 15-75 bp in length, 15-60 bp in length, 15-50

bp in length, 15-40 bp in length, 15-30 bp in length, 16-30 bp in length, 20-30 bp in length, 21-30 bp in length, 22-30 bp in length, 23-30 bp in length, 24-30 bp in length, 25-30 bp in length, or 26-30 bp in length. Border sequences from T-DNA from *Agrobacterium sp.* are well known in the art and include those described in Lacroix et al. (2008), Tzfira and Citovsky (2006) and Glevin (2003).

Whilst traditionally only *Agrobacterium sp.* have been used to transfer genes to plants cells, there are now a large number of systems which have been identified/developed which act in a similar manner to *Agrobacterium sp.* Several non-*Agrobacterium* species have recently been genetically modified to be competent for gene transfer (Chung et al., 2006; Broothaerts et al., 2005). These include *Rhizobium sp.* NGR234, *Sinorhizobium meliloti* and *Mezorhizobium loti*. The bacteria are made competent for gene transfer by providing the bacteria with the machinery needed for the transformation process, that is, a set of virulence genes encoded by an *Agrobacterium* Ti-plasmid and the T-DNA segment residing on a separate, small binary plasmid. Bacteria engineered in this way are capable of transforming different plant tissues (leaf disks, calli and oval tissue), monocots or dicots, and various different plant species (e.g., tobacco, rice).

Direct transfer of eukaryotic expression plasmids from bacteria to eukaryotic hosts was first achieved several decades ago by the fusion of mammalian cells and protoplasts of plasmid-carrying *Escherichia coli* (Schaffner, 1980). Since then, the number of bacteria capable of delivering genes into mammalian cells has steadily increased (Weiss, 2003), being discovered by four groups independently (Sizemore et al. 1995; Courvalin et al., 1995; Powell et al., 1996; Darji et al., 1997).

Attenuated *Shigella flexneri*, *Salmonella typhimurium* or *E. coli* that had been rendered invasive by the virulence plasmid (pWR100) of *S. flexneri* have been shown to be able to transfer expression plasmids after invasion of host cells and intracellular death due to metabolic attenuation. Mucosal application, either nasally or orally, of such recombinant *Shigella* or *Salmonella* induced immune responses against the antigen that was encoded by the expression plasmids. In the meantime, the list of bacteria that was shown to be able to transfer expression plasmids to mammalian host cells *in vitro* and *in vivo* has been more than doubled and has been documented for *S. typhi*, *S. choleraesuis*, *Listeria monocytogenes*, *Yersinia pseudotuberculosis*, and *Y. enterocolitica* (Fennelly et al., 1999; Shiau et al., 2001; Dietrich et al., 1998; Hense et al., 2001; Al-Mariri et al., 2002).

In general, it could be assumed that all bacteria that are able to enter the cytosol of the host cell (like *S. flexneri* or *L. monocytogenes*) and lyse within this cellular compartment, should be able to transfer DNA. This is known as 'abortive' or 'suicidal' invasion as the bacteria have to lyse for the DNA transfer to occur (Grillot-

Courvalin et al., 1999). In addition, even many of the bacteria that remain in the phagocytic vacuole (like *S. typhimurium*) may also be able to do so. Thus, recombinant laboratory strains of *E. coli* that have been engineered to be invasive but are unable of phagosomal escape, could deliver their plasmid load to the nucleus of
5 the infected mammalian cell nevertheless (Grillot-Courvalin et al., 1998). Furthermore, *Agrobacterium tumefaciens* has recently also been shown to introduce transgenes into mammalian cells (Kunik et al., 2001).

As used herein, the terms "transfection", "transformation" and variations thereof are generally used interchangeably. "Transfected" or "transformed" cells may
10 have been manipulated to introduce the polynucleotide(s) of interest, or may be progeny cells derived therefrom.

Recombinant Cells

The invention also provides a recombinant cell, for example, a recombinant
15 plant cell, which is a host cell transformed with one or more polynucleotides or vectors defined herein, or combination thereof. The term "recombinant cell" is used interchangeably with the term "transgenic cell" herein. Suitable cells of the invention include any cell that can be transformed with a polynucleotide or recombinant vector of the invention, encoding for example, a polypeptide or enzyme described herein.
20 The cell is preferably a cell which is thereby capable of being used for producing lipid. The recombinant cell may be a cell in culture, a cell *in vitro*, or in an organism such as for example, a plant, or in an organ such as, for example, a seed or a leaf. Preferably, the cell is in a plant, more preferably in the seed of a plant. In one embodiment, the recombinant cell is a non-human cell.

25 Host cells into which the polynucleotide(s) are introduced can be either untransformed cells or cells that are already transformed with at least one nucleic acid. Such nucleic acids may be related to lipid synthesis, or unrelated. Host cells of the present invention either can be endogenously (i.e., naturally) capable of producing polypeptide(s) defined herein, in which case the recombinant cell derived therefrom
30 has an enhanced capability of producing the polypeptide(s), or can be capable of producing said polypeptide(s) only after being transformed with at least one polynucleotide of the invention. In an embodiment, a recombinant cell of the invention has an enhanced capacity to produce non-polar lipid.

Host cells of the present invention can be any cell capable of producing at least
35 one protein described herein, and include bacterial, fungal (including yeast), parasite, arthropod, animal, algal, and plant cells. The cells may be prokaryotic or eukaryotic. Preferred host cells are yeast, algal and plant cells. In a preferred embodiment, the plant cell is a seed cell, in particular, a cell in a cotyledon or endosperm of a seed. In

one embodiment, the cell is an animal cell. The animal cell may be of any type of animal such as, for example, a non-human animal cell, a non-human vertebrate cell, a non-human mammalian cell, or cells of aquatic animals such as fish or crustacea, invertebrates, insects, etc. Non limiting examples of arthropod cells include insect
5 cells such as *Spodoptera frugiperda* (Sf) cells, for example, Sf9, Sf21, *Trichoplusia ni* cells, and *Drosophila* S2 cells. An example of a bacterial cell useful as a host cell of the present invention is *Synechococcus* spp. (also known as *Synechocystis* spp.), for example *Synechococcus elongatus*. Examples of algal cells useful as host cells of the present invention include, for example, *Chlamydomonas* sp. (for example,
10 *Chlamydomonas reinhardtii*), *Dunaliella* sp., *Haematococcus* sp., *Chlorella* sp., *Thraustochytrium* sp., *Schizochytrium* sp., and *Volvox* sp.

Host cells for expression of the instant nucleic acids may include microbial hosts that grow on a variety of feedstocks, including simple or complex carbohydrates, organic acids and alcohols and/or hydrocarbons over a wide range of
15 temperature and pH values. Preferred microbial hosts are oleaginous organisms that are naturally capable of non-polar lipid synthesis.

The host cells may be of an organism suitable for a fermentation process, such as, for example, *Yarrowia lipolytica* or other yeasts.

20 Transgenic Plants

The invention also provides a plant comprising an exogenous polynucleotide or polypeptide of the invention, a cell of the invention, a vector of the invention, or a combination thereof. The term "plant" refers to whole plants, whilst the term "part thereof" refers to plant organs (e.g., leaves, stems, roots, flowers, fruit), single cells
25 (e.g., pollen), seed, seed parts such as an embryo, endosperm, scutellum or seed coat, plant tissue such as vascular tissue, plant cells and progeny of the same. As used herein, plant parts comprise plant cells.

As used herein, the term "plant" is used in its broadest sense. It includes, but is not limited to, any species of grass, ornamental or decorative plant, crop or cereal
30 (e.g., oilseed, maize, soybean), fodder or forage, fruit or vegetable plant, herb plant, woody plant, flower plant, or tree. It is not meant to limit a plant to any particular structure. It also refers to a unicellular plant (e.g., microalga). The term "part thereof" in reference to a plant refers to a plant cell and progeny of same, a plurality of plant cells that are largely differentiated into a colony (e.g., volvox), a structure that
35 is present at any stage of a plant's development, or a plant tissue. Such structures include, but are not limited to, leaves, stems, flowers, fruits, nuts, roots, seed, seed coat, embryos. The term "plant tissue" includes differentiated and undifferentiated tissues of plants including those present in leaves, stems, flowers, fruits, nuts, roots,

seed, for example, embryonic tissue, endosperm, dermal tissue (e.g., epidermis, periderm), vascular tissue (e.g., xylem, phloem), or ground tissue (comprising parenchyma, collenchyma, and/or sclerenchyma cells), as well as cells in culture (e.g., single cells, protoplasts, callus, embryos, etc.). Plant tissue may be *in planta*, in organ
5 culture, tissue culture, or cell culture.

A "transgenic plant", "genetically modified plant" or variations thereof refers to a plant that contains a transgene not found in a wild-type plant of the same species, variety or cultivar. Transgenic plants as defined in the context of the present invention include plants and their progeny which have been genetically modified
10 using recombinant techniques to cause production of at least one polypeptide defined herein in the desired plant or part thereof. Transgenic plant parts has a corresponding meaning.

The terms "seed" and "grain" are used interchangeably herein. "Grain" refers to mature grain such as harvested grain or grain which is still on a plant but ready for
15 harvesting, but can also refer to grain after imbibition or germination, according to the context. Mature grain commonly has a moisture content of less than about 18-20%. In a preferred embodiment, the moisture content of the grain is at a level which is generally regarded as safe for storage, preferably between 5% and 15%, between 6% and 8%, between 8% and 10%, or between 12% and 15%. "Developing seed" as used
20 herein refers to a seed prior to maturity, typically found in the reproductive structures of the plant after fertilisation or anthesis, but can also refer to such seeds prior to maturity which are isolated from a plant. Mature seed commonly has a moisture content of less than about 18-20%. In a preferred embodiment, the moisture content of the seed is at a level which is generally regarded as safe for storage, preferably
25 between 5% and 15%, between 6% and 8%, between 8% and 10%, or between 12% and 15%.

As used herein, the term "plant storage organ" refers to a part of a plant specialized to store energy in the form of for example, proteins, carbohydrates, lipid. Examples of plant storage organs are seed, fruit, tuberous roots, and tubers. A
30 preferred plant storage organ of the invention is seed.

As used herein, the term "phenotypically normal" refers to a genetically modified plant or part thereof, particularly a storage organ such as a seed, tuber or fruit of the invention not having a significantly reduced ability to grow and reproduce when compared to an unmodified plant or plant thereof. In an embodiment, the
35 genetically modified plant or part thereof which is phenotypically normal comprises a recombinant polynucleotide encoding a silencing suppressor operably linked to a plant storage organ specific promoter and has an ability to grow or reproduce which is essentially the same as a corresponding plant or part thereof not comprising said

polynucleotide. Preferably, the biomass, growth rate, germination rate, storage organ size, seed size and/or the number of viable seeds produced is not less than 90% of that of a plant lacking said recombinant polynucleotide when grown under identical conditions. This term does not encompass features of the plant which may be different to the wild-type plant but which do not effect the usefulness of the plant for commercial purposes such as, for example, a ballerina phenotype of seedling leaves.

Plants provided by or contemplated for use in the practice of the present invention include both monocotyledons and dicotyledons. In preferred embodiments, the plants of the present invention are crop plants (for example, cereals and pulses, maize, wheat, potatoes, tapioca, rice, sorghum, millet, cassava, barley, or pea), or other legumes. The plants may be grown for production of edible roots, tubers, leaves, stems, flowers or fruit. The plants may be vegetable or ornamental plants. The plants of the invention may be: *Acrocomia aculeata* (macauba palm), *Arabidopsis thaliana*, *Aracinis hypogaea* (peanut), *Astrocaryum murumuru* (murumuru), *Astrocaryum vulgare* (tucumã), *Attalea geraensis* (Indaiá-rateiro), *Attalea humilis* (American oil palm), *Attalea oleifera* (andaiá), *Attalea phalerata* (uricuri), *Attalea speciosa* (babassu), *Avena sativa* (oats), *Beta vulgaris* (sugar beet), *Brassica sp.* such as *Brassica carinata*, *Brassica juncea*, *Brassica napobrassica*, *Brassica napus* (canola), *Camelina sativa* (false flax), *Cannabis sativa* (hemp), *Carthamus tinctorius* (safflower), *Caryocar brasiliense* (pequi), *Cocos nucifera* (Coconut), *Crambe abyssinica* (Abyssinian kale), *Cucumis melo* (melon), *Elaeis guineensis* (African palm), *Glycine max* (soybean), *Gossypium hirsutum* (cotton), *Helianthus sp.* such as *Helianthus annuus* (sunflower), *Hordeum vulgare* (barley), *Jatropha curcas* (physic nut), *Joannesia princeps* (arara nut-tree), *Lemna sp.* (duckweed) such as *Lemna aequinoctialis*, *Lemna disperma*, *Lemna ecuadoriensis*, *Lemna gibba* (swollen duckweed), *Lemna japonica*, *Lemna minor*, *Lemna minuta*, *Lemna obscura*, *Lemna paucicostata*, *Lemna perpusilla*, *Lemna tenera*, *Lemna trisulca*, *Lemna turionifera*, *Lemna valdiviana*, *Lemna yungensis*, *Licania rigida* (oiticica), *Linum usitatissimum* (flax), *Lupinus angustifolius* (lupin), *Mauritia flexuosa* (buriti palm), *Maximiliana maripa* (inaja palm), *Miscanthus sp.* such as *Miscanthus x giganteus* and *Miscanthus sinensis*, *Nicotiana sp.* (tobacco) such as *Nicotiana tabacum* or *Nicotiana benthamiana*, *Oenocarpus bacaba* (bacaba-do-azeite), *Oenocarpus bataua* (patauã), *Oenocarpus distichus* (bacaba-de-leque), *Oryza sp.* (rice) such as *Oryza sativa* and *Oryza glaberrima*, *Panicum virgatum* (switchgrass), *Paraqueiba paraensis* (mari), *Persea amencana* (avocado), *Pongamia pinnata* (Indian beech), *Populus trichocarpa*, *Ricinus communis* (castor), *Saccharum sp.* (sugarcane), *Sesamum indicum* (sesame), *Solanum tuberosum* (potato), *Sorghum sp.* such as *Sorghum bicolor*, *Sorghum vulgare*, *Theobroma grandiflorum* (cupuassu), *Trifolium sp.*, *Trithrinax brasiliensis*

(Brazilian needle palm), *Triticum* sp. (wheat) such as *Triticum aestivum*, *Zea mays* (corn), alfalfa (*Medicago sativa*), rye (*Secale cereale*), sweet potato (*Lopmoea batatus*), cassava (*Manihot esculenta*), coffee (*Cofea* spp.), pineapple (*Anana comosus*), citris tree (*Citrus* spp.), cocoa (*Theobroma cacao*), tea (*Camellia senensis*),
 5 banana (*Musa* spp.), avocado (*Persea americana*), fig (*Ficus casica*), guava (*Psidium guajava*), mango (*Mangifer indica*), olive (*Olea europaea*), papaya (*Carica papaya*), cashew (*Anacardium occidentale*), macadamia (*Macadamia intergrifolia*) and almond (*Prunus amygdalus*).

Other preferred plants include C4 grasses such as, in addition to those
 10 mentioned above, *Andropogon gerardi*, *Boiteloua curtispindula*, *B. gracilis*, *Buchloe dactyloides*, *Schizachyrium scoparium*, *Sorghastrum nutans*, *Sporobolus cryptandrus*; C3 grasses such as *Elymus canadensis*, the legumes *Lespedeza capitata* and *Petalostemum villosum*, the forb *Aster azureus*; and woody plants such as *Quercus ellipsoidalis* and *Q. macrocarpa*. Other preferred plants include C3 grasses.

15 In a preferred embodiment, the plant is an angiosperm.

In an embodiment, the plant is an oilseed plant, preferably an oilseed crop plant. As used herein, an "oilseed plant" is a plant species used for the commercial production of lipid from the seeds of the plant. The oilseed plant may be, for example, oil-seed rape (such as canola), maize, sunflower, safflower, soybean,
 20 sorghum, flax (linseed) or sugar beet. Furthermore, the oilseed plant may be other *Brassicas*, cotton, peanut, poppy, rutabaga, mustard, castor bean, sesame, safflower, or nut producing plants. The plant may produce high levels of lipid in its fruit such as olive, oil palm or coconut. Horticultural plants to which the present invention may be applied are lettuce, endive, or vegetable *Brassicas* including cabbage, broccoli, or
 25 cauliflower. The present invention may be applied in tobacco, cucurbits, carrot, strawberry, tomato, or pepper.

In a preferred embodiment, the transgenic plant is homozygous for each and every gene that has been introduced (transgene) so that its progeny do not segregate for the desired phenotype. The transgenic plant may also be heterozygous for the
 30 introduced transgene(s), preferably uniformly heterozygous for the transgene such as for example, in F1 progeny which have been grown from hybrid seed. Such plants may provide advantages such as hybrid vigour, well known in the art.

Where relevant, the transgenic plants may also comprise additional transgenes encoding enzymes involved in the production of non-polar lipid such as, but not
 35 limited to LPAAT, LPCAT, PAP, or a phospholipid:diacylglycerol acyltransferase (PDAT1, PDAT2 or PDAT3; see for example, Ghosal et al., 2007), or a combination of two or more thereof. The transgenic plants of the invention may also express oleosin from an exogenous polynucleotide.

Transformation of plants

Transgenic plants can be produced using techniques known in the art, such as those generally described in Slater et al., *Plant Biotechnology - The Genetic Manipulation of Plants*, Oxford University Press (2003), and Christou and Klee, *Handbook of Plant Biotechnology*, John Wiley and Sons (2004).

As used herein, the terms "stably transforming", "stably transformed" and variations thereof refer to the integration of the polynucleotide into the genome of the cell such that they are transferred to progeny cells during cell division without the need for positively selecting for their presence. Stable transformants, or progeny thereof, can be selected by any means known in the art such as Southern blots on chromosomal DNA, or *in situ* hybridization of genomic DNA.

Agrobacterium-mediated transfer is a widely applicable system for introducing genes into plant cells because DNA can be introduced into cells in whole plant tissues, plant organs, or explants in tissue culture, for either transient expression, or for stable integration of the DNA in the plant cell genome. The use of *Agrobacterium*-mediated plant integrating vectors to introduce DNA into plant cells is well known in the art (see for example, US 5177010, US 5104310, US 5004863, or US 5159135). The region of DNA to be transferred is defined by the border sequences, and the intervening DNA (T-DNA) is usually inserted into the plant genome. Further, the integration of the T-DNA is a relatively precise process resulting in few rearrangements. In those plant varieties where *Agrobacterium*-mediated transformation is efficient, it is the method of choice because of the facile and defined nature of the gene transfer. Preferred *Agrobacterium* transformation vectors are capable of replication in *E. coli* as well as *Agrobacterium*, allowing for convenient manipulations as described (Klee et al., In: *Plant DNA Infectious Agents*, Hohn and Schell, eds., Springer-Verlag, New York, pp. 179-203 (1985)).

Acceleration methods that may be used include for example, microprojectile bombardment and the like. One example of a method for delivering transforming nucleic acid molecules to plant cells is microprojectile bombardment. This method has been reviewed by Yang et al., *Particle Bombardment Technology for Gene Transfer*, Oxford Press, Oxford, England (1994). Non-biological particles (microprojectiles) that may be coated with nucleic acids and delivered into cells by a propelling force. Exemplary particles include those comprised of tungsten, gold, platinum, and the like. A particular advantage of microprojectile bombardment, in addition to it being an effective means of reproducibly transforming monocots, is that neither the isolation of protoplasts, nor the susceptibility of *Agrobacterium* infection are required. An illustrative embodiment of a method for delivering DNA into *Zea mays* cells by acceleration is a biolistics α -particle delivery system, that can be used

to propel particles coated with DNA through a screen such as a stainless steel or Nytex screen, onto a filter surface covered with corn cells cultured in suspension. A particle delivery system suitable for use with the present invention is the helium acceleration PDS-1000/He gun available from Bio-Rad Laboratories.

5 For the bombardment, cells in suspension may be concentrated on filters. Filters containing the cells to be bombarded are positioned at an appropriate distance below the microprojectile stopping plate. If desired, one or more screens are also positioned between the gun and the cells to be bombarded.

Alternatively, immature embryos or other target cells may be arranged on solid
10 culture medium. The cells to be bombarded are positioned at an appropriate distance below the microprojectile stopping plate. If desired, one or more screens are also positioned between the acceleration device and the cells to be bombarded. Through the use of techniques set forth herein, one may obtain up to 1000 or more foci of cells transiently expressing a marker gene. The number of cells in a focus that express the
15 gene product 48 hours post-bombardment often range from one to ten and average one to three.

In bombardment transformation, one may optimize the pre-bombardment culturing conditions and the bombardment parameters to yield the maximum numbers of stable transformants. Both the physical and biological parameters for
20 bombardment are important in this technology. Physical factors are those that involve manipulating the DNA/microprojectile precipitate or those that affect the flight and velocity of either the macro- or microprojectiles. Biological factors include all steps involved in manipulation of cells before and immediately after bombardment, the osmotic adjustment of target cells to help alleviate the trauma associated with
25 bombardment, and also the nature of the transforming DNA such as linearized DNA or intact supercoiled plasmids. It is believed that pre-bombardment manipulations are especially important for successful transformation of immature embryos.

In another alternative embodiment, plastids can be stably transformed. Methods disclosed for plastid transformation in higher plants include particle gun
30 delivery of DNA containing a selectable marker and targeting of the DNA to the plastid genome through homologous recombination (US 5,451,513, US 5,545,818, US 5,877,402, US 5,932,479, and WO 99/05265).

Accordingly, it is contemplated that one may wish to adjust various aspects of the bombardment parameters in small scale studies to fully optimize the conditions.
35 One may particularly wish to adjust physical parameters such as gap distance, flight distance, tissue distance, and helium pressure. One may also minimize the trauma reduction factors by modifying conditions that influence the physiological state of the recipient cells and that may therefore influence transformation and integration

efficiencies. For example, the osmotic state, tissue hydration and the subculture stage, or cell cycle of the recipient cells, may be adjusted for optimum transformation. The execution of other routine adjustments will be known to those of skill in the art in light of the present disclosure.

5 Transformation of plant protoplasts can be achieved using methods based on calcium phosphate precipitation, polyethylene glycol treatment, electroporation, and combinations of these treatments. Application of these systems to different plant varieties depends upon the ability to regenerate that particular plant strain from protoplasts. Illustrative methods for the regeneration of cereals from protoplasts are
10 described (Fujimura et al., 1985; Toriyama et al., 1986; Abdullah et al., 1986).

Other methods of cell transformation can also be used and include but are not limited to the introduction of DNA into plants by direct DNA transfer into pollen, by direct injection of DNA into reproductive organs of a plant, or by direct injection of DNA into the cells of immature embryos followed by the rehydration of desiccated
15 embryos.

The regeneration, development, and cultivation of plants from single plant protoplast transformants or from various transformed explants is well known in the art (Weissbach et al., In: Methods for Plant Molecular Biology, Academic Press, San Diego, Calif., (1988)). This regeneration and growth process typically includes the
20 steps of selection of transformed cells, culturing those individualized cells through the usual stages of embryonic development through the rooted plantlet stage. Transgenic embryos and seeds are similarly regenerated. The resulting transgenic rooted shoots are thereafter planted in an appropriate plant growth medium such as soil.

The development or regeneration of plants containing the foreign, exogenous
25 gene is well known in the art. Preferably, the regenerated plants are self-pollinated to provide homozygous transgenic plants. Otherwise, pollen obtained from the regenerated plants is crossed to seed-grown plants of agronomically important lines. Conversely, pollen from plants of these important lines is used to pollinate regenerated plants. A transgenic plant of the present invention containing a desired
30 polynucleotide is cultivated using methods well known to one skilled in the art.

Methods for transforming dicots, primarily by use of *Agrobacterium tumefaciens*, and obtaining transgenic plants have been published for cotton (US 5,004,863, US 5,159,135, US 5,518,908), soybean (US 5,569,834, US 5,416,011), Brassica (US 5,463,174), peanut (Cheng et al., 1996), and pea (Grant et al., 1995).

35 Methods for transformation of cereal plants such as wheat and barley for introducing genetic variation into the plant by introduction of an exogenous nucleic acid and for regeneration of plants from protoplasts or immature plant embryos are well known in the art, see for example, CA 2,092,588, AU 61781/94, AU 667939, US

6,100,447, WO 97/048814, US 5,589,617, US 6,541,257, and other methods are set out in WO 99/14314. Preferably, transgenic wheat or barley plants are produced by *Agrobacterium tumefaciens* mediated transformation procedures. Vectors carrying the desired polynucleotide may be introduced into regenerable wheat cells of tissue
5 cultured plants or explants, or suitable plant systems such as protoplasts.

The regenerable wheat cells are preferably from the scutellum of immature embryos, mature embryos, callus derived from these, or the meristematic tissue.

To confirm the presence of the transgenes in transgenic cells and plants, a polymerase chain reaction (PCR) amplification or Southern blot analysis can be
10 performed using methods known to those skilled in the art. Expression products of the transgenes can be detected in any of a variety of ways, depending upon the nature of the product, and include Western blot and enzyme assay. One particularly useful way to quantitate protein expression and to detect replication in different plant tissues is to use a reporter gene such as GUS. Once transgenic plants have been obtained,
15 they may be grown to produce plant tissues or parts having the desired phenotype. The plant tissue or plant parts, may be harvested, and/or the seed collected. The seed may serve as a source for growing additional plants with tissues or parts having the desired characteristics. Preferably, the vegetative plant parts are harvested at a time when the yield of non-polar lipids are at their highest. In one embodiment, the
20 vegetative plant parts are harvested about at the time of flowering.

A transgenic plant formed using *Agrobacterium* or other transformation methods typically contains a single genetic locus on one chromosome. Such transgenic plants can be referred to as being hemizygous for the added gene(s). More preferred is a transgenic plant that is homozygous for the added gene(s), that is, a
25 transgenic plant that contains two added genes, one gene at the same locus on each chromosome of a chromosome pair. A homozygous transgenic plant can be obtained by self-fertilising a hemizygous transgenic plant, germinating some of the seed produced and analyzing the resulting plants for the gene of interest.

It is also to be understood that two different transgenic plants that contain two
30 independently segregating exogenous genes or loci can also be crossed (mated) to produce offspring that contain both sets of genes or loci. Selfing of appropriate F1 progeny can produce plants that are homozygous for both exogenous genes or loci. Back-crossing to a parental plant and out-crossing with a non-transgenic plant are also contemplated, as is vegetative propagation. Descriptions of other breeding methods
35 that are commonly used for different traits and crops can be found in Fehr, In: Breeding Methods for Cultivar Development, Wilcox J. ed., American Society of Agronomy, Madison Wis. (1987).

TILLING

In one embodiment, TILLING (Targeting Induced Local Lesions IN Genomes) can be used to produce plants in which endogenous genes are knocked out, for example genes encoding a DGAT, *sn-1* glycerol-3-phosphate acyltransferase (GPAT),
 5 1-acyl-glycerol-3-phosphate acyltransferase (LPAAT), acyl-CoA:lysophosphatidylcholine acyltransferase (LPCAT), phosphatidic acid phosphatase (PAP), or a combination of two or more thereof.

In a first step, introduced mutations such as novel single base pair changes are induced in a population of plants by treating seeds (or pollen) with a chemical
 10 mutagen, and then advancing plants to a generation where mutations will be stably inherited. DNA is extracted, and seeds are stored from all members of the population to create a resource that can be accessed repeatedly over time.

For a TILLING assay, PCR primers are designed to specifically amplify a single gene target of interest. Specificity is especially important if a target is a
 15 member of a gene family or part of a polyploid genome. Next, dye-labeled primers can be used to amplify PCR products from pooled DNA of multiple individuals. These PCR products are denatured and reannealed to allow the formation of mismatched base pairs. Mismatches, or heteroduplexes, represent both naturally occurring single nucleotide polymorphisms (SNPs) (i.e., several plants from the
 20 population are likely to carry the same polymorphism) and induced SNPs (i.e., only rare individual plants are likely to display the mutation). After heteroduplex formation, the use of an endonuclease, such as *CelI*, that recognizes and cleaves mismatched DNA is the key to discovering novel SNPs within a TILLING population.

25 Using this approach, many thousands of plants can be screened to identify any individual with a single base change as well as small insertions or deletions (1-30 bp) in any gene or specific region of the genome. Genomic fragments being assayed can range in size anywhere from 0.3 to 1.6 kb. At 8-fold pooling, 1.4 kb fragments (discounting the ends of fragments where SNP detection is problematic due to noise)
 30 and 96 lanes per assay, this combination allows up to a million base pairs of genomic DNA to be screened per single assay, making TILLING a high-throughput technique. TILLING is further described in Slade and Knauf (2005), and Henikoff et al. (2004).

In addition to allowing efficient detection of mutations, high-throughput TILLING technology is ideal for the detection of natural polymorphisms. Therefore,
 35 interrogating an unknown homologous DNA by heteroduplexing to a known sequence reveals the number and position of polymorphic sites. Both nucleotide changes and small insertions and deletions are identified, including at least some repeat number polymorphisms. This has been called Ecotilling (Comai et al., 2004).

Each SNP is recorded by its approximate position within a few nucleotides. Thus, each haplotype can be archived based on its mobility. Sequence data can be obtained with a relatively small incremental effort using aliquots of the same amplified DNA that is used for the mismatch-cleavage assay. The left or right
5 sequencing primer for a single reaction is chosen by its proximity to the polymorphism. Sequencher software performs a multiple alignment and discovers the base change, which in each case confirmed the gel band.

Ecotilling can be performed more cheaply than full sequencing, the method currently used for most SNP discovery. Plates containing arrayed ecotypic DNA can
10 be screened rather than pools of DNA from mutagenized plants. Because detection is on gels with nearly base pair resolution and background patterns are uniform across lanes, bands that are of identical size can be matched, thus discovering and genotyping SNPs in a single step. In this way, ultimate sequencing of the SNP is simple and efficient, made more so by the fact that the aliquots of the same PCR
15 products used for screening can be subjected to DNA sequencing.

Enhancing Exogenous RNA Levels and Stabilized Expression

Post-transcriptional gene silencing (PTGS) is a nucleotide sequence-specific defense mechanism that can target both cellular and viral mRNAs for degradation.
20 PTGS occurs in plants or fungi stably or transiently transformed with a recombinant polynucleotide(s) and results in the reduced accumulation of RNA molecules with sequence similarity to the introduced polynucleotide. "Post-transcriptional" refers to a mechanism operating at least partly, but not necessarily exclusively, after production of an initial RNA transcript, for example during processing of the initial RNA
25 transcript, or concomitant with splicing or export of the RNA to the cytoplasm, or within the cytoplasm by complexes associated with Argonaute proteins.

RNA molecule levels can be increased, and/or RNA molecule levels stabilized over numerous generations or under different environmental conditions, by limiting the expression of a silencing suppressor in a storage organ of a plant or part thereof.
30 As used herein, a "silencing suppressor" is any polynucleotide or polypeptide that can be expressed in a plant cell that enhances the level of expression product from a different transgene in the plant cell, particularly, over repeated generations from the initially transformed plant. In an embodiment, the silencing suppressor is a viral silencing suppressor or mutant thereof. A large number of viral silencing suppressors
35 are known in the art and include, but are not limited to P19, V2, P38, Pe-Po and RPV-P0. Examples of suitable viral silencing suppressors include those described in WO 2010/057246. A silencing suppressor may be stably expressed in a plant or part thereof of the present invention.

As used herein, the term "stably expressed" or variations thereof refers to the level of the RNA molecule being essentially the same or higher in progeny plants over repeated generations, for example, at least three, at least five, or at least ten generations, when compared to corresponding plants lacking the exogenous polynucleotide encoding the silencing suppressor. However, this term(s) does not exclude the possibility that over repeated generations there is some loss of levels of the RNA molecule when compared to a previous generation, for example, not less than a 10% loss per generation.

The suppressor can be selected from any source e.g. plant, viral, mammal, etc.

10 The suppressor may be, for example, flock house virus B2, pothos latent virus P14, pothos latent virus AC2, African cassava mosaic virus AC4, bhendi yellow vein mosaic disease C2, bhendi yellow vein mosaic disease C4, bhendi yellow vein mosaic disease β C1, tomato chlorosis virus p22, tomato chlorosis virus CP, tomato chlorosis virus CPm, tomato golden mosaic virus AL2, tomato leaf curl Java virus β C1, tomato

15 yellow leaf curl virus V2, tomato yellow leaf curl virus-China C2, tomato yellow leaf curl China virus Y10 isolate β C1, tomato yellow leaf curl Israeli isolate V2, mungbean yellow mosaic virus-Vigna AC2, hibiscus chlorotic ringspot virus CP, turnip crinkle virus P38, turnip crinkle virus CP, cauliflower mosaic virus P6, beet yellows virus p21, citrus tristeza virus p20, citrus tristeza virus p23, citrus tristeza

20 virus CP, cowpea mosaic virus SCP, sweet potato chlorotic stunt virus p22, cucumber mosaic virus 2b, tomato aspermy virus HC-Pro, beet curly top virus L2, soil borne wheat mosaic virus 19K, barley stripe mosaic virus Gammab, poa semilatifolia virus Gammab, peanut clump pecluvirus P15, rice dwarf virus Pns10, cucurbit aphid borne yellows virus P0, beet western yellows virus P0, potato virus X P25, cucumber vein

25 yellowing virus P1b, plum pox virus HC-Pro, sugarcane mosaic virus HC-Pro, potato virus Y strain HC-Pro, tobacco etch virus P1/HC-Pro, turnip mosaic virus P1/HC-Pro, cocksfoot mottle virus P1, cocksfoot mottle virus-Norwegian isolate P1, rice yellow mottle virus P1, rice yellow mottle virus-Nigerian isolate P1, rice hoja blanca virus NS3, rice stripe virus NS3, crucifer infecting tobacco mosaic virus 126K,

30 crucifer infecting tobacco mosaic virus p122, tobacco mosaic virus p122, tobacco mosaic virus 126, tobacco mosaic virus 130K, tobacco rattle virus 16K, tomato bushy stunt virus P19, tomato spotted wilt virus NSs, apple chlorotic leaf spot virus P50, grapevine virus A p10, grapevine leafroll associated virus-2 homolog of BYV p21, as well as variants/mutants thereof. The list above provides the virus from which the

35 suppressor can be obtained and the protein (e.g., B2, P14, etc.), or coding region designation for the suppressor from each particular virus. Other candidate silencing suppressors may be obtained by examining viral genome sequences for polypeptides encoded at the same position within the viral genome, relative to the structure of a

related viral genome comprising a known silencing suppressor, as is appreciated by a person of skill in the art.

Silencing suppressors can be categorized based on their mode of action. Suppressors such as V2 which preferentially bind to a double-stranded RNA molecule which has overhanging 5' ends relative to a corresponding double-stranded RNA molecule having blunt ends are particularly useful for enhancing transgene expression when used in combination with gene silencing (exogenous polynucleotide encoding a dsRNA). Other suppressors such as p19 which preferentially bind a dsRNA molecule which is 21 base pairs in length relative to a dsRNA molecule of a different length can also allow transgene expression in the presence of an exogenous polynucleotide encoding a dsRNA, but generally to a lesser degree than, for example, V2. This allows the selection of an optimal combination of dsRNA, silencing suppressor and over-expressed transgene for a particular purpose. Such optimal combinations can be identified using a method of the invention.

In an embodiment, the silencing suppressor preferentially binds to a double-stranded RNA molecule which has overhanging 5' ends relative to a corresponding double-stranded RNA molecule having blunt ends. In this context, the corresponding double-stranded RNA molecule preferably has the same nucleotide sequence as the molecule with the 5' overhanging ends, but without the overhanging 5' ends. Binding assays are routinely performed, for example in *in vitro* assays, by any method as known to a person of skill in the art.

Multiple copies of a suppressor may be used. Different suppressors may be used together (e. g., in tandem).

Essentially any RNA molecule which is desirable to be expressed in a plant storage organ can be co-expressed with the silencing suppressor. The RNA molecule may influence an agronomic trait, insect resistance, disease resistance, herbicide resistance, sterility, grain characteristics, and the like. The encoded polypeptides may be involved in metabolism of lipid, starch, carbohydrates, nutrients, etc., or may be responsible for the synthesis of proteins, peptides, lipids, waxes, starches, sugars, carbohydrates, flavors, odors, toxins, carotenoids, hormones, polymers, flavonoids, storage proteins, phenolic acids, alkaloids, lignins, tannins, celluloses, glycoproteins, glycolipids, etc.

In a particular example, the plants produced increased levels of enzymes for lipid production in plants such as Brassicas, for example oilseed rape or sunflower, safflower, flax, cotton, soya bean or maize.

Plant Biomass

An increase in the total lipid content of plant biomass equates to greater energy content, making its use in the production of biofuel more economical.

Plant biomass is the organic materials produced by plants, such as leaves, roots, seeds, and stalks. Plant biomass is a complex mixture of organic materials, such as carbohydrates, fats, and proteins, along with small amounts of minerals, such as sodium, phosphorus, calcium, and iron. The main components of plant biomass are carbohydrates (approximately 75%, dry weight) and lignin (approximately 25%), which can vary with plant type. The carbohydrates are mainly cellulose or hemicellulose fibers, which impart strength to the plant structure, and lignin, which holds the fibers together. Some plants also store starch (another carbohydrate polymer) and fats as sources of energy, mainly in seeds and roots (such as corn, soybeans, and potatoes).

Plant biomass typically has a low energy density as a result of both its physical form and moisture content. This makes it inconvenient and inefficient for storage and transport, and also usually unsuitable for use without some kind of pre-processing.

There are a range of processes available to convert it into a more convenient form including: 1) physical pre-processing (for example, grinding) or 2) conversion by thermal (for example, combustion, gasification, pyrolysis) or chemical (for example, anaerobic digestion, fermentation, composting, transesterification) processes. In this way, the biomass is converted into what can be described as a biomass fuel.

Combustion

Combustion is the process by which flammable materials are allowed to burn in the presence of air or oxygen with the release of heat. The basic process is oxidation. Combustion is the simplest method by which biomass can be used for energy, and has been used to provide heat. This heat can itself be used in a number of ways: 1) space heating, 2) water (or other fluid) heating for central or district heating or process heat, 3) steam raising for electricity generation or motive force. When the flammable fuel material is a form of biomass the oxidation is of predominantly the carbon (C) and hydrogen (H) in the cellulose, hemicellulose, lignin, and other molecules present to form carbon dioxide (CO₂) and water (H₂O).

Gasification

Gasification is a partial oxidation process whereby a carbon source such as plant biomass, is broken down into carbon monoxide (CO) and hydrogen (H₂), plus carbon dioxide (CO₂) and possibly hydrocarbon molecules such as methane (CH₄). If

the gasification takes place at a relatively low temperature, such as 700°C to 1000°C, the product gas will have a relatively high level of hydrocarbons compared to high temperature gasification. As a result it may be used directly, to be burned for heat or electricity generation via a steam turbine or, with suitable gas clean up, to run an internal combustion engine for electricity generation. The combustion chamber for a simple boiler may be close coupled with the gasifier, or the producer gas may be cleaned of longer chain hydrocarbons (tars), transported, stored and burned remotely. A gasification system may be closely integrated with a combined cycle gas turbine for electricity generation (IGCC - integrated gasification combined cycle). Higher temperature gasification (1200°C to 1600°C) leads to few hydrocarbons in the product gas, and a higher proportion of CO and H₂. This is known as synthesis gas (syngas or biosyngas) as it can be used to synthesize longer chain hydrocarbons using techniques such as Fischer-Tropsch (FT) synthesis. If the ratio of H₂ to CO is correct (2:1) FT synthesis can be used to convert syngas into high quality synthetic diesel biofuel which is compatible with conventional fossil diesel and diesel engines.

Pyrolysis

As used herein, the term "pyrolysis" means a process that uses slow heating in the absence of oxygen to produce gaseous, oil and char products from biomass. Pyrolysis is a thermal or thermo-chemical conversion of lipid-based, particularly triglyceride-based, materials. The products of pyrolysis include gas, liquid and a solid char, with the proportions of each depending upon the parameters of the process. Lower temperatures (around 400°C) tend to produce more solid char (slow pyrolysis), whereas somewhat higher temperatures (around 500°C) produce a much higher proportion of liquid (bio-oil), provided the vapour residence time is kept down to around 1s or less. After this, secondary reactions take place and increase the gas yield. The bio-oil produced by fast (higher temperature) pyrolysis is a dark brown, mobile liquid with a heating value about half that of conventional fuel oil. It can be burned directly, co-fired, upgraded to other fuels or gasified.

Pyrolysis involves direct thermal cracking of the lipids or a combination of thermal and catalytic cracking. At temperatures of about 400-500°C, cracking occurs, producing short chain hydrocarbons such as alkanes, alkenes, alkadienes, aromatics, olefins and carboxylic acid, as well as carbon monoxide and carbon dioxide.

Four main catalyst types can be used including transition metal catalysts, molecular sieve type catalysts, activated alumina and sodium carbonate (Maher et al., 2007). Examples are given in US 4102938. Alumina (Al₂O₃) activated by acid is an effective catalyst (US 5233109). Molecular sieve catalysts are porous, highly crystalline structures that exhibit size selectivity, so that molecules of only certain

sizes can pass through. These include zeolite catalysts such as ZSM-5 or HZSM-5 which are crystalline materials comprising AlO_4 and SiO_4 and other silica-alumina catalysts. The activity and selectivity of these catalysts depends on the acidity, pore size and pore shape, and typically operate at 300-500°C. Transition metal catalysts are described for example in US 4992605. Sodium carbonate catalyst has been used in the pyrolysis of oils (Dandik and Aksoy, 1998).

Transesterification

"Transesterification" as used herein is the conversion of lipids, principally triacylglycerols, into fatty acid methyl esters or ethyl esters using short chain alcohols such as methanol or ethanol, in the presence of a catalyst such as alkali or acid. Methanol is used more commonly due to low cost and availability. The catalysts may be homogeneous catalysts, heterogeneous catalysts or enzymatic catalysts. Homogeneous catalysts include ferric sulphate followed by KOH. Heterogeneous catalysts include CaO , K_3PO_4 , and WO_3/ZrO_2 . Enzymatic catalysts include Novozyme 435 produced from *Candida antarctica*.

Anaerobic Digestion

Anaerobic digestion is the process whereby bacteria break down organic material in the absence of air, yielding a biogas containing methane. The products of this process are biogas (principally methane (CH_4) and carbon dioxide (CO_2)), a solid residue (fibre or digestate) that is similar, but not identical, to compost and a liquid liquor that can be used as a fertilizer. The methane can be burned for heat or electricity generation. The solid residue of the anaerobic digestion process can be used as a soil conditioner or alternatively can be burned as a fuel, or gasified.

Anaerobic digestion is typically performed on biological material in an aqueous slurry. However there are an increasing number of dry digesters. Mesophilic digestion takes place between 20°C and 40°C and can take a month or two to complete. Thermophilic digestion takes place from 50-65°C and is faster, but the bacteria are more sensitive.

Fermentation

Conventional fermentation processes for the production of bioalcohol make use of the starch and sugar components of plant crops. Second generation bioalcohol precedes this with acid and/or enzymatic hydrolysis of hemicellulose and cellulose into fermentable saccharides to make use of a much larger proportion of available biomass. More detail is provided below under the heading "*Fermentation processes for lipid production*".

Composting

Composting is the aerobic decomposition of organic matter by microorganisms. It is typically performed on relatively dry material rather than a slurry. Instead of, or in addition to, collecting the flammable biogas emitted, the exothermic nature of the composting process can be exploited and the heat produced used, usually using a heat pump.

Production of Non-Polar Lipids

Techniques that are routinely practiced in the art can be used to extract, process, purify and analyze the non-polar lipids produced by cells, organisms or parts thereof of the instant invention. Such techniques are described and explained throughout the literature in sources such as, Fereidoon Shahidi, Current Protocols in Food Analytical Chemistry, John Wiley & Sons, Inc. (2001) D1.1.1-D1.1.11, and Perez-Vich et al. (1998).

Production of seedoil

Typically, plant seeds are cooked, pressed, and/or extracted to produce crude seedoil, which is then degummed, refined, bleached, and deodorized. Generally, techniques for crushing seed are known in the art. For example, oilseeds can be tempered by spraying them with water to raise the moisture content to, for example, 8.5%, and flaked using a smooth roller with a gap setting of 0.23 to 0.27 mm. Depending on the type of seed, water may not be added prior to crushing. Application of heat deactivates enzymes, facilitates further cell rupturing, coalesces the lipid droplets, and agglomerates protein particles, all of which facilitate the extraction process.

In an embodiment, the majority of the seedoil is released by passage through a screw press. Cakes expelled from the screw press are then solvent extracted for example, with hexane, using a heat traced column. Alternatively, crude seedoil produced by the pressing operation can be passed through a settling tank with a slotted wire drainage top to remove the solids that are expressed with the seedoil during the pressing operation. The clarified seedoil can be passed through a plate and frame filter to remove any remaining fine solid particles. If desired, the seedoil recovered from the extraction process can be combined with the clarified seedoil to produce a blended crude seedoil.

Once the solvent is stripped from the crude seedoil, the pressed and extracted portions are combined and subjected to normal lipid processing procedures (i.e., degumming, caustic refining, bleaching, and deodorization).

In an embodiment, the oil and/or protein content of the seed is analysed by near-infrared reflectance spectroscopy as described in Hom et al. (2007).

Degumming

5 Degumming is an early step in the refining of oils and its primary purpose is the removal of most of the phospholipids from the oil, which may be present as approximately 1-2% of the total extracted lipid. Addition of ~2% of water, typically containing phosphoric acid, at 70–80°C to the crude oil results in the separation of most of the phospholipids accompanied by trace metals and pigments. The insoluble
10 material that is removed is mainly a mixture of phospholipids and triacylglycerols and is also known as lecithin. Degumming can be performed by addition of concentrated phosphoric acid to the crude seedoil to convert non-hydratable phosphatides to a hydratable form, and to chelate minor metals that are present. Gum is separated from the seedoil by centrifugation. The seedoil can be refined by addition of a sufficient
15 amount of a sodium hydroxide solution to titrate all of the fatty acids and removing the soaps thus formed.

Alkali refining

Alkali refining is one of the refining processes for treating crude oil,
20 sometimes also referred to as neutralization. It usually follows degumming and precedes bleaching. Following degumming, the seedoil can be treated by the addition of a sufficient amount of an alkali solution to titrate all of the fatty acids and phosphoric acids, and removing the soaps thus formed. Suitable alkaline materials include sodium hydroxide, potassium hydroxide, sodium carbonate, lithium hydroxide,
25 calcium hydroxide, calcium carbonate and ammonium hydroxide. This process is typically carried out at room temperature and removes the free fatty acid fraction. Soap is removed by centrifugation or by extraction into a solvent for the soap, and the neutralised oil is washed with water. If required, any excess alkali in the oil may be neutralized with a suitable acid such as hydrochloric acid or sulphuric acid.

30

Bleaching

Bleaching is a refining process in which oils are heated at 90–120°C for 10–30 minutes in the presence of a bleaching earth (0.2–2.0%) and in the absence of oxygen by operating with nitrogen or steam or in a vacuum. This step in oil processing is
35 designed to remove unwanted pigments (carotenoids, chlorophyll, gossypol etc), and the process also removes oxidation products, trace metals, sulphur compounds and traces of soap.

Deodorization

Deodorization is a treatment of oils and fats at a high temperature (200–260°C) and low pressure (0.1–1 mm Hg). This is typically achieved by introducing steam into the seedoil at a rate of about 0.1 ml/minute/100 ml of seedoil.

5 Deodorization can be performed by heating the seedoil to 260°C under vacuum, and slowly introducing steam into the seedoil at a rate of about 0.1 ml/minute/100 ml of seedoil. After about 30 minutes of sparging, the seedoil is allowed to cool under vacuum. The seedoil is typically transferred to a glass container and flushed with argon before being stored under refrigeration. If the amount of seedoil is limited, the

10 seedoil can be placed under vacuum for example, in a Parr reactor and heated to 260°C for the same length of time that it would have been deodorized. This treatment improves the colour of the seedoil and removes a majority of the volatile substances or odorous compounds including any remaining free fatty acids, monoacylglycerols and oxidation products.

15

Winterisation

Winterization is a process sometimes used in commercial production of oils for the separation of oils and fats into solid (stearin) and liquid (olein) fractions by crystallization at sub-ambient temperatures. It was applied originally to cottonseed oil

20 to produce a solid-free product. It is typically used to decrease the saturated fatty acid content of oils.

Plant biomass for the production of lipids

Parts of plants involved in photosynthesis (e.g., and stems and leaves of higher

25 plants and aquatic plants such as algae) can also be used to produce lipid. Independent of the type of plant, there are several methods for extracting lipids from green biomass. One way is physical extraction, which often does not use solvent extraction. It is a "traditional" way using several different types of mechanical extraction. Expeller pressed extraction is a common type, as are the screw press and

30 ram press extraction methods. The amount of lipid extracted using these methods varies widely, depending upon the plant material and the mechanical process employed. Mechanical extraction is typically less efficient than solvent extraction described below.

In solvent extraction, an organic solvent (e.g., hexane) is mixed with at least

35 the genetically modified plant green biomass, preferably after the green biomass is dried and ground. Of course, other parts of the plant besides the green biomass (e.g., lipid-containing seeds) can be ground and mixed in as well. The solvent dissolves the lipid in the biomass and the like, which solution is then separated from the biomass by

mechanical action (e.g., with the pressing processes above). This separation step can also be performed by filtration (e.g., with a filter press or similar device) or centrifugation etc. The organic solvent can then be separated from the non-polar lipid (e.g., by distillation). This second separation step yields non-polar lipid from the
5 plant and can yield a re-usable solvent if one employs conventional vapor recovery.

If, for instance, vegetative tissue as described herein, is not to be used immediately to extract, and/or process, the lipid it is preferably handled post-harvest to ensure the lipid content does not decrease, or such that any decrease in lipid content is minimized as much as possible (see, for example, Christie, 1993). In one
10 embodiment, the vegetative tissue is frozen as soon as possible after harvesting using, for example, dry ice or liquid nitrogen. In another embodiment, the vegetative tissue is stored at a cold temperature, for example -20°C or -60°C in an atmosphere of nitrogen.

15 *Algae for the production of lipids*

Algae can produce 10 to 100 times as much mass as terrestrial plants in a year. In addition to being a prolific organism, algae are also capable of producing oils and starches that can be converted into biofuels.

The specific algae most useful for biofuel production are known as microalgae,
20 consisting of small, often unicellular, types. These algae can grow almost anywhere. With more than 100,000 known species of diatoms (a type of alga), 40,000 known species of green plant-like algae, and smaller numbers of other algae species, algae will grow rapidly in nearly any environment, with almost any kind of water. Specifically, useful algae can be grown in marginal areas with limited or poor quality
25 water, such as in the arid and mostly empty regions of the American Southwest. These areas also have abundant sunshine for photosynthesis. In short, algae can be an ideal organism for production of biofuels - efficient growth, needing no premium land or water, not competing with food crops, needing much smaller amounts of land than food crops, and storing energy in a desirable form.

30 Algae can store energy in its cell structure in the form of either oil or starch. Stored oil can be as much as 60% of the weight of the algae. Certain species which are highly prolific in oil or starch production have been identified, and growing conditions have been tested. Processes for extracting and converting these materials to fuels have also been developed.

35 The most common oil-producing algae can generally include, or consist essentially of, the diatoms (bacillariophytes), green algae (chlorophytes), blue-green algae (cyanophytes), and golden-brown algae (chrysophytes). In addition a fifth group known as haptophytes may be used. Groups include brown algae and

heterokonts. Specific non-limiting examples algae include the Classes: Chlorophyceae, Eustigmatophyceae, Prymnesiophyceae, Bacillariophyceae. Bacillariophytes capable of oil production include the genera Amphipleura, Amphora, Chaetoceros, Cyclotella, Cymbella, Fragilaria, Hantzschia, Navicula, Nitzschia, 5 Phaeodactylum, and Thalassiosira. Specific non-limiting examples of chlorophytes capable of oil production include Ankistrodesmus, Botryococcus, Chlorella, Chlorococcum, Dunaliella, Monoraphidium, Oocystis, Scenedesmus, and Tetraselmis. In one aspect, the chlorophytes can be Chlorella or Dunaliella. Specific non-limiting examples of cyanophytes capable of oil production include Oscillatoria and 10 Synechococcus. A specific example of chrysophytes capable of oil production includes Boekelovia. Specific non-limiting examples of haptophytes include Isochysis and Pleurochysis.

Specific algae useful in the present invention include, for example, *Chlamydomonas* sp. such as *Chlamydomonas reinhardtii*, *Dunaliella* sp. such as 15 *Dunaliella salina*, *Dunaliella tertiolecta*, *D. acidophila*, *D. bardawil*, *D. bioculata*, *D. lateralis*, *D. maritima*, *D. minuta*, *D. parva*, *D. peircei*, *D. polymorpha*, *D. primolecta*, *D. pseudosalina*, *D. quartolecta*, *D. viridis*, *Haematococcus* sp., *Chlorella* sp. such as *Chlorella vulgaris*, *Chlorella sorokiniana* or *Chlorella protothecoides*, *Thraustochytrium* sp., *Schizochytrium* sp., *Volvox* sp., *Nannochloropsis* sp., 20 *Botryococcus braunii* which can contain over 60wt% lipid, *Phaeodactylum tricornutum*, *Thalassiosira pseudonana*, *Isochrysis* sp., *Pavlova* sp., *Chlorococcum* sp., *Ellipsoidion* sp., *Neochloris* sp., *Scenedesmus* sp.

Further, the oil-producing algae of the present invention can include a combination of an effective amount of two or more strains in order to maximize 25 benefits from each strain. As a practical matter, it can be difficult to achieve 100% purity of a single strain of algae or a combination of desired algae strains. However, when discussed herein, the oil-producing algae is intended to cover intentionally introduced strains of algae, while foreign strains are preferably minimized and kept below an amount which would detrimentally affect yields of desired oil-producing 30 algae and algal oil. Undesirable algae strains can be controlled and/or eliminated using any number of techniques. For example, careful control of the growth environment can reduce introduction of foreign strains. Alternatively, or in addition to other techniques, a virus selectively chosen to specifically target only the foreign strains can be introduced into the growth reservoirs in an amount which is effective to 35 reduce and/or eliminate the foreign strain. An appropriate virus can be readily identified using conventional techniques. For example, a sample of the foreign algae will most often include small amounts of a virus which targets the foreign algae. This virus can be isolated and grown in order to produce amounts which would effectively

control or eliminate the foreign algae population among the more desirable oil-producing algae.

Algaculture is a form of aquaculture involving the farming of species of algae (including microalgae, also referred to as phytoplankton, microphytes, or planktonic algae, and macroalgae, commonly known as seaweed).

Commercial and industrial algae cultivation has numerous uses, including production of food ingredients, food, and algal fuel.

Mono or mixed algal cultures can be cultured in open-ponds (such as raceway-type ponds and lakes) or photobioreactors.

Algae can be harvested using microscreens, by centrifugation, by flocculation (using for example, chitosan, alum and ferric chloride) and by froth flotation. Interrupting the carbon dioxide supply can cause algae to flocculate on its own, which is called "autoflocculation". In froth flotation, the cultivator aerates the water into a froth, and then skims the algae from the top. Ultrasound and other harvesting methods are currently under development.

Lipid may be separated from the algae by mechanical crushing. When algae is dried it retains its lipid content, which can then be "pressed" out with an oil press. Since different strains of algae vary widely in their physical attributes, various press configurations (screw, expeller, piston, etc.) work better for specific algae types.

Osmotic shock is sometimes used to release cellular components such as lipid from algae. Osmotic shock is a sudden reduction in osmotic pressure and can cause cells in a solution to rupture.

Ultrasonic extraction can accelerate extraction processes, in particular enzymatic extraction processes employed to extract lipid from algae. Ultrasonic waves are used to create cavitation bubbles in a solvent material. When these bubbles collapse near the cell walls, the resulting shock waves and liquid jets cause those cells walls to break and release their contents into a solvent.

Chemical solvents (for example, hexane, benzene, petroleum ether) are often used in the extraction of lipids from algae. Soxhlet extraction can be used to extract lipids from algae through repeated washing, or percolation, with an organic solvent under reflux in a special glassware.

Enzymatic extraction may be used to extract lipids from algae. Enzymatic extraction uses enzymes to degrade the cell walls with water acting as the solvent. The enzymatic extraction can be supported by ultrasonication.

Supercritical CO₂ can also be used as a solvent. In this method, CO₂ is liquefied under pressure and heated to the point that it becomes supercritical (having properties of both a liquid and a gas), allowing it to act as a solvent.

Fermentation processes for lipid production

As used herein, the term the "fermentation process" refers to any fermentation process or any process comprising a fermentation step. A fermentation process includes, without limitation, fermentation processes used to produce alcohols (e.g., ethanol, methanol, butanol), organic acids (e.g., citric acid, acetic acid, itaconic acid, lactic acid, gluconic acid), ketones (e.g., acetone), amino acids (e.g., glutamic acid), gases (e.g., H₂ and CO₂), antibiotics (e.g., penicillin and tetracycline), enzymes, vitamins (e.g., riboflavin, beta-carotene), and hormones. Fermentation processes also include fermentation processes used in the consumable alcohol industry (e.g., beer and wine), dairy industry (e.g., fermented dairy products), leather industry and tobacco industry. Preferred fermentation processes include alcohol fermentation processes, as are well known in the art. Preferred fermentation processes are anaerobic fermentation processes, as are well known in the art. Suitable fermenting cells, typically microorganisms that are able to ferment, that is, convert, sugars such as glucose or maltose, directly or indirectly into the desired fermentation product.

Examples of fermenting microorganisms include fungal organisms such as yeast, preferably an oleaginous organism. As used herein, an "oleaginous organism" is one which accumulates at least 25% of its dry weight as triglycerides. As used herein, "yeast" includes *Saccharomyces* spp., *Saccharomyces cerevisiae*, *Saccharomyces carlbergensis*, *Candida* spp., *Kluveromyces* spp., *Pichia* spp., *Hansenula* spp., *Trichoderma* spp., *Lipomyces starkey*, and *Yarrowia lipolytica*. Preferred yeast include *Yarrowia lipolytica* or other oleaginous yeasts and strains of the *Saccharomyces* spp., and in particular, *Saccharomyces cerevisiae*.

In one embodiment, the fermenting microorganism is a transgenic organism that comprises one or more exogenous polynucleotides, wherein the transgenic organism has an increased level of one or more non-polar lipids when compared to a corresponding organism lacking the one or more exogenous polynucleotides. The transgenic microorganism is preferably grown under conditions that optimize activity of fatty acid biosynthetic genes and fatty acid acyltransferase genes. This leads to production of the greatest and the most economical yield of lipid. In general, media conditions that may be optimized include the type and amount of carbon source, the type and amount of nitrogen source, the carbon-to-nitrogen ratio, the oxygen level, growth temperature, pH, length of the biomass production phase, length of the lipid accumulation phase and the time of cell harvest.

Fermentation media must contain a suitable carbon source. Suitable carbon sources may include, but are not limited to: monosaccharides (e.g., glucose, fructose), disaccharides (e.g., lactose, sucrose), oligosaccharides, polysaccharides (e.g., starch, cellulose or mixtures thereof), sugar alcohols (e.g., glycerol) or mixtures from

renewable feedstocks (e.g., cheese whey permeate, cornsteep liquor, sugar beet molasses, barley malt). Additionally, carbon sources may include alkanes, fatty acids, esters of fatty acids, monoglycerides, diglycerides, triglycerides, phospholipids and various commercial sources of fatty acids including vegetable oils (e.g., soybean oil) and animal fats. Additionally, the carbon substrate may include one-carbon substrates (e.g., carbon dioxide, methanol, formaldehyde, formate, carbon-containing amines) for which metabolic conversion into key biochemical intermediates has been demonstrated. Hence it is contemplated that the source of carbon utilized in the present invention may encompass a wide variety of carbon-containing substrates and will only be limited by the choice of the host microorganism. Although all of the above mentioned carbon substrates and mixtures thereof are expected to be suitable in the present invention, preferred carbon substrates are sugars and/or fatty acids. Most preferred is glucose and/or fatty acids containing between 10-22 carbons.

Nitrogen may be supplied from an inorganic (e.g., $(\text{NH}_4)_2\text{SO}_4$) or organic source (e.g., urea, glutamate). In addition to appropriate carbon and nitrogen sources, the fermentation media may also contain suitable minerals, salts, cofactors, buffers, vitamins and other components known to those skilled in the art suitable for the growth of the microorganism and promotion of the enzymatic pathways necessary for lipid production.

A suitable pH range for the fermentation is typically between about pH 4.0 to pH 8.0, wherein pH 5.5 to pH 7.0 is preferred as the range for the initial growth conditions. The fermentation may be conducted under aerobic or anaerobic conditions, wherein microaerobic conditions are preferred.

Typically, accumulation of high levels of lipid in the cells of oleaginous microorganisms requires a two-stage process, since the metabolic state must be "balanced" between growth and synthesis/storage of fats. Thus, most preferably, a two-stage fermentation process is necessary for the production of lipids in microorganisms. In this approach, the first stage of the fermentation is dedicated to the generation and accumulation of cell mass and is characterized by rapid cell growth and cell division. In the second stage of the fermentation, it is preferable to establish conditions of nitrogen deprivation in the culture to promote high levels of lipid accumulation. The effect of this nitrogen deprivation is to reduce the effective concentration of AMP in the cells, thereby reducing the activity of the NAD-dependent isocitrate dehydrogenase of mitochondria. When this occurs, citric acid will accumulate, thus forming abundant pools of acetyl-CoA in the cytoplasm and priming fatty acid synthesis. Thus, this phase is characterized by the cessation of cell division followed by the synthesis of fatty acids and accumulation of TAGs.

Although cells are typically grown at about 30°C, some studies have shown increased synthesis of unsaturated fatty acids at lower temperatures. Based on process economics, this temperature shift should likely occur after the first phase of the two-stage fermentation, when the bulk of the microorganism's growth has occurred.

5 It is contemplated that a variety of fermentation process designs may be applied, where commercial production of lipids using the instant nucleic acids is desired. For example, commercial production of lipid from a recombinant microbial host may be produced by a batch, fed-batch or continuous fermentation process.

A batch fermentation process is a closed system wherein the media
10 composition is set at the beginning of the process and not subject to further additions beyond those required for maintenance of pH and oxygen level during the process. Thus, at the beginning of the culturing process the media is inoculated with the desired organism and growth or metabolic activity is permitted to occur without adding additional substrates (i.e., carbon and nitrogen sources) to the medium. In
15 batch processes the metabolite and biomass compositions of the system change constantly up to the time the culture is terminated. In a typical batch process, cells moderate through a static lag phase to a high-growth log phase and finally to a stationary phase, wherein the growth rate is diminished or halted. Left untreated, cells in the stationary phase will eventually die. A variation of the standard batch process
20 is the fed-batch process, wherein the substrate is continually added to the fermentor over the course of the fermentation process. A fed-batch process is also suitable in the present invention. Fed-batch processes are useful when catabolite repression is apt to inhibit the metabolism of the cells or where it is desirable to have limited amounts of substrate in the media at any one time. Measurement of the substrate
25 concentration in fed-batch systems is difficult and therefore may be estimated on the basis of the changes of measurable factors such as pH, dissolved oxygen and the partial pressure of waste gases (e.g., CO₂). Batch and fed-batch culturing methods are common and well known in the art and examples may be found in Brock, In Biotechnology: A Textbook of Industrial Microbiology, 2^{sup}.nd ed., Sinauer
30 Associates, Sunderland, Mass., (1989); or Deshpande (1992).

Commercial production of lipid using the instant cells may also be accomplished by a continuous fermentation process, wherein a defined media is continuously added to a bioreactor while an equal amount of culture volume is removed simultaneously for product recovery. Continuous cultures generally
35 maintain the cells in the log phase of growth at a constant cell density. Continuous or semi-continuous culture methods permit the modulation of one factor or any number of factors that affect cell growth or end product concentration. For example, one approach may limit the carbon source and allow all other parameters to moderate

metabolism. In other systems, a number of factors affecting growth may be altered continuously while the cell concentration, measured by media turbidity, is kept constant. Continuous systems strive to maintain steady state growth and thus the cell growth rate must be balanced against cell loss due to media being drawn off the culture. Methods of modulating nutrients and growth factors for continuous culture processes, as well as techniques for maximizing the rate of product formation, are well known in the art of industrial microbiology and a variety of methods are detailed by Brock, *supra*.

Fatty acids, including PUFAs, may be found in the host microorganism as free fatty acids or in esterified forms such as acylglycerols, phospholipids, sulfolipids or glycolipids, and may be extracted from the host cell through a variety of means well-known in the art.

In general, means for the purification of fatty acids, including PUFAs, may include extraction with organic solvents, sonication, supercritical fluid extraction (e.g., using carbon dioxide), saponification and physical means such as presses, or combinations thereof. Of particular interest is extraction with methanol and chloroform in the presence of water (Bligh and Dyer, 1959). Where desirable, the aqueous layer can be acidified to protonate negatively-charged moieties and thereby increase partitioning of desired products into the organic layer. After extraction, the organic solvents can be removed by evaporation under a stream of nitrogen. When isolated in conjugated forms, the products may be enzymatically or chemically cleaved to release the free fatty acid or a less complex conjugate of interest, and can then be subject to further manipulations to produce a desired end product. Desirably, conjugated forms of fatty acids are cleaved with potassium hydroxide.

If further purification is necessary, standard methods can be employed. Such methods may include extraction, treatment with urea, fractional crystallization, HPLC, fractional distillation, silica gel chromatography, high-speed centrifugation or distillation, or combinations of these techniques. Protection of reactive groups such as the acid or alkenyl groups, may be done at any step through known techniques (e.g., alkylation, iodination). Methods used include methylation of the fatty acids to produce methyl esters. Similarly, protecting groups may be removed at any step. Desirably, purification of fractions containing GLA, STA, ARA, DHA and EPA may be accomplished by treatment with urea and/or fractional distillation.

An example of the use of plant biomass for the production of a biomass slurry using yeast is described in WO 2011/100272.

Uses of Lipids

The lipids produced by the methods described have a variety of uses. In some embodiments, the lipids are used as food oils. In other embodiments, the lipids are refined and used as lubricants or for other industrial uses such as the synthesis of plastics. In some preferred embodiments, the lipids are refined to produce biodiesel.

5

Biofuel

As used herein the term "biofuel" includes biodiesel and bioalcohol. Biodiesel can be made from oils derived from plants, algae and fungi. Bioalcohol is produced from the fermentation of sugar. This sugar can be extracted directly from plants (e.g.,
10 sugarcane), derived from plant starch (e.g., maize or wheat) or made from cellulose (e.g., wood, leaves or stems).

Biofuels currently cost more to produce than petroleum fuels. In addition to processing costs, biofuel crops require planting, fertilising, pesticide and herbicide applications, harvesting and transportation. Plants, algae and fungi of the present
15 invention may reduce production costs of biofuel.

General methods for the production of biofuel can be found in, for example, Maher and Bressler, 2007; Greenwell et al., 2010; Karmakar et al., 2010; Alonso et al., 2010; Lee and Mohamed, 2010; Liu et al., 2010a; Gong and Jiang, 2011; Endalew et al., 2011; Semwal et al., 2011.

20

Bioalcohol

The production of biologically produced alcohols, for example, ethanol, propanol and butanol is well known. Ethanol is the most common bioalcohol.

The basic steps for large scale production of ethanol are: 1) microbial (for example, yeast) fermentation of sugars, 2) distillation, 3) dehydration, and optionally
25 4) denaturing. Prior to fermentation, some crops require saccharification or hydrolysis of carbohydrates such as cellulose and starch into sugars. Saccharification of cellulose is called cellulolysis. Enzymes can be used to convert starch into sugar.

30

Fermentation

Bioalcohol is produced by microbial fermentation of the sugar. Microbial fermentation will currently only work directly with sugars. Two major components of plants, starch and cellulose, are both made up of sugars, and can in principle be converted to sugars for fermentation.

35

Distillation

For the ethanol to be usable as a fuel, the majority of the water must be removed. Most of the water is removed by distillation, but the purity is limited to 95-96% due to the formation of a low-boiling water-ethanol azeotrope with maximum (95.6% m/m (96.5% v/v) ethanol and 4.4% m/m (3.5% v/v) water). This mixture is called hydrous ethanol and can be used as a fuel alone, but unlike anhydrous ethanol, hydrous ethanol is not miscible in all ratios with gasoline, so the water fraction is typically removed in further treatment in order to burn in combination with gasoline in gasoline engines.

10 *Dehydration*

Water can be removed from an azeotropic ethanol/water mixture by dehydration. Azeotropic distillation, used in many early fuel ethanol plants, consists of adding benzene or cyclohexane to the mixture. When these components are added to the mixture, it forms a heterogeneous azeotropic mixture in vapor-liquid-liquid equilibrium, which when distilled produces anhydrous ethanol in the column bottom, and a vapor mixture of water and cyclohexane/benzene. When condensed, this becomes a two-phase liquid mixture. Another early method, called extractive distillation, consists of adding a ternary component which will increase ethanol's relative volatility. When the ternary mixture is distilled, it will produce anhydrous ethanol on the top stream of the column.

A third method has emerged and has been adopted by the majority of modern ethanol plants. This new process uses molecular sieves to remove water from fuel ethanol. In this process, ethanol vapor under pressure passes through a bed of molecular sieve beads. The bead's pores are sized to allow absorption of water while excluding ethanol. After a period of time, the bed is regenerated under vacuum or in the flow of inert atmosphere (e.g. N₂) to remove the absorbed water. Two beds are often used so that one is available to absorb water while the other is being regenerated.

30 *Biodiesel*

The production of biodiesel, or alkyl esters, is well known. There are three basic routes to ester production from lipids: 1) Base catalysed transesterification of the lipid with alcohol; 2) Direct acid catalysed esterification of the lipid with methanol; and 3) Conversion of the lipid to fatty acids, and then to alkyl esters with acid catalysis.

Any method for preparing fatty acid alkyl esters and glyceryl ethers (in which one, two or three of the hydroxy groups on glycerol are etherified) can be used. For example, fatty acids can be prepared, for example, by hydrolyzing or saponifying

triglycerides with acid or base catalysts, respectively, or using an enzyme such as a lipase or an esterase. Fatty acid alkyl esters can be prepared by reacting a fatty acid with an alcohol in the presence of an acid catalyst. Fatty acid alkyl esters can also be prepared by reacting a triglyceride with an alcohol in the presence of an acid or base catalyst. Glycerol ethers can be prepared, for example, by reacting glycerol with an alkyl halide in the presence of base, or with an olefin or alcohol in the presence of an acid catalyst.

In some preferred embodiments, the lipids are transesterified to produce methyl esters and glycerol. In some preferred embodiments, the lipids are reacted with an alcohol (such as methanol or ethanol) in the presence of a catalyst (for example, potassium or sodium hydroxide) to produce alkyl esters. The alkyl esters can be used for biodiesel or blended with petroleum based fuels.

The alkyl esters can be directly blended with diesel fuel, or washed with water or other aqueous solutions to remove various impurities, including the catalysts, before blending. It is possible to neutralize acid catalysts with base. However, this process produces salt. To avoid engine corrosion, it is preferable to minimize the salt concentration in the fuel additive composition. Salts can be substantially removed from the composition, for example, by washing the composition with water.

In another embodiment, the composition is dried after it is washed, for example, by passing the composition through a drying agent such as calcium sulfate.

In yet another embodiment, a neutral fuel additive is obtained without producing salts or using a washing step, by using a polymeric acid, such as Dowex 50™, which is a resin that contains sulfonic acid groups. The catalyst is easily removed by filtration after the esterification and etherification reactions are complete.

Plant triacylglycerols as a biofuel source

Use of plant triacylglycerols for the production of biofuel is reviewed in Durrett et al. (2008). Briefly, plant oils are primarily composed of various triacylglycerols (TAGs), molecules that consist of three fatty acid chains (usually 18 or 16 carbons long) esterified to glycerol. The fatty acyl chains are chemically similar to the aliphatic hydrocarbons that make up the bulk of the molecules found in petrol and diesel. The hydrocarbons in petrol contain between 5 and 12 carbon atoms per molecule, and this volatile fuel is mixed with air and ignited with a spark in a conventional engine. In contrast, diesel fuel components typically have 10–15 carbon atoms per molecule and are ignited by the very high compression obtained in a diesel engine. However, most plant TAGs have a viscosity range that is much higher than that of conventional diesel: $17.3\text{--}32.9\text{ mm}^2\text{s}^{-1}$ compared to $1.9\text{--}4.1\text{ mm}^2\text{s}^{-1}$, respectively (ASTM D975; Knothe and Steidley, 2005). This higher viscosity results

in poor fuel atomization in modern diesel engines, leading to problems derived from incomplete combustion such as carbon deposition and coking (Ryan et al., 1984). To overcome this problem, TAGs are converted to less viscous fatty acid esters by esterification with a primary alcohol, most commonly methanol. The resulting fuel is commonly referred to as biodiesel and has a dynamic viscosity range from 1.9 to 6.0 mm²s⁻¹ (ASTM D6751). The fatty acid methyl esters (FAMES) found in biodiesel have a high energy density as reflected by their high heat of combustion, which is similar, if not greater, than that of conventional diesel (Knothe, 2005). Similarly, the cetane number (a measure of diesel ignition quality) of the FAMES found in biodiesel exceeds that of conventional diesel (Knothe, 2005).

Plant oils are mostly composed of five common fatty acids, namely palmitate (16:0), stearate (18:0), oleate (18:1), linoleate (18:2) and linolenate (18:3), although, depending on the particular species, longer or shorter fatty acids may also be major constituents. These fatty acids differ from each other in terms of acyl chain length and number of double bonds, leading to different physical properties. Consequently, the fuel properties of biodiesel derived from a mixture of fatty acids are dependent on that composition. Altering the fatty acid profile can therefore improve fuel properties of biodiesel such as cold-temperature flow characteristics, oxidative stability and NO_x emissions. Altering the fatty acid composition of TAGs may reduce the viscosity of the plant oils, eliminating the need for chemical modification, thus improving the cost-effectiveness of biofuels.

Most plant oils are derived from triacylglycerols stored in seeds. However, the present invention provides methods for also increasing oil content in vegetative tissues. The plant tissues of the present invention have an increased total lipid yield. Furthermore, the level of oleic acid is increased significantly while the polyunsaturated fatty acid alpha linolenic acid was reduced.

Once a leaf is developed, it undergoes a developmental change from sink (absorbing nutrients) to source (providing sugars). In food crops, most sugars are translocated out of source leaves to support growth of new leaves, roots and fruits. Because translocation of carbohydrate is an active process, there is a loss of carbon and energy during translocation. Furthermore, after the developing seed takes up carbon from the plant, there are additional carbon and energy losses associated with the conversion of carbohydrate into the oil, protein or other major components of the seed (Goffman et al., 2005). Plants of the present invention increase the energy content of leaves and/or stems such that the whole above-ground plant may be harvested and used to produce biofuel.

Algae as a biofuel source

Algae store oil inside the cell body, sometimes but not always in vesicles. This oil can be recovered in several relatively simple ways, including solvents, heat, and/or pressure. However, these methods typically recover only about 80% to 90% of the stored oil. Processes which offer more effective oil extraction methods which can recover close to 100% of the stored oil at low cost as known in the art. These processes include or consist of depolymerizing, such as biologically breaking the walls of the algal cell and/or oil vesicles, if present, to release the oil from the oil-producing algae.

10 In addition, a large number of viruses exist which invade and rupture algae cells, and can thereby release the contents of the cell in particular stored oil or starch. Such viruses are an integral part of the algal ecosystem, and many of the viruses are specific to a single type of algae. Specific examples of such viruses include the chlorella virus PBCV-1 (Paramecium Bursaria Chlorella Virus) which is specific to
15 certain Chlorella algae, and cyanophages such as SM-1, P-60, and AS-1 specific to the blue-green algae Synechococcus. The particular virus selected will depend on the particular species of algae to be used in the growth process. One aspect of the present invention is the use of such a virus to rupture the algae so that oil contained inside the algae cell wall can be recovered. In another detailed aspect of the present invention, a
20 mixture of biological agents can be used to rupture the algal cell wall and/or oil vesicles.

Mechanical crushing, for example, an expeller or press, a hexane or butane solvent recovery step, supercritical fluid extraction, or the like can also be useful in extracting the oil from oil vesicles of the oil-producing algae. Alternatively,
25 mechanical approaches can be used in combination with biological agents in order to improve reaction rates and/or separation of materials. Regardless of the particular biological agent or agents chosen such can be introduced in amounts which are sufficient to serve as the primary mechanism by which algal oil is released from oil vesicles in the oil-producing algae, i.e. not a merely incidental presence of any of
30 these.

Once the oil has been released from the algae it can be recovered or separated
16 from a slurry of algae debris material, for example, cellular residue, oil, enzyme, by-products, etc. This can be done by sedimentation or centrifugation, with centrifugation generally being faster. Starch production can follow similar separation
35 processes.

An algal feed can be formed from a biomass feed source as well as an algal feed source. Biomass from either algal or terrestrial sources can be depolymerized in a variety of ways such as, but not limited to saccharification, hydrolysis or the like.

The source material can be almost any sufficiently voluminous cellulose, lignocellulose, polysaccharide or carbohydrate, glycoprotein, or other material making up the cell wall of the source material.

5 The fermentation stage can be conventional in its use of yeast to ferment sugar to alcohol. The fermentation process produces carbon dioxide, alcohol, and algal husks. All of these products can be used elsewhere in the process and systems of the present invention, with substantially no unused material or wasted heat. Alternatively, if ethanol is so produced, it can be sold as a product or used to produce ethyl acetate for the transesterification process. Similar considerations would apply to alcohols
10 other than ethanol.

Algal oil can be converted to biodiesel through a process of direct hydrogenation or transesterification of the algal oil. Algal oil is in a similar form as most vegetable oils, which are in the form of triglycerides. A triglyceride consists of three fatty acid chains, one attached to each of the three carbon atoms in a glycerol
15 backbone. This form of oil can be burned directly. However, the properties of the oil in this form are not ideal for use in a diesel engine, and without modification, the engine will soon run poorly or fail. In accordance with the present invention, the triglyceride is converted into biodiesel, which is similar to but superior to petroleum diesel fuel in many respects.

20 One process for converting the triglyceride to biodiesel is transesterification, and includes reacting the triglyceride with alcohol or other acyl acceptor to produce free fatty acid esters and glycerol. The free fatty acids are in the form of fatty acid alkyl esters (FAAE).

With the chemical process, additional steps are needed to separate the catalyst
25 and clean the fatty acids. In addition, if ethanol is used as the acyl acceptor, it must be essentially dry to prevent production of soap via saponification in the process, and the glycerol must be purified. The biological process, by comparison, can accept ethanol in a less dry state and the cleaning and purification of the biodiesel and glycerol are much easier.

30 Transesterification often uses a simple alcohol, typically methanol derived from petroleum. When methanol is used the resultant biodiesel is called fatty acid methyl ester (FAME) and most biodiesel sold today, especially in Europe, is FAME. However, ethanol can also be used as the alcohol in transesterification, in which case the biodiesel is fatty acid ethyl ester (FAEE). In the US, the two types are usually not
35 distinguished, and are collectively known as fatty acid alkyl esters (FAAE); which as a generic term can apply regardless of the acyl acceptor used. Direct hydrogenation can also be utilized to convert at least a portion of the algal oil to a biodiesel. As such, in one aspect, the biodiesel product can include an alkane.

The algal triglyceride can also be converted to biodiesel by direct hydrogenation. In this process, the products are alkane chains, propane, and water. The glycerol backbone is hydrogenated to propane, so there is substantially no glycerol produced as a byproduct. Furthermore, no alcohol or transesterification catalysts are needed. All of the biomass can be used as feed for the oil-producing algae with none needed for fermentation to produce alcohol for transesterification. The resulting alkanes are pure hydrocarbons, with no oxygen, so the biodiesel produced in this way has a slightly higher energy content than the alkyl esters, degrades more slowly, does not attract water, and has other desirable chemical properties.

Feedstuffs

The present invention includes compositions which can be used as feedstuffs. For purposes of the present invention, "feedstuffs" include any food or preparation for human or animal consumption (including for enteral and/or parenteral consumption) which when taken into the body: (1) serve to nourish or build up tissues or supply energy, and/or (2) maintain, restore or support adequate nutritional status or metabolic function. Feedstuffs of the invention include nutritional compositions for babies and/or young children.

Feedstuffs of the invention comprise for example, a cell of the invention, a plant of the invention, the plant part of the invention, the seed of the invention, an extract of the invention, the product of a method of the invention, the product of a fermentation process of the invention, or a composition along with a suitable carrier(s). The term "carrier" is used in its broadest sense to encompass any component which may or may not have nutritional value. As the person skilled in the art will appreciate, the carrier must be suitable for use (or used in a sufficiently low concentration) in a feedstuff, such that it does not have deleterious effect on an organism which consumes the feedstuff.

The feedstuff of the present invention comprises a lipid produced directly or indirectly by use of the methods, cells or organisms disclosed herein. The composition may either be in a solid or liquid form. Additionally, the composition may include edible macronutrients, vitamins, and/or minerals in amounts desired for a particular use. The amounts of these ingredients will vary depending on whether the composition is intended for use with normal individuals or for use with individuals having specialized needs such as individuals suffering from metabolic disorders and the like.

Examples of suitable carriers with nutritional value include, but are not limited to, macronutrients such as edible fats, carbohydrates and proteins. Examples of such

edible fats include, but are not limited to, coconut oil, borage oil, fungal oil, black current oil, soy oil, and mono- and di-glycerides. Examples of such carbohydrates include, but are not limited to, glucose, edible lactose, and hydrolyzed starch. Additionally, examples of proteins which may be utilized in the nutritional
5 composition of the invention include, but are not limited to, soy proteins, electrodialysed whey, electrodialysed skim milk, milk whey, or the hydrolysates of these proteins.

With respect to vitamins and minerals, the following may be added to the feedstuff compositions of the present invention, calcium, phosphorus, potassium,
10 sodium, chloride, magnesium, manganese, iron, copper, zinc, selenium, iodine, and vitamins A, E, D, C, and the B complex. Other such vitamins and minerals may also be added.

The components utilized in the feedstuff compositions of the present invention can be of semi-purified or purified origin. By semi-purified or purified is meant a
15 material which has been prepared by purification of a natural material or by *de novo* synthesis.

A feedstuff composition of the present invention may also be added to food even when supplementation of the diet is not required. For example, the composition may be added to food of any type, including, but not limited to, margarine, modified
20 butter, cheeses, milk, yogurt, chocolate, candy, snacks, salad oils, cooking oils, cooking fats, meats, fish and beverages.

The genus *Saccharomyces spp* is used in both brewing of beer and wine making and also as an agent in baking, particularly bread. Yeast is a major constituent of vegetable extracts. Yeast is also used as an additive in animal feed. It
25 will be apparent that genetically modified yeast strains can be provided which are adapted to synthesize lipid as described herein. These yeast strains can then be used in food stuffs and in wine and beer making to provide products which have enhanced lipid content.

Additionally, lipid produced in accordance with the present invention or host
30 cells transformed to contain and express the subject genes may also be used as animal food supplements to alter an animal's tissue or milk fatty acid composition to one more desirable for human or animal consumption. Examples of such animals include sheep, cattle, horses and the like.

Furthermore, feedstuffs of the invention can be used in aquaculture to increase
35 the levels of fatty acids in fish for human or animal consumption.

Preferred feedstuffs of the invention are the plants, seed and other plant parts such as leaves, fruits and stems which may be used directly as food or feed for humans or other animals. For example, animals may graze directly on such plants

grown in the field, or be fed more measured amounts in controlled feeding. The invention includes the use of such plants and plant parts as feed for increasing the polyunsaturated fatty acid levels in humans and other animals.

Compositions

- 5 The present invention also encompasses compositions, particularly pharmaceutical compositions, comprising one or more lipids produced using the methods of the invention.

10 A pharmaceutical composition may comprise one or more of the lipids, in combination with a standard, well-known, non-toxic pharmaceutically-acceptable carrier, adjuvant or vehicle such as phosphate-buffered saline, water, ethanol, polyols, vegetable oils, a wetting agent, or an emulsion such as a water/oil emulsion. The composition may be in either a liquid or solid form. For example, the composition may be in the form of a tablet, capsule, ingestible liquid, powder, topical ointment or cream. Proper fluidity can be maintained for example, by the maintenance of the
15 required particle size in the case of dispersions and by the use of surfactants. It may also be desirable to include isotonic agents for example, sugars, sodium chloride, and the like. Besides such inert diluents, the composition can also include adjuvants such as wetting agents, emulsifying and suspending agents, sweetening agents, flavoring agents and perfuming agents.

- 20 Suspensions, in addition to the active compounds, may comprise suspending agents such as ethoxylated isostearyl alcohols, polyoxyethylene sorbitol and sorbitan esters, microcrystalline cellulose, aluminum metahydroxide, bentonite, agar-agar, and tragacanth, or mixtures of these substances.

25 Solid dosage forms such as tablets and capsules can be prepared using techniques well known in the art. For example, lipid produced in accordance with the present invention can be tableted with conventional tablet bases such as lactose, sucrose, and cornstarch in combination with binders such as acacia, cornstarch or gelatin, disintegrating agents such as potato starch or alginic acid, and a lubricant such as stearic acid or magnesium stearate. Capsules can be prepared by incorporating
30 these excipients into a gelatin capsule along with antioxidants and the relevant lipid(s).

For intravenous administration, the lipids produced in accordance with the present invention or derivatives thereof may be incorporated into commercial formulations.

- 35 A typical dosage of a particular fatty acid is from 0.1 mg to 20 g, taken from one to five times per day (up to 100 g daily) and is preferably in the range of from about 10 mg to about 1, 2, 5, or 10 g daily (taken in one or multiple doses). As known in the art, a minimum of about 300 mg/day of fatty acid, especially polyunsaturated

fatty acid, is desirable. However, it will be appreciated that any amount of fatty acid will be beneficial to the subject.

Possible routes of administration of the pharmaceutical compositions of the present invention include for example, enteral and parenteral. For example, a liquid
5 preparation may be administered orally. Additionally, a homogenous mixture can be completely dispersed in water, admixed under sterile conditions with physiologically acceptable diluents, preservatives, buffers or propellants to form a spray or inhalant.

The dosage of the composition to be administered to the subject may be determined by one of ordinary skill in the art and depends upon various factors such
10 as weight, age, overall health, past history, immune status, etc., of the subject.

Additionally, the compositions of the present invention may be utilized for cosmetic purposes. The compositions may be added to pre-existing cosmetic compositions, such that a mixture is formed, or a fatty acid produced according to the invention may be used as the sole "active" ingredient in a cosmetic composition.

15

Polypeptides

The terms "polypeptide" and "protein" are generally used interchangeably.

A polypeptide or class of polypeptides may be defined by the extent of identity (% identity) of its amino acid sequence to a reference amino acid sequence, or by
20 having a greater % identity to one reference amino acid sequence than to another. The % identity of a polypeptide to a reference amino acid sequence is typically determined by GAP analysis (Needleman and Wunsch, 1970; GCG program) with parameters of a gap creation penalty = 5, and a gap extension penalty = 0.3. The query sequence is at least 100 amino acids in length and the GAP analysis aligns the
25 two sequences over a region of at least 100 amino acids. Even more preferably, the query sequence is at least 250 amino acids in length and the GAP analysis aligns the two sequences over a region of at least 250 amino acids. Even more preferably, the GAP analysis aligns two sequences over their entire length. The polypeptide or class of polypeptides may have the same enzymatic activity as, or a different activity than,
30 or lack the activity of, the reference polypeptide. Preferably, the polypeptide has an enzymatic activity of at least 10% of the activity of the reference polypeptide.

As used herein a "biologically active fragment" is a portion of a polypeptide of the invention which maintains a defined activity of a full-length reference polypeptide for example, MGAT activity. Biologically active fragments as used herein exclude
35 the full-length polypeptide. Biologically active fragments can be any size portion as long as they maintain the defined activity. Preferably, the biologically active fragment maintains at least 10% of the activity of the full length polypeptide.

With regard to a defined polypeptide or enzyme, it will be appreciated that % identity figures higher than those provided herein will encompass preferred embodiments. Thus, where applicable, in light of the minimum % identity figures, it is preferred that the polypeptide/enzyme comprises an amino acid sequence which is

5 at least 60%, more preferably at least 65%, more preferably at least 70%, more preferably at least 75%, more preferably at least 80%, more preferably at least 85%, more preferably at least 90%, more preferably at least 91%, more preferably at least 92%, more preferably at least 93%, more preferably at least 94%, more preferably at least 95%, more preferably at least 96%, more preferably at least 97%, more

10 preferably at least 98%, more preferably at least 99%, more preferably at least 99.1%, more preferably at least 99.2%, more preferably at least 99.3%, more preferably at least 99.4%, more preferably at least 99.5%, more preferably at least 99.6%, more preferably at least 99.7%, more preferably at least 99.8%, and even more preferably at least 99.9% identical to the relevant nominated SEQ ID NO.

15 Amino acid sequence mutants of the polypeptides defined herein can be prepared by introducing appropriate nucleotide changes into a nucleic acid defined herein, or by *in vitro* synthesis of the desired polypeptide. Such mutants include for example, deletions, insertions, or substitutions of residues within the amino acid sequence. A combination of deletions, insertions and substitutions can be made to

20 arrive at the final construct, provided that the final polypeptide product possesses the desired characteristics.

Mutant (altered) polypeptides can be prepared using any technique known in the art, for example, using directed evolution or rational design strategies (see below). Products derived from mutated/altered DNA can readily be screened using

25 techniques described herein to determine if they possess fatty acid acyltransferase activity, for example, MGAT, DGAT, or GPAT/phosphatase activity.

In designing amino acid sequence mutants, the location of the mutation site and the nature of the mutation will depend on characteristic(s) to be modified. The sites for mutation can be modified individually or in series for example, by (1)

30 substituting first with conservative amino acid choices and then with more radical selections depending upon the results achieved, (2) deleting the target residue, or (3) inserting other residues adjacent to the located site.

Amino acid sequence deletions generally range from about 1 to 15 residues, more preferably about 1 to 10 residues and typically about 1 to 5 contiguous residues.

35 Substitution mutants have at least one amino acid residue in the polypeptide removed and a different residue inserted in its place. The sites of greatest interest for substitutional mutagenesis include sites identified as the active site(s). Other sites of interest are those in which particular residues obtained from various strains or species

are identical. These positions may be important for biological activity. These sites, especially those falling within a sequence of at least three other identically conserved sites, are preferably substituted in a relatively conservative manner. Such conservative substitutions are shown in Table 1 under the heading of "exemplary
5 substitutions".

In a preferred embodiment a mutant/variant polypeptide has only, or not more than, one or two or three or four conservative amino acid changes when compared to a naturally occurring polypeptide. Details of conservative amino acid changes are provided in Table 1. As the skilled person would be aware, such minor changes can
10 reasonably be predicted not to alter the activity of the polypeptide when expressed in a recombinant cell.

Table 1. Exemplary substitutions.

Original Residue	Exemplary Substitutions
Ala (A)	val; leu; ile; gly
Arg (R)	lys
Asn (N)	gln; his
Asp (D)	glu
Cys (C)	ser
Gln (Q)	asn; his
Glu (E)	asp
Gly (G)	pro, ala
His (H)	asn; gln
Ile (I)	leu; val; ala
Leu (L)	ile; val; met; ala; phe
Lys (K)	arg
Met (M)	leu; phe
Phe (F)	leu; val; ala
Pro (P)	gly
Ser (S)	thr
Thr (T)	ser
Trp (W)	tyr
Tyr (Y)	trp; phe
Val (V)	ile; leu; met; phe, ala

Directed Evolution

In directed evolution, random mutagenesis is applied to a protein, and a selection regime is used to pick out variants that have the desired qualities, for example, increased fatty acid acyltransferase activity. Further rounds of mutation and selection are then applied. A typical directed evolution strategy involves three steps:

1) *Diversification*: The gene encoding the protein of interest is mutated and/or recombined at random to create a large library of gene variants. Variant gene libraries can be constructed through error prone PCR (see, for example, Leung, 1989; Cadwell and Joyce, 1992), from pools of DNaseI digested fragments prepared from parental templates (Stemmer, 1994a; Stemmer, 1994b; Crameri et al., 1998; Coco et al., 2001) from degenerate oligonucleotides (Ness et al., 2002, Coco, 2002) or from mixtures of both, or even from undigested parental templates (Zhao et al., 1998; Eggert et al., 2005; Jézéque et al., 2008) and are usually assembled through PCR. Libraries can also be made from parental sequences recombined *in vivo* or *in vitro* by either homologous or non-homologous recombination (Ostermeier et al., 1999; Volkov et al., 1999; Sieber et al., 2001). Variant gene libraries can also be constructed by sub-cloning a gene of interest into a suitable vector, transforming the vector into a "mutator" strain such as the *E. coli* XL-1 red (Stratagene) and propagating the transformed bacteria for a suitable number of generations. Variant gene libraries can also be constructed by subjecting the gene of interest to DNA shuffling (i.e., *in vitro* homologous recombination of pools of selected mutant genes by random fragmentation and reassembly) as broadly described by Harayama (1998).

2) *Selection*: The library is tested for the presence of mutants (variants) possessing the desired property using a screen or selection. Screens enable the identification and isolation of high-performing mutants by hand, while selections automatically eliminate all nonfunctional mutants. A screen may involve screening for the presence of known conserved amino acid motifs. Alternatively, or in addition, a screen may involve expressing the mutated polynucleotide in a host organism or part thereof and assaying the level of fatty acid acyltransferase activity by, for example, quantifying the level of resultant product in lipid extracted from the organism or part thereof, and determining the level of product in the extracted lipid from the organism or part thereof relative to a corresponding organism or part thereof lacking the mutated polynucleotide and optionally, expressing the parent (unmutated) polynucleotide. Alternatively, the screen may involve feeding the organism or part thereof labelled substrate and determining the level of substrate or product in the organism or part thereof relative to a corresponding organism or part thereof lacking the mutated polynucleotide and optionally, expressing the parent (unmutated) polynucleotide.

3) *Amplification*: The variants identified in the selection or screen are replicated many fold, enabling researchers to sequence their DNA in order to understand what mutations have occurred.

Together, these three steps are termed a "round" of directed evolution. Most experiments will entail more than one round. In these experiments, the "winners" of the previous round are diversified in the next round to create a new library. At the end of the experiment, all evolved protein or polynucleotide mutants are characterized using biochemical methods.

10 *Rational Design*

A protein can be designed rationally, on the basis of known information about protein structure and folding. This can be accomplished by design from scratch (*de novo* design) or by redesign based on native scaffolds (see, for example, Hellinga, 1997; and Lu and Berry, Protein Structure Design and Engineering, Handbook of Proteins 2, 1153-1157 (2007)). Protein design typically involves identifying sequences that fold into a given or target structure and can be accomplished using computer models. Computational protein design algorithms search the sequence-conformation space for sequences that are low in energy when folded to the target structure. Computational protein design algorithms use models of protein energetics to evaluate how mutations would affect a protein's structure and function. These energy functions typically include a combination of molecular mechanics, statistical (i.e. knowledge-based), and other empirical terms. Suitable available software includes IPRO (Iterative Protein Redesign and Optimization), EGAD (A Genetic Algorithm for Protein Design), Rosetta Design, Sharpen, and Abalone.

Also included within the scope of the invention are polypeptides defined herein which are differentially modified during or after synthesis for example, by biotinylation, benzylation, glycosylation, acetylation, phosphorylation, amidation, derivatization by known protecting/blocking groups, proteolytic cleavage, linkage to an antibody molecule or other cellular ligand, etc. These modifications may serve to increase the stability and/or bioactivity of the polypeptide of the invention.

Identification of Fatty Acid Acyltransferases

In one aspect, the invention provides a method for identifying a nucleic acid molecule encoding a fatty acid acyltransferase having an increased ability to produce MAG, DAG and/or TAG in a cell.

The method comprises obtaining a cell comprising a nucleic acid molecule encoding a fatty acid acyltransferase operably linked to a promoter which is active in

the cell. The nucleic acid molecule may encode a naturally occurring fatty acid acyltransferase such as MGAT, GPAT and/or DGAT, or a mutant(s) thereof. Mutants may be engineered using standard procedures in the art (see above) such as by performing random mutagenesis, targeted mutagenesis, or saturation mutagenesis on known genes of interest, or by subjecting different genes to DNA shuffling. For example, a polynucleotide comprising a sequence selected from any one of SEQ ID NOs:1 to 44 which encodes a MGAT may be mutated and/or recombined at random to create a large library of gene variants (mutants) using for example, error-prone PCR and/or DNA shuffling. Mutants may be selected for further investigation on the basis that they comprise a conserved amino acid motif. For example, in the case of a candidate nucleic acid encoding a MGAT, a skilled person may determine whether it comprises a sequence as provided in SEQ ID NOs:220, 221, 222, 223, and/or 224 before testing whether the nucleic acid encodes a functional MGAT mutant (by for example, transfection into a host cell, such as a plant cell and assaying for fatty acid acyltransferase (i.e., MGAT) activity as described herein).

Direct PCR sequencing of the nucleic acid or a fragment thereof may be used to determine the exact nucleotide sequence and deduce the corresponding amino acid sequence and thereby identify conserved amino acid sequences. Degenerate primers based on conserved amino acid sequences may be used to direct PCR amplification. Degenerate primers can also be used as probes in DNA hybridization assays. Alternatively, the conserved amino acid sequence(s) may be detected in protein hybridization assays that utilize for example, an antibody that specifically binds to the conserved amino acid sequences(s), or a substrate that specifically binds to the conserved amino acid sequences(s) such as, for example, a lipid that binds FLXLXXN (a putative neutral lipid binding domain; SEQ ID NO:224).

In one embodiment, the nucleic acid molecule comprises a sequence of nucleotides encoding a MGAT. The sequence of nucleotides may i) comprise a sequence selected from any one of SEQ ID NOs:1 to 44, ii) encode a polypeptide comprising amino acids having a sequence as provided in any one of SEQ ID NOs:45 to 82, or a biologically active fragment thereof, iii) be at least 50% identical to i) or ii), or iv) hybridize to any one of i) to iii) under stringent conditions. In another or additional embodiment, the nucleic acid molecule comprises a sequence of nucleotides encoding one or more conserved DGAT2 and/or MGAT1/2 amino acid sequences as provided in SEQ ID NOs:220, 221, 222, 223, and 224. In a preferred embodiment, the nucleic acid molecule comprises a sequence of nucleotides encoding the conserved amino acid sequences provided in SEQ ID NO:220 and/or SEQ ID NO:224.

In another embodiment, the nucleic acid molecule comprises a sequence of nucleotides encoding a GPAT, preferably a GPAT which has phosphatase activity. The sequence of nucleotides may i) comprise a sequence selected from any one of SEQ ID NOs:84 to 141, ii) encode a polypeptide comprising amino acids having a sequence as provided in any one of SEQ ID NOs:144 to 201, or a biologically active fragment thereof, iii) be at least 50% identical to i) or ii), or iv) hybridize to any one of i) to iii) under stringent conditions. In another or additional embodiment, the nucleic acid molecule comprises a sequence of nucleotides encoding one or more conserved GPAT amino acid sequences as provided in SEQ ID NOs:225, 226, and 227, or a sequence of amino acids which is at least 50%, preferably at least 60%, more preferably at least 65% identical thereto.

In another embodiment, the nucleic acid molecule comprises a sequence of nucleotides encoding a DGAT2. The sequence of nucleotides may comprise i) a sequence of nucleotides selected from any one of SEQ ID NO:204 to 211, ii) encode a polypeptide comprising amino acids having a sequence as provided in any one of SEQ ID NO:212 to 219, or a biologically active fragment thereof, iii) be at least 50% identical to i) or ii), or iv) hybridize to any one of i) to iii) under stringent conditions. In a preferred embodiment, the DGAT2 comprises a sequence of nucleotides of SEQ ID NO:204 and/or a sequence of nucleotides encoding a polypeptide comprising amino acids having a sequence as provided in SEQ ID NO:212.

A cell comprising a nucleic acid molecule encoding a fatty acid acyltransferase operably linked to a promoter which is active in the cell may be obtained using standard procedures in the art such as by introducing the nucleic acid molecule into a cell by, for example, calcium phosphate precipitation, polyethylene glycol treatment, electroporation, and combinations of these treatments. Other methods of cell transformation can also be used and include, but are not limited to, the introduction of DNA into plants by direct DNA transfer or injection. Transformed plant cells may also be obtained using *Agrobacterium*-mediated transfer and acceleration methods as described herein.

The method further comprises determining if the level of MAG, DAG and/or TAG produced in the cell is increased when compared to a corresponding cell lacking the nucleic acid using known techniques in the art such as those exemplified in Example 1. For instance, lipids can be extracted in a chloroform/methanol solution, dried and separated by thin layer chromatography (TLC). Identities of TAG, DAG, MAG, free fatty acid, and other lipids can be verified with internal lipid standards after staining with iodine vapor. The resultant chromatograms can be analyzed using a PhosphorImager and the amount of MAG, DAG and TAG quantified on the basis of the known amount of internal standards, or alternatively, the cells may be fed *sn*-2

monooleoylglycerol[¹⁴C] or [¹⁴C]glycerol-3-phosphate and associated radioactivity quantitated by liquid scintillation counting (i.e., the amount of labelled MAG, DAG and TAG is quantified).

The method further comprises identifying a nucleic acid molecule encoding a fatty acid acyltransferase having an increased ability to produce MAG, DAG and/or TAG in a cell. In a preferred embodiment, the fatty acid acyltransferase catalyzes an enzyme reaction in the MGAT pathway. In a further preferred embodiment, DAG is increased via the MGAT pathway (i.e., acylation of MAG with fatty acyl-CoA is catalysed by a MGAT to form DAG). In another or additional embodiment, the substrate MAG is produced by a GPAT which also has phosphatase activity and/or DAG is acylated with fatty acyl-CoA by a DGAT and/or a MGAT having DGAT activity to form TAG.

Gloss

Certain aspects of the invention relate to measuring the glossiness of vegetative material as a marker for the level of lipid in the material, with higher glossiness levels being associated with higher lipid levels.

The gloss of the vegetative material can be determined using known procedures. Glossmeters (reflectometers) provide a quantifiable way of measuring gloss intensity ensuring consistency of measurement by defining the precise illumination and viewing conditions. The configuration of both illumination source and observation reception angles allows measurement over a small range of the overall reflection angle. The measurement results of a glossmeter are related to the amount of reflected light from a black glass standard with a defined refractive index. The ratio of reflected to incident light for the specimen, compared to the ratio for the gloss standard, is recorded as gloss units.

The measurement scale, Gloss Units (GU), of a glossmeter is a scaling based on a highly polished reference black glass standard with a defined refractive index having a specular reflectance of 100GU at the specified angle. This standard is used to establish an upper point calibration of 100 with the lower end point established at 0 on a perfectly matt surface. This scaling is suitable for most non-metallic materials.

The optimal or expected level of glossiness of vegetative material is likely to vary between plant species. The skilled person can readily analyse the lipid content of vegetative material of different plants of the invention and identify a suitable pre-determined level of glossiness that can be used as a standard in the field for assessing the best time to harvest a vegetative material from a particular plant species.

EXAMPLES

Example 1. General materials and methods

Expression of genes in plant cells in a transient expression system

Genes were expressed in plant cells using a transient expression system essentially as described by Voinnet et al. (2003) and Wood et al. (2009). Binary vectors containing the coding region to be expressed by a strong constitutive e35S promoter containing a duplicated enhancer region were introduced into *Agrobacterium tumefaciens* strain AGL1. A chimeric binary vector, 35S:p19, for expression of the p19 viral silencing suppressor was separately introduced into AGL1, as described in WO2010/057246. A chimeric binary vector, 35S:V2, for expression of the V2 viral silencing suppressor was separately introduced into AGL1. The recombinant cells were grown to stationary phase at 28°C in LB broth supplemented with 50 mg/L kanamycin and 50 mg/L rifampicin. The bacteria were then pelleted by centrifugation at 5000 g for 5 min at room temperature before being resuspended to OD600 = 1.0 in an infiltration buffer containing 10 mM MES pH 5.7, 10 mM MgCl₂ and 100 µM acetosyringone. The cells were then incubated at 28°C with shaking for 3 hours after which the OD600 was measured and a volume of each culture, including the viral suppressor construct 35S:p19 or 35S:V2, required to reach a final concentration of OD600 = 0.125 added to a fresh tube. The final volume was made up with the above buffer. Leaves were then infiltrated with the culture mixture and the plants were typically grown for a further three to five days after infiltration before leaf discs were recovered for either purified cell lysate preparation or total lipid isolation.

Purified leaf lysate assay

Nicotiana benthamiana leaf tissues previously infiltrated as described above were ground in a solution containing 0.1 M potassium phosphate buffer (pH 7.2) and 0.33 M sucrose using a glass homogenizer. Leaf homogenate was centrifuged at 20,000 g for 45 minutes at 4°C after which each supernatant was collected. Protein content in each supernatant was measured according to Bradford (1976) using a Wallac1420 multi-label counter and a Bio-Rad Protein Assay dye reagent (Bio-Rad Laboratories, Hercules, CA USA). Acyltransferase assays used 100 µg protein according to Cao et al. (2007) with some modifications. The reaction medium contained 100 mM Tris-HCl (pH 7.0), 5 mM MgCl₂, 1 mg/mL BSA (fatty acid-free), 200 mM sucrose, 40 mM cold oleoyl-CoA, 16.4 µM *sn*-2 monooleoylglycerol [¹⁴C] (55mCi/mmol, American Radiochemicals, Saint Louis, MO USA) or 6.0 µM [¹⁴C]glycerol-3-phosphate (G-3-P) disodium salt (150 mCi/mmol, American Radiochemicals). The assays were carried out for 7.5, 15, or 30 minutes.

Lipid analysis

In summary, the methods used for analysing lipids in seeds or vegetative tissues were as follows:

5 Arabidopsis seed and any other similar sized seed:

(i) Fatty acid composition -direct methylation of fatty acids in seeds, without crushing of seeds.

(ii) Total fatty acid or TAG quantitation - direct methylation of fatty acids in seeds, without crushing of seeds, with use of a 17:0 TAG standard.

10

Canola seed, Camelina seed, and any other larger sized seeds:

(i) Single seed fatty acid composition - direct methylation of fatty acids in seed after breaking seed coat.

15 (ii) Pooled seed-fatty acid composition of total extracted lipid - crushing seeds in CHCl₃/ MeOH and methylation of aliquots of the extracted lipid.

(iii) Pooled seed-total lipid content (seed oil content) - two times lipid extraction for complete recovery of seed lipids after crushing seeds from known amount of dessicated seeds, with methylation of lipids from known amount of seeds together with 17:0 fatty acids as internal standard.

20 (iv) Pooled seed-purified TAG quantitation - two times lipid extraction for complete recovery of seed lipids after crushing seeds, from known amount of dessicated seeds, TAG fractionation from the lipid using TLC, and direct methylation of TAG *in silica* using 17:0 TAG as internal standard.

25 Leaf samples:

(i) Fatty acid composition of total lipid - direct methylation of fatty acids in freeze-dried samples.

(ii) Total lipid quantitation - direct methylation of fatty acids in known weight of freeze-dried samples, with 17:0 FFA.

30 (iii) TAG quantitation - because of the presence of substantial amounts of polar lipids in leaves, TAG was fractionated by TLC from extracted total lipids, and methylated in the presence of 17:0 TAG internal standard. Steps: Freeze dry samples, weighing, lipid extraction, fractionation of TAG from known amount of total lipids, direct methylation of TAG *in silica* together with 17:0 TAG as internal standard.

35 The methods are detailed as follows:

Analysis of oil content in Arabidopsis seeds

Where seed oil content was to be determined in small seeds such as *Arabidopsis* seeds, seeds were dried in a desiccator for 24 hours and approximately 4 mg of seed was transferred to a 2 ml glass vial containing Teflon-lined screw cap. 0.05 mg triheptadecanoin dissolved in 0.1 ml toluene was added to the vial as internal standard. Seed FAME were prepared by adding 0.7 ml of 1N methanolic HCl (Supelco) to the vial containing seed material. Crushing of the seeds was not necessary with small seeds such as *Arabidopsis* seeds. The mixture was vortexed briefly and incubated at 80°C for 2 hours. After cooling to room temperature, 0.3 ml of 0.9% NaCl (w/v) and 0.1 ml hexane was added to the vial and mixed well for 10 minutes in a Heidolph Vibramax 110. The FAME was collected into a 0.3 ml glass insert and analysed by GC with a flame ionization detector (FID) as mentioned earlier.

The peak area of individual FAME were first corrected on the basis of the peak area responses of a known amount of the same FAMES present in a commercial standard GLC-411 (NU-CHEK PREP, INC., USA). GLC-411 contains equal amounts of 31 fatty acids (% by weight), ranging from C8:0 to C22:6. In case of fatty acids which were not present in the standard, the peak area responses of the most similar FAME was taken. For example, the peak area response of FAMES of 16:1d9 was used for 16:1d7 and the FAME response of C22:6 was used for C22:5. The corrected areas were used to calculate the mass of each FAME in the sample by comparison to the internal standard mass. Oil is stored mainly in the form of TAG and its weight was calculated based on FAME weight. Total moles of glycerol was determined by calculating moles of each FAME and dividing total moles of FAMES by three. TAG was calculated as the sum of glycerol and fatty acyl moieties using a relation: % oil by weight = $100 \times ((41 \times \text{total mol FAME} / 3) + (\text{total g FAME} - (15 \times \text{total mol FAME}))) / \text{g seed}$, where 41 and 15 are molecular weights of glycerol moiety and methyl group, respectively.

Analysis of oil content in Camelina seeds and canola seeds by extraction

After harvest at plant maturity, *Camelina* or canola seeds were dessicated by storing the seeds for 24 hours at room temperature in a dessicator containing silica gel as dessicant. Moisture content of the seeds is typically 6-8%. Total lipids were extracted from known weights of the dessicated seeds by crushing the seeds using a mixture of chloroform and methanol (2/1 v/v) in an eppendorf tube using a Reicht tissue lyser (22 frequency/seconds for 3 minutes) and a metal ball. One volume of 0.1M KCl was added and the mixture shaken for 10 minutes. The lower non-polar phase was collected after centrifuging the mixture for 5 minutes at 3000 rpm. The

remaining upper (aqueous) phase was washed with 2 volumes of chloroform by mixing for 10 minutes. The second non-polar phase was also collected and pooled with the first. The solvent was evaporated from the lipids in the extract under nitrogen flow and the total dried lipid was dissolved in a known volume of chloroform.

To measure the amount of lipid in the extracted material, a known amount of 17:0-TAG was added as internal standard and the lipids from the known amount of seeds incubated in 1 N methanolic-HCl (Supelco) for 2 hours at 80°C. FAME thus made were extracted in hexane and analysed by GC. Individual FAMES were quantified on the basis of the amount of 17:0 TAG-FAME. Individual FAMES weights, after subtraction of weights of the esterified methyl groups from FAME, were converted into moles by dividing by molecular weights of individual FAMES. Total moles of all FAMES were divided by three to calculate moles of TAG and therefore glycerol. Then, moles of TAG were converted in to weight of TAG. Finally, the percentage oil content on a seed weight basis was calculated using seed weights, assuming that all of the extracted lipid is TAG or equivalent to TAG for the purpose of calculating oil content. This method was based on Li et al., (2006). Seeds other than Camelina or canola seeds that are of a similar size can also be analysed by this method.

Canola and other seed oil content can also be measured by nuclear magnetic resonance techniques (Rossell and Pritchard, 1991), for example, by a pulsed wave NMS 100 Minispec (Bruker Pty Ltd Scientific Instruments, Germany), or by near infrared reflectance spectroscopy such as using a NIRSystems Model 5000 monochromator. The NMR method can simultaneously measure moisture content. Moisture content can also be measured on a sample from a batch of seeds by drying the seeds in the sample for 18 hours at about 100°C, according to Li et al., (2006).

Where fatty acid composition is to be determined for the oil in canola seed, the direct methylation method used for *Arabidopsis* seed (above) can be used, modified with the addition of cracking of the canola seedcoat. This method extracts sufficient oil from the seed to allow fatty acid composition analysis.

Analysis of lipids from leaf lysate assays

Lipids from the lysate assays were extracted using chloroform:methanol:0.1 M KCl (2:1:1) and recovered. The different lipid classes in the samples were separated on Silica gel 60 thin layer chromatography (TLC) plates (MERCK, Darmstadt, Germany) impregnated with 10% boric acid. The solvent system used to fractionate TAG from the lipid extract consisted of chloroform/acetone (90/10 v/v). Individual lipid classes were visualized by exposing the plates to iodine vapour and identified by

running parallel authentic standards on the same TLC plate. The plates were exposed to phosphor imaging screens overnight and analysed by a Fujifilm FLA-5000 phosphorimager before liquid scintillation counting for DPM quantification.

5 Total lipid isolation and fractionation

Tissues including leaf samples were freeze-dried, weighed (dry weight) and total lipids extracted as described by Bligh and Dyer (1959) or by using chloroform:methanol:0.1 M KCl (CMK; 2:1:1) as a solvent. Total lipids were extracted from *N. benthamiana* leaf samples, after freeze drying, by adding 900 µL of a chloroform/methanol (2/1 v/v) mixture per 1 cm diameter leaf sample. 0.8 µg DAGE was added per 0.5 mg dry leaf weight as internal standard when TLC-FID analysis was to be performed. Samples were homogenized using an IKA ultra-turrax tissue lyser after which 500 µL 0.1 M KCl was added. Samples were vortexed, centrifuged for 5 min and the lower phase was collected. The remaining upper phase was extracted a second time by adding 600 µL chloroform, vortexing and centrifuging for 5 min. The lower phase was recovered and pooled into the previous collection. Lipids were dried under a nitrogen flow and resuspended in 2 µL chloroform per mg leaf dry weight. Total lipids of *N. tabacum* leaves or leaf samples were extracted as above with some modifications. If 4 or 6 leaf discs (each approx 1 cm² surface area) were combined, 1.6 ml of CMK solvent was used, whereas if 3 or less leaf discs were combined, 1.2 ml CMK was used. Freeze dried leaf tissues were homogenized in an eppendorf tube containing a metallic ball using a Reicht tissue lyser (Qiagen) for 3 minutes at 20 frequency/sec.

25 Separation of neutral lipids via TLC and transmethylation

Known volumes of total leaf extracts such as, for example, 30 µL, were loaded on a TLC silica gel 60 plate (1x20 cm) (Merck KGaA, Germany). The neutral lipids were separated via TLC in an equilibrated development tank containing a hexane/DEE/acetic acid (70/30/1 v/v/v) solvent system. The TAG bands were visualised by iodine vapour, scraped from the TLC plate, transferred to 2 mL GC vials and dried with N₂. 750 µL of 1N methanolic-HCl (Supelco analytical, USA) was added to each vial together with a known amount of C17:0 TAG, such as, for example, 30 µg, as internal standard for quantification.

When analysing the effect on oleic acid levels of specific gene combinations, TAG and polar lipids bands were collected from the TLC plates. Next, 15 µg of C17:0 internal standard was added to samples such as TAG samples, polar lipid samples and 20 µL of the total lipid extracts. Following drying under N₂, 70 µL toluene and 700 µL methanolic HCl were added.

Lipid samples for fatty acid composition analysis by GC were transmethylated by incubating the mixtures at 80°C for 2 hours in the presence of the methanolic-HCl. After cooling samples to room temperature, the reaction was stopped by adding 350 µl H₂O. Fatty acyl methyl esters (FAME) were extracted from the mixture by adding 5 350 µl hexane, vortexing and centrifugation at 1700 rpm for 5 min. The upper hexane phase was collected and transferred into GC vials with 300 µl conical inserts. After evaporation, the samples were resuspended in 30 µl hexane. One µl was injected into the GC.

The amount of individual and total fatty acids (TFA) present in the lipid 10 fractions was quantified by GC by determining the area under each peak and calculated by comparison with the peak area for the known amount of internal standard. TAG content in leaf was calculated as the sum of glycerol and fatty acyl moieties in the TAG fraction using a relation: % TAG by weigh = $100 \times ((41 \times \text{total mol FAME}/3) + (\text{total g FAME} - (15 \times \text{total mol FAME}))) / \text{g leaf dry weight}$, where 41 15 and 15 are molecular weights of glycerol moiety and methyl group, respectively.

Capillary gas-liquid chromatography (GC)

FAME were analysed by GC using an Agilent Technologies 7890A GC (Palo Alto, California, USA) equipped with an SGE BPX70 (70% cyanopropyl 20 polysilphenylene-siloxane) column (30 m x 0.25 mm i.d., 0.25 µm film thickness), an FID, a split/splitless injector and an Agilent Technologies 7693 Series auto sampler and injector. Helium was used as the carrier gas. Samples were injected in split mode (50:1 ratio) at an oven temperature of 150°C. After injection, the oven temperature was held at 150°C for 1 min, then raised to 210°C at 3°C.min⁻¹ and finally to 240°C 25 at 50°C.min⁻¹. Peaks were quantified with Agilent Technologies ChemStation software (Rev B.04.03 (16), Palo Alto, California, USA) based on the response of the known amount of the external standard GLC-411 (Nuccheck) and C17:0-Me internal standard.

30 Quantification of TAG via Iatroscan

One µL of lipid extract was loaded on one Chromarod-SII for TLC-FID IatroscanTM (Mitsubishi Chemical Medience Corporation – Japan). The Chromarod rack was then transferred into an equilibrated developing tank containing 70 mL of a hexane/CHCl₃/2-propanol/formic acid (85/10.716/0.567/0.0567 v/v/v/v) solvent 35 system. After 30 min of incubation, the Chromarod rack was dried for 3 min at 100°C and immediately scanned on an Iatroscan MK-6s TLC-FID analyser (Mitsubishi Chemical Medience Corporation – Japan). Peak areas of DAGE internal standard and

TAG were integrated using SIC-480II integration software (Version:7.0-E SIC System instruments Co., LTD – Japan).

TAG quantification was carried out in two steps. First, DAGE was scanned in all samples to correct the extraction yields after which concentrated TAG samples
5 were selected and diluted. Next, TAG was quantified in diluted samples with a second scan according to the external calibration using glyceryl trilinoleate as external standard (Sigma-Aldrich).

Quantification of TAG in leaf samples by GC

10 The peak area of individual FAME were first corrected on the basis of the peak area responses of known amounts of the same FAMES present in a commercial standard GLC-411 (NU-CHEK PREP, Inc., USA). The corrected areas were used to calculate the mass of each FAME in the sample by comparison to the internal standard. Since oil is stored primarily in the form of TAG, the amount of oil was
15 calculated based on the amount of FAME in each sample. Total moles of glycerol were determined by calculating the number of moles of FAMES and dividing total moles of FAMES by three. The amount of TAG was calculated as the sum of glycerol and fatty acyl moieties using the formula: % oil by weight = $100 \times ((41 \times \text{total mol FAME} / 3) + (\text{total g FAME} - (15 \times \text{total mol FAME}))) / \text{g leaf dry weight}$, where 41 and 15
20 were the molecular weights of glycerol moiety and methyl group, respectively.

DGAT assay in *Saccharomyces cerevisiae* H1246

Saccharomyces cerevisiae strain H1246 is completely devoid of DGAT activity and lacks TAG and sterol esters as a result of knockout mutations in four
25 genes (DGA1, LRO1, ARE1, ARE2). The addition of free fatty acid (e.g. 1 mM 18:1^{Δ9}) to H1246 growth media is toxic in the absence of DGAT activity. Growth on such media can therefore be used as an indicator or selection for the presence of DGAT activity in this yeast strain.

S. cerevisiae H1246 was transformed with the pYES2 construct (negative
30 control), a construct encoding *Arabidopsis thaliana* DGAT1 in pYES2, or a construct encoding *Mus musculus* MGAT2 in pYES2. Transformants were fed [¹⁴C]18:1^{Δ9} free fatty acids.

In a separate experiment, *S. cerevisiae* H1246 was transformed with the pYES2 construct (negative control), a construct encoding *Bernardia pulchella* DGAT1 in
35 pYES2, or a construct encoding *M. musculus* MGAT1 in pYES2 and fed 18:1^{Δ9} free fatty acids. *S. cerevisiae* S288C wild type strain transformed with pYES2 served as a positive control.

Yeast transformants were resuspended in sterile mQ water and diluted to OD₆₀₀=1. Samples were further diluted in four consecutive dilutions, each at 1/10. 2 µl of each dilution was spotted on each of the plates (YNBD, YNBG, YNBG+FA) together with 2 µL mQ water and 2 µL of an untransformed H1246 cell suspension (OD₆₀₀=1). Plates were incubated for 6 days at 30°C before scoring growth.

Plate medium, 40 mL media per plate

- YNBD: minimal dropout medium lacking uracil and containing 2% glucose, 0.01% NP40 and 100 µL ethanol.
- 10 • YNBG: minimal dropout medium lacking uracil and containing 2% galactose, 1% raffinose, 0.01% NP40 and 100 µL ethanol.
- YNBG+FA: minimal dropout medium lacking uracil and containing 2% galactose, 1% raffinose, 0.01% NP40 and 1mM C18:1^{Δ9} dissolved in 100 µl ethanol.

15 **Example 2. Constitutive expression of a monoacylglycerol acyltransferase in plant cells**

MGAT1

The enzyme activity of the monoacylglycerol acyltransferase 1 (MGAT1) encoded by the gene from *M. musculus* (Yen et al., 2002) and *A. thaliana* diacylglycerol acyltransferase (DGAT1) (Bouvier-Nave et al., 2000), used here as a comparison with MGAT1, were demonstrated in *N. benthamiana* leaf tissue using a transient expression system as described in Example 1.

A vector designated 35S-pORE04 was made by inserting a *Pst*I fragment containing a 35S promoter into the *Sfo*I site of vector pORE04 after T4 DNA polymerase treatment to blunt the ends (Coutu et al., 2007). A chimeric DNA encoding the *M. musculus* MGAT1, codon-optimised for *Brassica napus*, was synthesized by Geneart and designated 0954364_MGAT_pMA. A chimeric DNA designated 35S:MGAT1 and encoding the *M. musculus* MGAT1 (Genbank Accession No. Q91ZV4) for expression in plant cells was made by inserting the entire coding region of 0954364_MGAT_pMA, contained within an *Eco*RI fragment, into 35S-pORE04 at the *Eco*RI site. The vector containing the 35S:MGAT1 construct was designated as pJP3184. Similarly, a chimeric DNA 35S:DGAT1 encoding the *A. thaliana* DGAT1 (Genbank Accession No. AAF19262) for expression in plant cells was made by inserting the entire coding region of pXZP513E, contained within a *Bam*HI-*Eco*RV fragment, into 35S-pORE04 at the *Bam*HI-*Eco*RV site. The vector containing the 35S:DGAT1 construct was designated pJP2078.

The chimeric vectors were introduced into *A. tumefaciens* strain AGL1 and cells from cultures of these infiltrated into leaf tissue of *N. benthamiana* plants in a

24°C growth room. In order to allow direct comparisons between samples and to reduce inter-leaf variation, samples being compared were infiltrated on either side of the same leaf. Experiments were performed in triplicate. Following infiltration, the plants were grown for a further three days before leaf discs were taken, freeze-dried, and lipids extracted from the samples were fractionated and quantified as described in Example 1. This analysis revealed that the MGAT1 and DGAT1 genes were functioning to increase leaf oil levels in *N. benthamiana* as follows.

Leaf tissue transformed with the 35S:p19 construct only (negative control) contained an average of 4 µg free fatty acid (FFA) derived from DAG/mg dry leaf weight and 5 µg FFA derived from TAG/mg dry leaf weight. Leaf tissue transformed with the 35S:p19 and 35S:DGAT1 constructs (control for expression of DGAT1) contained an average of 4 µg FFA derived from DAG/mg dry leaf weight and 22 µg FFA derived from TAG/mg dry leaf weight. Leaf tissue transformed with the 35S:p19 and 35S:MGAT1 constructs contained an average of 8 µg FFA derived from DAG/mg dry leaf weight and 44 µg FFA derived from TAG/mg dry leaf weight. Leaf tissue transformed with the 35S:p19, 35S:DGAT1 and 35S:MGAT1 constructs did not contain DAG or TAG levels higher than those observed in the 35S:p19 and 35S:MGAT1 infiltration (Figure 2). Also, a decrease in the level of saturates in seeds was noted after MGAT expression when compared with either the p19 control or DGAT1 samples (Table 2).

The data described above demonstrated that the MGAT1 enzyme was far more active than the DGAT1 enzyme in promoting both DAG and TAG accumulation in leaf tissue. Expression of the MGAT1 gene resulted in twice as much TAG and DAG accumulation in leaf tissue compared to when the DGAT1 was expressed. This result was highly surprising and unexpected, considering that the MGAT is an enzyme expressed in mouse intestine, a vastly different biological system than plant leaves. This study was the first demonstration of ectopic MGAT expression in a plant cell.

Leaf samples infiltrated with *M. musculus* MGAT1 accumulated double the DAG and TAG relative to leaf tissue infiltrated with *A. thaliana* DGAT1 alone. The efficiency of the production of TAG was also surprising and unexpected given that the mouse MGAT has only very low activity as a DGAT. Leaf tissue infiltrated with genes encoding both MGAT1 and DGAT1 did not accumulate significantly more TAG than the MGAT1-only leaf sample. Figure 1 is a representation of various TAG accumulation pathways, most of which converge at DAG, a central molecule in lipid synthesis. For instance, MAG, DAG and TAG can be inter-converted via various enzyme activities including transacylation, lipase, MGAT, DGAT and PDAT. A decrease in the level of saturates was also noted after MGAT expression.

MGAT2

A chimeric DNA designated 35S:MGAT2 and encoding the *M. musculus* MGAT2 for expression in plant cells was made by inserting the entire MGAT2 coding region, contained within an *EcoRI* fragment, into 35S-pORE04 at the *EcoRI* site. The enzyme activity of the monoacylglycerol acyltransferase 2 (MGAT2) encoded by the gene from *M. musculus* (Yen, 2003) (Genbank Accession No. Q80W94) and *A. thaliana* DGAT1 (Bouvier-Nave et al., 2000), used here as a comparison with MGAT2, was also demonstrated in *N. benthamiana* leaf tissue using a transient expression system as described in Example 1.

Compared with controls, DGAT1 expression increased leaf TAG 5.9-fold, MGAT2 by 7.3-fold and the combination of MGAT2+DGAT1 by 9.8-fold (Figure 3). The ability of MGAT2 alone to yield such significant increases in TAG was unexpected for a number of reasons. Firstly, the amount of substrate MAG present in leaf tissue is known to be low and large increases in TAG accumulation from this substrate would not be expected. Secondly, the addition of MGAT activity alone (i.e., addition of MGAT2 which does not have DGAT activity) would be expected to yield DAG, not TAG, especially in a leaf environment where little native DGAT activity is usually present.

Discussion

The present inventors have surprisingly demonstrated that the transgenic expression of a MGAT gene results in significant increases in lipid yield in plant cells. The present inventors understand that Tumaney et al. had isolated a DGAT with some MGAT activity and that they were not successful in attempts to clone a gene encoding a MGAT as defined herein. Tumaney et al. (2001) reported MGAT activity in peanut and isolated an enzyme responsible for this activity. However, Tumaney et al. did not publish results of tests for DGAT activity and it therefore seems that the enzyme reported was a DGAT with some MGAT activity. Indeed, previous work had failed to identify any MGAT activity in other species (Stobart et al., 1997). Furthermore, it was surprising that the enzyme isolated by Tumaney et al. was a soluble, cytosolic, enzyme rather than a membrane-bound enzyme.

Recently, researchers have identified a microsomal membrane-bound monoacylglycerol acyltransferase (MGAT) from immature peanut (*Arachis hypogaea*) seeds. The MGAT could be solubilized from microsomal membranes using a combination of a chaotropic agent and a zwitterionic detergent, and a functionally active 14S multiprotein complex was isolated and characterized. Oleosin3 (OLE3) was identified as part of the multiprotein complex, which is capable

of performing bifunctional activities such as acylating monoacylglycerol (MAG) to diacylglycerol (DAG) and phospholipase A2 (PLA2; Parthibane et al., 2012).

Example 3. Biochemical demonstration of transgenic MGAT activity in leaf extracts

Cell lysates were made from *N. benthamiana* leaf tissue that had been infiltrated with 35S:MGAT1, 35S:MGAT2 and 35S:DGAT1, as described in Example 1. Separate leaf infiltrations were performed, each in triplicate, for strains containing the 35S:p19 construct only (negative control), the 35S:MGAT2 strain together with the 35S:p19 strain, and a mixture of the 35S:MGAT2 and 35S:DGAT1 *Agrobacterium* strains with the 35S:p19 strain. The triplicate samples were harvested after three days and a purified cell lysate prepared by mechanical tissue lysis and centrifugation. The MGAT activities of the purified cell lysates were compared by feeding [¹⁴C]MAG to the lysates as described in Example 1. The data are shown in Figure 4.

Little MGAT activity was observed in the 35S:p19 control sample, since most of the radioactivity remained in MAG throughout the assay. In contrast, the majority of the labelled MAG in the 35S:MGAT2 sample was rapidly converted to DAG (Figure 4, central panel), indicating strong MGAT activity expressed from the 35S:MGAT2 construct. Furthermore, a significant amount of TAG was also produced. The TAG production observed in the 35S:MGAT2 sample was likely due to native *N. benthamiana* DGAT activity, or produced by another TAG synthesis route. The amount of TAG production was greatly increased by the further addition of 35S:DGAT1 (Figure 4, right hand panel), indicating that the MGAT2 enzyme produced DAG which was accessible for conversion to TAG by DGAT1 in plant vegetative tissues.

Example 4. Biochemical demonstration of the production of MGAT-accessible MAG in leaf extracts

In the *in vitro* assays described in Example 3 using leaf lysates, the substrates (*sn*-2 MAG and oleoyl-CoA) were exogenously supplied, whereas *in vivo* MGAT activity in intact plant tissues would require the native presence of these substrates. The presence of low levels of MAG in various plant tissues has been reported previously (Hirayama and Hujii, 1965; Panekina et al., 1978; Lakshminarayana et al., 1984; Perry & Harwood, 1993). To test whether the MGAT2 could access MAG produced by native plant pathways, the above experiment was repeated but this time feeding [¹⁴C]G-3-P to the lysates. The resultant data are shown schematically in Figure 5.

The production of labelled MAG was observed in all samples, indicating the *de novo* production of MAG from the G-3-P in plant leaf lysates. Labelled DAG and TAG products were also observed in all samples although these were relatively low in the 35S:p19 control sample, indicating that the production of these neutral lipids by the endogenous Kennedy pathway was relatively low in this sample. In contrast, the majority of the label in the MGAT2 and MGAT2 + DGAT1 samples appeared in the DAG and TAG pools, indicating that the exogenously added MGAT catalysed conversion of the MAG that had been produced from the labelled G-3-P by a native plant pathway.

- Examples 2 to 4 demonstrate several key points: 1) Leaf tissue can synthesise MAG from G-3-P such that the MAG is accessible to an exogenous MGAT expressed in the leaf tissue; 2) Even an MGAT which is derived from mammalian intestine can function in plant tissues, not known to possess an endogenous MGAT, requiring a successful interaction with other plant factors involved in lipid synthesis; 3) DAG produced by the exogenous MGAT activity is accessible to a plant DGAT, or an exogenous DGAT, to produce TAG; and 4) the expression of an exogenous MGAT can yield greatly increased TAG levels in plant tissues, levels which are at least as great as that yielded by exogenous *A. thaliana* DGAT1 expression.

Example 5. Expression of DGAT1, MGAT1 and MGAT2 in yeast

- Chimeric yeast expression vectors were constructed by inserting genes encoding the *A. thaliana* DGAT1, *M. musculus* MGAT1 and *M. musculus* MGAT2 into the pYES2 vector to yield pYES2:DGAT1, pYES2:MGAT1 and pYES2:MGAT2. These constructs were transformed in *Saccharomyces cerevisiae* strain H1246 which is completely devoid of DGAT activity and lacks TAG and sterol esters as a result of knockout mutations in four genes (*DGA1*, *LRO1*, *ARE1*, *ARE2*). Yeast strain H1246 is capable of synthesizing DAG from exogenously added fatty acids, but is unable to convert the DAG to TAG because of the knockout mutations. The transformed yeast cultures were fed [¹⁴C]18:1^{Δ9} before total lipids were extracted and fractionated by TLC as described in Example 1. An autoradiogram of a representative TLC plate is shown in Figure 6.

- TAG formation, indicating the presence of DGAT activity, was observed for the yeast cells containing either DGAT1 (positive control) and the mammalian MGAT1, but not in cells containing the MGAT2 encoded by the native *M. musculus* coding region. It was concluded that MGAT1 from mouse also had DGAT activity in yeast cells, and therefore functioned as a dual function MGAT/DGAT enzyme, whereas MGAT2 did not have detectable DGAT activity and was therefore solely an MGAT. A construct which included an MGAT2 coding region which was codon

optimization for expression in yeast exhibited MGAT activity (production of DAG) when tested *in vitro* using yeast microsomes and labelled MAG substrate, whereas a similar construct which was codon-optimised for expression in *B. napus* did not show DAG production in the yeast microsomes. This experiment showed the benefit of
5 codon-optimisation for the organism in which heterologous coding regions were to be expressed.

Example 6. Expression of a monoacylglycerol acyltransferase in plants, seeds
10 **and fungi**

Expression of MGAT1 in *Arabidopsis thaliana* seeds

A gene encoding *M. musculus* MGAT1 and under the control of a seed-specific promoter (FP1, a truncated *Brassica napus* napin promoter) was used to generate stably transformed *A. thaliana* plants and progeny seeds. The vector designated
15 pJP3174 was made by inserting a *Sall* fragment containing an *EcoRI* site flanked by the FP1 promoter and *Glycine max* lectin polyadenylation signal into the *Sall*-*XhoI* site of vector pCW141. The pCW141 vector also contained an FP1-driven, intron-interrupted, seed-secreted GFP as a screenable marker gene. The chimeric gene designated FP1:MGAT1-GFP was made by inserting the entire coding region of the
20 construct 0954364_MGAT_pMA, contained within an *EcoRI* fragment, into pJP3174 at the *EcoRI* site, generating pJP3179. This chimeric vector was introduced into *A. tumefaciens* strain AGL1 and cells from culture of the transformed *Agrobacterium* used to treat *A. thaliana* (ecotype Columbia) plants using the floral dip method for transformation (Clough and Bent, 1998). After maturation, the seeds from the treated
25 plants were viewed under a Leica MZFLIII dissection microscope and ebq 100 mercury lamp. Fifteen transgenic seeds (strongly GFP positive) and fifteen non-transgenic (GFP negative) seeds were isolated and each set pooled. The GFP positive and GFP negative pools were analysed for total fatty acid content as described in Example 1. This analysis provided the average fatty acid content and composition for
30 seeds transformed with the MGAT construct, but in a population which may have contained both hemizygous and homozygous transformed seeds.

This analysis revealed that the MGAT1 gene was functioning to increase seed oil levels in *A. thaliana* seed with the fifteen non-transgenic seeds (control, the same as wild-type) containing an average of 69.4 µg total fatty acids while the fifteen
35 transgenic seeds transformed with the GFP gene, and therefore likely to contain the FP1:MGAT1 genetic construct, contained an average of 71.9 µg total fatty acids. This was an increase of 3.5% in the oil content relative to the control (wild-type). The analysis also revealed that the MGAT gene was functioning to enrich polyunsaturated

fatty acids in the seed, as seen from the fatty acid composition of the total extracted lipid obtained from the seeds. In particular, the amount of ALA present as a percentage of the total fatty acid extracted from the seeds increasing from 16.0 to 19.6%. Similarly, the percentage of the fatty acid 20:2n6 increased from 1.25% to 1.90% and the fatty acid 20:3n3 increased from 0.26% to 0.51% (Table 2).

Table 2. Effect of MGAT expression on seed fatty acid composition.

FA profile (% of TFA)									
	16:00	16:01	16:3w3	C18:0	C18:1d9	C18:1d11	C18:2	C18:3	
Sample									
Control	7.41	0.36	0.12	3.00	15.26	1.98	30.93	15.98	
MGAT1	7.11	0.32	0.11	2.95	13.86	1.51	28.87	19.59	
	C20:0	20:1d11	20:1iso	20:2n6	20:3n3	C22:0	C22:1	C24:0	24:1d15
Sample									Total
Control	1.86	17.95	1.74	1.25	0.26	0.57	0.98	0.20	0.17
MGAT1	1.90	17.22	1.71	1.90	0.51	0.57	1.52	0.19	0.17

A further experiment was performed where the FP1:MGAT1-GFP chimeric DNA was modified to remove the GFP gene. This genetic construct, designated FP1:MGAT1, was transformed into an *A. thaliana* line which was mutant for FAD2.

- 15 The total fatty acid content of the T₂ seed from antibiotic resistant T₁ plants, as well as parental lines grown alongside these plants, was determined according to Example 1. The data is shown in Table 3. The average total fatty acids of the seed from the control lines was 347.9 µg/100 seeds whereas the average of the transgenic seeds was 381.0 µg/100 seeds. When the data for the control line C6 was excluded for
- 20 determining the average, the average for the controls was 370 µg/100 seeds. The oil content in the transgenic seeds represented an increase of about 3% in relative terms compared to the oil content in the untransformed seeds.

Table 3. *Arabidopsis thaliana* T₂ FP1:MGAT1 transgenic and parental control seed fatty acid profiles and total fatty acid quantification.

Sample	C16:0	C16:1	C18:0	C18:1	C18:1d11	C18:2	C18:3	C20:0	20:1d11	C22:0	C24:0	24:1d15	µg FA/100 seeds
C7	6.2	0.5	2.5	81.3	4.2	0.6	2.7	0.7	0.7	0.3	0.2	0.1	442.5
C4	6.3	0.4	2.4	81.7	3.9	0.5	2.5	0.9	0.6	0.4	0.2	0.1	403.8
C8	6.4	0.5	2.6	81.1	4.1	0.6	2.6	0.8	0.6	0.4	0.2	0.1	403.2
C2	6.2	0.5	2.4	81.4	4.1	0.6	2.7	0.8	0.6	0.4	0.2	0.1	377.0
C1	6.4	0.5	2.4	80.6	4.1	0.7	3.3	0.8	0.6	0.4	0.2	0.1	344.8
C3	6.4	0.5	2.6	80.0	4.1	0.6	3.5	0.8	0.6	0.4	0.2	0.2	314.3
C5	6.3	0.5	2.6	80.7	4.4	0.6	2.4	0.7	0.6	0.9	0.2	0.1	310.6
C6	6.7	0.7	2.7	77.2	5.0	0.8	4.3	0.9	0.7	0.4	0.3	0.2	186.8
M23	5.9	0.4	2.0	81.4	5.0	0.8	2.4	0.7	0.7	0.5	0.2	0.2	455.7
M10	6.0	0.4	2.4	82.3	4.2	0.7	2.2	0.7	0.6	0.3	0.2	0.1	437.7
M22	5.9	0.4	2.2	81.4	4.8	0.8	2.4	0.7	0.6	0.4	0.2	0.2	425.0
M25	6.0	0.4	2.2	81.7	4.6	0.7	2.4	0.8	0.6	0.3	0.2	0.1	406.7
M8	6.0	0.4	2.2	81.6	4.5	0.8	2.5	0.7	0.6	0.3	0.2	0.2	404.5
M14	5.7	0.4	2.1	81.8	4.6	0.8	2.5	0.7	0.6	0.3	0.2	0.2	396.4
M26	6.2	0.4	2.2	81.8	4.4	0.8	2.2	0.7	0.6	0.4	0.2	0.1	393.0
M6	5.9	0.4	2.2	81.8	4.5	0.8	2.4	0.7	0.6	0.3	0.2	0.2	392.9
M5	5.9	0.5	2.2	80.9	4.8	0.9	2.6	0.7	0.7	0.5	0.2	0.2	389.7

M17	6.3	0.4	2.3	75.1	4.7	4.8	4.4	0.7	0.6	0.3	0.2	0.1	388.4
M9	6.1	0.4	2.3	81.8	4.4	0.7	2.4	0.7	0.6	0.3	0.2	0.1	388.2
M20	6.2	0.4	2.2	81.5	4.7	0.7	2.3	0.7	0.6	0.3	0.2	0.1	379.1
M12	6.2	0.4	2.2	81.6	4.4	1.0	2.2	0.7	0.6	0.3	0.2	0.1	374.7
M18	6.2	0.4	2.4	81.3	4.7	0.7	2.3	0.8	0.5	0.3	0.2	0.1	369.1
M24	6.1	0.4	2.2	81.6	4.6	0.7	2.3	0.9	0.6	0.3	0.2	0.1	361.7
M7	5.9	0.4	2.3	81.9	4.5	0.7	2.3	0.7	0.6	0.3	0.2	0.1	359.4
M4	6.2	0.4	2.4	81.3	4.4	0.7	2.5	0.8	0.5	0.3	0.2	0.2	352.3
M13	6.1	0.4	2.3	81.4	4.5	0.8	2.4	0.7	0.6	0.3	0.2	0.1	352.0
M16	6.1	0.4	2.2	81.8	4.4	0.7	2.4	0.7	0.6	0.3	0.2	0.2	340.5
M19	6.1	0.4	2.3	80.9	4.9	0.8	2.7	0.7	0.5	0.3	0.2	0.1	318.6
M3	6.0	0.5	2.3	81.1	4.7	0.9	2.7	0.7	0.6	0.3	0.2	0.2	316.6

The coding region of the mouse MGAT2 gene, codon-optimised for expression in plant cells, was substituted for the MGAT1 coding region in the constructs mentioned above, and introduced into *Arabidopsis*. Thirty plants of each transgenic line (T1 and T2 plants, giving rise to T2- and T3-generation seeds) were grown in the greenhouse in a randomly arranged distribution and compared to control plants. Seeds from the transgenic plants were increased in their oil content relative to the control seeds (Figure 7). The average TAG percentage of the T3 transgenic seeds represented a relative increase of about 8% compared to the TAG percentage in the untransformed seeds (Table 4). A significant increase was observed in the level of polyunsaturated fatty acids in the TAG of the transgenic seeds, in particular of ALA, and a decrease in saturated fatty acid levels such as palmitic and stearic acids. Moreover, the increased TAG levels and altered fatty acid composition was more pronounced in the T3 generation than in the T2 seeds, presumably due to the homozygous state of the transgene in the T3 seeds.

Table 4. TAG levels and fatty acid composition in TAG extracted from *Arabidopsis thaliana* T2 and T3 seeds expressing MGAT2 compared to untransformed control seed.

	C14:0	C16:0	C16:1	C18:0	C18:1d9	C18:1d11	C18:2	C18:3
Control	0.1	8.1	0.3	3.3	13.2	1.8	28.4	20.0
T2 seeds	0.1	7.2	0.2	2.8	13.0	1.3	27.9	24.3
T3 seeds	0.1	6.1	0.2	2.5	12.3	1.2	28.4	26.7
	C20:0	C20:1	20:1iso	20:2n6	C22:1	C24:0	24:1	% TAG by Seed weight
	2.2	17.2	1.9	1.8	1.2	0.3	0.2	32.4
	1.6	14.8	3.1	1.8	1.2	0.3	0.2	37.9
	1.8	15.3	1.5	1.9	1.5	0.4	0.2	40.2

Expression of MGAT1 in *Brassica napus* seeds

The vector FP1:MGAT1 used for the expression of *M. musculus* MGAT1 in *Arabidopsis thaliana* seeds was used to generate transformed *B. napus* plants. The vector was introduced into *A. tumefaciens* strain AGL1 via standard electroporation procedures. Cultures were grown overnight at 28°C in LB medium with agitation at 150 rpm. The bacterial cells were collected by centrifugation at 4000 rpm for 5 minutes, washed with Winans' AB (Winans, 1988) and re-suspended in 10 mL of Winans' AB medium (pH 5.2) and grown with kanamycin (50 mg/L) and rifampicin

(25 mg/L) overnight with the addition of 100 μ M acetosyringone. Two hours before infection of the *Brassica* cells, spermidine (120 mg/L) was added and the final density of the bacteria adjusted to an OD 600nm of 0.3-0.4 with fresh AB media. Freshly isolated cotyledonary petioles from 8-day old *B. napus* seedlings grown on 1/2 MS (Murashige-Skoog, 1962) or hypocotyl segments preconditioned by 3-4 days on MS media with 1 mg/L thidiazuron (TDZ) + 0.1 mg/L alpha-naphthaleneacetic acid (NAA) were infected with 10 mL *Agrobacterium* cultures for 5 minutes. Explants (cotyledonary petiole and hypocotyl) infected with *Agrobacterium* were then blotted on sterile filter paper to remove the excess *Agrobacterium* and transferred to co-cultivation media (MS media with 1 mg/L TDZ + 0.1 mg/L NAA + 100 μ M acetosyringone) supplemented with or without different antioxidants (L-cysteine 50 mg/L and ascorbic 15 mg/L). All the plates were sealed with parafilm and incubated in the dark at 23-24°C for 48 hours.

The co-cultivated explants (cotyledonary petiole and hypocotyl) were then washed with sterile distilled water + 500 mg/L cefotaxime + 50 mg/L timentin for 10 minutes, rinsed in sterile distilled water for 10 minutes, blotted dry on sterile filter paper, transferred to pre-selection media (MS + 1 mg/L TDZ + 0.1 mg/L NAA + 20 mg/L adenine sulphate (ADS) + 1.5 mg/L AgNO₃ + 250 mg/L cefotaxime and 50 mg/L timentin) and cultured for five days at 24°C with a 16hour/8hour photoperiod. They were then transferred to selection media (MS + 1 mg/L TDZ + 0.1 mg/L NAA + 20 mg/L ADS + 1.5 mg/L AgNO₃ + 250 mg/L cefotaxime and 50 mg/L timentin) with 1.5 mg/L glufosinate ammonium and cultured for 4 weeks at 24°C with 16hour/8hour photoperiod with a biweekly subculture onto the same media. Explants with green callus were transferred to shoot initiation media (MS + 1 mg/L kinetin + 20 mg/L ADS + 1.5 mg/L AgNO₃ + 250 mg/L cefotaxime + 50 mg/L timentin + 1.5 mg/L glufosinate ammonium) and cultured for another 2-3 weeks. Shoots emerging from the resistant explants were transferred to shoot elongation media (MS media with 0.1 mg/L gibberelic acid + 20 mg/L ADS + 1.5 mg/L AgNO₃ + 250 mg/L cefotaxime + 1.5 mg/L glufosinate ammonium and cultured for another two weeks. Healthy shoots 2-3 cm long were selected and transferred to rooting media (1/2 MS with 1 mg/L NAA + 20 mg/L ADS + 1.5 mg/L AgNO₃ + 250 mg/L cefotaxime) and cultured for 2-3 weeks. Well established shoots with roots were transferred to pots (seedling raising mix) and grown in a growth cabinet for two weeks and subsequently transferred to glasshouse. Sixteen individual transformants in the cultivar Oscar were confirmed to be transgenic for the FP1:MGAT1 construct and grew normally under glasshouse conditions. Plant growth appeared normal and the plants were fertile, flowering and setting seed normally. The plants were grown to maturity and seeds obtained from transformed plants were harvested. Seeds from some of the transformed plants were

- analysed for seed oil content and fatty acid composition. Data from these preliminary analyses showed variability in the oil content and fatty acid composition, probably due to the plants being grown at different times and under different environmental conditions. To reduce variability, T1 plants which express MGAT1 are produced and
- 5 grown under the same conditions as control (wild-type, cultivar Oscar) plants of the same genotype, and the oil content compared.

Expression of MGAT1 in *Gossypium hirsutum* seeds

- The same seed-specific chimeric gene used for the expression of *M. musculus*
- 10 MGAT1 in *Arabidopsis thaliana* seeds was used to generate transformed *Gossypium hirsutum* plants. The vector designated FP1:MGAT1 was introduced into *A. tumefaciens* strain AGL1 via standard electroporation procedures and cells from the *Agrobacterium* culture used to introduce the chimeric DNAs into cells of *Gossypium hirsutum*, variety Coker315. Cotyledons excised from 10-day old cotton seedlings
- 15 were used as explants and infected and co-cultivated with *A. tumefaciens* for a period of two days. This was followed by a six-week selection on MS medium (Murashige and Skoog, 1962) containing 0.1 mg/L 2,4-D, 0.1 mg/L kinetin, 50 mg/L kanamycin sulphate, and 25 mg/L cefotaxime. Healthy calli derived from the cotyledon explants were transferred to MS medium containing 5 mg/L 6-(γ,γ -dimethylallylamino)-purine
- 20 (2ip), 0.1 mg/L naphthalene acetic acid (NAA), 25 mg/L kanamycin, and 250 mg/L cefotaxime for a second period of six weeks at 28°C. Somatic embryos that formed after about six to ten weeks of incubation were germinated and maintained on the same medium, but without added phytohormone or antibiotics. Plantlets developed from the somatic embryos were transferred to soil and maintained in a glasshouse
- 25 once leaves and roots were developed, with 28°C/20°C (day/night) growth temperature. Ten independent primary transgenic plants (T0) containing the FP1-MGAT1 construct were grown in the glasshouse, flowered and produced bolls containing seeds. The seeds were harvested on maturity. To enhance the reliability of the oil content analysis, 5 plants were established from each of the 10 primary
- 30 transgenic plants and the mature T2 seeds are subjected to the analysis of oil content. The seed-specific expression of MGAT1 increases oil content and increases the percentage of polyunsaturated fatty acids in the cotton seedoil.

Expression of a MGAT1 and MGAT2 genes in *N. benthamiana* plants after stable

35 transformation

N. benthamiana was stably transformed with the 35S:MGAT1 construct described in Example 2. 35S:MGAT1 was introduced into *A. tumefaciens* strain AGL1 via standard electroporation procedure. The transformed cells were grown on

solid LB media supplemented with kanamycin (50 mg/L) and rifampicin (25 mg/L) and incubated at 28°C for two days. A single colony was used to initiate fresh culture. Following 48 hours vigorous culture, the cells were collected by centrifugation at 2,000x g and the supernatant was removed. The cells were
5 resuspended in fresh solution containing 50% LB and 50% MS medium at the density of $OD_{600}=0.5$.

Leaf samples of *N. benthamiana* grown *in vitro* were excised and cut into square sections around 0.5-1 cm² in size with a sharp scalpel while immersed in the *A. tumefaciens* solution. The wounded *N. benthamiana* leaf pieces submerged in *A.*
10 *tumefaciens* were allowed to stand at room temperature for 10 minutes prior to being blotted dry on a sterile filter paper and transferred onto MS plates without supplement. Following a co-cultivation period of two days at 24°C, the explants were washed three times with sterile, liquid MS medium, then blotted dry with sterile filter paper and placed on the selective MS agar supplemented with 1.0 mg/L
15 benzylaminopurine (BAP), 0.25 mg/L indoleacetic acid (IAA), 50 mg/L kanamycin and 250 mg/L cefotaxime. The plates were incubated at 24°C for two weeks to allow for shoot development from the transformed *N. benthamiana* leaf pieces.

To establish rooted transgenic plants *in vitro*, healthy green shoots were cut off and transferred into 200 mL tissue culture pots containing MS agar medium
20 supplemented with 25 µg/L IAA, 50 mg/L kanamycin and 250 mg/L cefotaxime. Sufficiently large leaf discs were taken from transgenic shoots and freeze-dried for TAG fractionation and quantification analysis as described in Example 1 (Table 5). The best 35S:MGAT1 *N. benthamiana* plant had a TAG content of 204.85 µg/100 mg dry weight leaf tissue compared with an average TAG content of 85.02 µg/100 mg
25 dry weight leaf tissue in the control lines, representing an increase in TAG content of 241%.

N. benthamiana was also stably transformed with the 35S:MGAT2 construct described in Example 2 and a control binary vector pORE4 (Table 6). The best 35S:MGAT2 *N. benthamiana* plant had a TAG content of 79.0 µg/100 mg dry weight
30 leaf tissue compared with a TAG content of 9.5 µg/100 mg dry weight leaf tissue in the control line at the same developmental stage, representing an increase in TAG content of 731%. The fatty acid profile of the TAG fractions was also altered with significantly reduced levels of the saturated fatty acids 16:0 and 18:0, and increased levels of the polyunsaturated fatty acids, particularly 18:3ω3 (ALA) (Table 6). The
35 fatty acid profile of the polar lipids from the same leaf samples were not significantly affected, indicating that the changes in the fatty acid composition of the non-polar lipids was real. The control plants in this experiment were smaller and different physiologically than in the previous experiment with the 35S:MGAT1 construct, and

this may have explained the different oil contents of the control plants from one experiment to the other. Experiments to directly compare the 35S:MGAT1 and 35S:MGAT2 constructs with control plants are performed using plants of the same size and physiology.

5 A new set of constitutive binary expression vectors was made using a 35S promoter with duplicated enhancer region (e35S). 35S:MGAT1#2 (pJP3346), 35S:MGAT2#2 (pJP3347) and 35S:DGAT1#2 (pJP3352) were made by first cloning the e35S promoter, contained within a *Bam*HI-*Eco*RI fragment, into pORE04 at the *Bam*HI-*Eco*RI sites to yield pJP3343. pJP3346 and pJP3347 were then produced by
10 cloning the MGAT1 and MGAT2 genes, respectively, into the *Eco*RI site of pJP3343. pJP3352 was produced by cloning the *A. thaliana* DGAT1, contained within a *Xho*I-*Asi*SI site, into the *Xho*I-*Asi*SI sites of pJP3343.

pJP3346, pJP3347 and pJP3352 in *Agrobacterium* strain AGL1 were used to transform *N. benthamiana* as described above. Fourteen confirmed transgenic plants
15 were recovered for pJP3346 and 22 for pJP3347. A number of kanamycin resistant, transformed shoots have been generated with pJP3352. Expression analysis of the transgenes was performed on the plants transformed with MGAT1 or MGAT2. Plants with high levels of expression were selected. Expression analysis on plants transformed with the *A. thaliana* DGAT1 is performed. The plants grow normally
20 and are grown to maturity. Seed is harvested when mature. Seed from high-expressing progeny are sown directly onto soil for lipid analysis of the T2 segregating population, which includes both homozygous and heterozygous plants. Oil content of leaves of plants expressing high levels of either MGAT1 or MGAT2 is significantly increased compared to plants transformed with *A. thaliana* DGAT1 or control plants.
25 MGAT2 transgenic plants showed a significant increase in the unsaturated fatty acid 18:1 and 1 % relative increase in total fatty acid content compared to the null events (Table 7).

pJP3346, pJP3347 and a control vector in AGL1 were also used to transform *A. thaliana* as described above. Twenty-five confirmed transgenic T2 plants
30 comprising the T-DNA from pJP3346 and 43 transgenic plants for pJP3347 were identified. Expression analysis was performed on the transgenic plants. Seeds from high-expressing progeny were harvested and sown directly onto soil. Lipid analysis including oil content of the leaves from T2 and T3 progeny was performed, including from segregants lacking the transgenes. The highest levels of TAG were obtained in
35 plants that are homozygous for the MGAT transgenes.

Table 5. Fatty acid profile and quantification of TAG in *Nicotiana benthamiana* leaf tissue stably transformed with the 35S:MGAT1 construct. 'M' samples are 35S:MGAT1 whilst 'C' samples are parental control plants.

Sample	C16:0	16:3w3	C18:0	C18:1	C18:1d11	C18:2	C18:3	C20:0	20:3n3	C22:0	C24:0	µg/100mg DW
M1	38.7	0.7	5.1	8.5	0.4	7.0	34.4	1.1	0.3	0.2	0.4	204.85
M8	33.2	0.8	4.4	8.1	0.3	6.5	42.8	0.9	0.2	0.2	0.2	184.20
M3	41.1	0.6	5.3	10.4	0.4	5.5	31.8	1.0	0.4	0.2	0.2	133.62
M2	42.5	0.5	5.2	7.4	0.0	4.8	34.4	1.0	0.2	0.3	0.2	133.57
M7	35.2	0.6	4.5	8.6	0.0	4.9	41.7	1.1	0.3	0.3	0.2	128.49
M5	49.1	0.6	6.4	9.0	0.4	3.7	16.9	1.1	0.0	0.5	0.7	107.39
M4	41.9	0.4	6.0	9.6	0.0	4.2	33.0	1.1	0.2	0.4	0.2	93.71
M6	41.4	0.4	5.8	8.2	0.0	4.3	34.6	1.1	0.2	0.3	0.2	88.38
C1	40.2	0.4	6.1	8.3	0.0	7.8	31.9	1.3	0.2	0.4	0.3	81.53
C2	39.9	0.6	5.5	7.1	0.0	6.9	35.4	1.1	0.3	0.4	0.3	88.52

Table 6. Fatty acid profile and quantification of TAG in *Nicotiana benthamiana* leaf tissue stably transformed with the 35S:MGAT2 construct. 'M' samples are 35S:MGAT2 whilst 'C' samples are parental control plants. Two leaves from each plant were taken and analysed separately.

Sample		C16:0	16:1d7	16:1d13t	C16:1	16:3n3	C18:0	C18:1	C18:1d11	C18:2	C18:3n	C20:0	µg/100mg DW
C, leaf 1	TAG	34.0	2.7	0.8	0.0	0.0	17.3	6.6	0.0	15.9	18.7	0.0	12.9
C, leaf 2	TAG	35.0	1.8	0.0	0.0	1.3	25.0	3.0	0.0	13.0	17.6	1.4	6.1
M, leaf 1	TAG	14.6	0.4	1.0	0.4	7.7	5.9	4.0	0.4	16.8	47.0	0.6	97.1
M, leaf 2	TAG	18.1	0.3	1.0	0.0	6.0	8.1	2.8	0.3	14.0	46.9	1.0	60.9
C, leaf 1	PL	13.4	0.0	3.0	0.2	7.4	2.0	2.5	0.4	8.4	61.4	0.3	2439.3
C, leaf 2	PL	10.3	0.0	2.4	0.2	9.7	1.4	2.0	0.3	9.5	63.3	0.0	4811.5
M, leaf 1	PL	11.6	0.0	2.4	0.2	8.7	1.9	2.4	0.3	8.7	63.0	0.0	3568.8
M, leaf 2	PL	10.7	0.0	2.4	0.2	9.5	1.6	1.9	0.3	9.2	63.3	0.0	3571.2

Table 7. Total fatty acid amount (TFA) and fatty acid composition in *Nicotiana benthamiana* leaf tissues stably transformed with the 35S:MGAT2 construct.

	16:0	16:1	16:3	18:0	18:1 d9	18:1 d11	18:2	18:3 w3	20:0	20:1	20:1 iso	22:0	22:1	24:0	24:1	TFA (ug/100 ug DW)
MGAT	14.5	1.8	5.2	2.1	6.3	0.8	11.3	53.7	0.4	0.5	0.2	0.2	1.2	0.2	1.2	4.0
Nulls	15.1	2.2	6.0	2.7	3.9	0.6	9.6	56.3	0.4	0.4	0.1	0.2	0.7	0.2	1.1	3.2

Thirty plants of each transgenic line were grown in a random arrangement in the greenhouse with parental control plants. T2 seeds were analysed for oil content and exhibited an increase of about 2% in the oil content (total fatty acid level) compared to the total fatty acid content of parental seeds (Figure 8).

5

Expression of MGAT1 in stably transformed *Trifolium repens* plants

A chimeric gene encoding *M. musculus* MGAT1 was used to transform *Trifolium repens*, another dicotyledonous plant. Vectors containing the chimeric genes 35S:MGAT1 and 35S:DGAT1 were introduced into *A. tumefaciens* via a standard electroporation procedure. Both vectors also contain a 35S:BAR selectable marker gene. The transformed *Agrobacterium* cells were grown on solid LB media supplemented with kanamycin (50 mg/L) and rifampicin (25 mg/L) and incubated at 28°C for two days. A single colony was used to initiate a fresh culture for each construct. Following 48 hours vigorous culture, the *Agrobacterium* cultures were used to treat *T. repens* (cv. Haifa) cotyledons that had been dissected from imbibed seed as described by Larkin et al. (1996). Following co-cultivation for three days the explants were exposed to 5 mg/L PPT to select transformed shoots and then transferred to rooting medium to form roots, before transfer to soil. A transformed plant containing MGAT1 was obtained. The 35S promoter is expressed constitutively in cells of the transformed plants. The oil content is increased in at least the vegetative tissues such as leaves.

Expression of MGAT1 in stably transformed *Hordeum vulgare*

A chimeric vector including *M. musculus* MGAT1 was used to produce stably transformed *Hordeum vulgare*, a monocotyledonous plant. Vectors containing the chimeric genes Ubi:MGAT1 and Ubi:DGAT1 were constructed by cloning the entire *M. musculus* MGAT1 and *A. thaliana* DGAT1 coding regions separately into pWVEC8-Ubi. Vectors containing the chimeric genes Ubi:MGAT1 and Ubi:DGAT1 were introduced into *A. tumefaciens* strain AGL1 via a standard electroporation procedure. Transformed *Agrobacterium* cells were grown on solid LB media supplemented with kanamycin (50 mg/L) and rifampicin (25 mg/L) and the plates incubated at 28°C for two days. A single colony of each was used to initiate fresh cultures.

Following 48 hours vigorous culture, the *Agrobacterium* cultures were used to transform cells in immature embryos of barley (cv. Golden Promise) according to published methods (Tingay et al., 1997; Bartlett et al., 2008) with some modifications. Briefly, embryos between 1.5 and 2.5 mm in length were isolated from immature caryopses and the embryonic axes removed. The resulting explants were co-

cultivated for 2-3 days with the transgenic *Agrobacterium* and then cultured in the dark for 4-6 weeks on media containing timentin and hygromycin to generate embryogenic callus before being moved to transition media in low light conditions for two weeks. Callus was then transferred to regeneration media to allow for the
5 regeneration of shoots and roots before transfer to soil. Transformed plants were obtained and transferred to the greenhouse. The MGAT1 coding region was expressed constitutively under the control of the Ubi promoter in cells of the transformed plants. Transgenic plants were generated and their tissues analysed for oil content. Due to the low number of transgenic events obtained in a first
10 transformation, no statistically significant conclusion could be drawn from the data.

The coding region of the mouse MGAT2 gene, codon optimised for expression in plant cells, is substituted for the MGAT1 in the constructs mentioned above, and introduced into *Hordeum* as described above. Vegetative tissues from the resultant transgenic plants are increased for oil content.

15

Expression of MGAT in yeast cells

A chimeric vector including *M. musculus* MGAT1 was used to transform yeast, in this example *Saccharomyces cerevisiae*, a fungal microbe suitable for production of oil by fermentation. A genetic construct Gal1:MGAT1 was made by
20 inserting the entire coding region of a construct designated 0954364_MGAT_pMA, contained within an *EcoRI* fragment, into pYES2 at the *EcoRI* site, generating pJP3301. Similarly, a genetic construct Gal1:DGAT1, used here as a comparison and separately encoding the enzyme *A. thaliana* DGAT1 was made by inserting the entire *A. thaliana* DGAT1 coding region into pYES2. These chimeric vectors were
25 introduced into *S. cerevisiae* strain S288C by heat shock and transformants were selected on yeast minimal medium (YMM) plates containing 2% raffinose as the sole carbon source. Clonal inoculum cultures were established in liquid YMM with 2% raffinose as the sole carbon source. Experimental cultures were inoculated from these in YMM medium containing 1% NP-40, to an initial OD600 of about 0.3. Cultures
30 were grown at 28°C with shaking (about 100 rpm) until OD600 was approximately 1.0. At this point, galactose was added to a final concentration of 2% (w/v). Cultures were incubated at 25°C with shaking for a further 48 hours prior to harvesting by centrifugation. Cell pellets were washed with water before being freeze-dried for lipid class fractionation and quantification analysis as described in Example 1. The
35 Gal promoter is expressed inducibly in the transformed yeast cells, increasing the oil content in the cells.

The coding region of the mouse MGAT2 gene, codon optimised for expression in yeast cells, is substituted for the MGAT1 in the constructs mentioned above, and

introduced into yeast. The resultant transgenic cells are increased for oil content. The genes are also introduced into the oleaginous yeast, *Yarrowia lipolytica*, to increase oil content.

5 Expression of MGAT in algal cells *Chlamydomonas reinhardtii*

A chimeric vector including *M. musculus* MGAT1 is used to stably transform algal cells. The genetic constructs designated 35S:MGAT1 is made by cloning the MGAT1 coding region into a cloning vector containing a Cauliflower mosaic virus 35S promoter cassette and a paramomycin-resistance gene (aminoglycoside-O-phosphotransferase VIII) expressed by a *C. reinhardtii* RBCS2 promoter. 35S:MGAT1 is introduced separately into a logarithmic culture of 5×10^7 cc503, a cell-wall-deficient strain of *Chlamydomonas reinhardtii* by a modified glass bead method (Kindle, 1990). Both vectors also contain the BLE resistance gene as a selectable marker gene. Briefly, a colony of non-transformed cells on a TAP agar plate kept at about 24°C is grown to about 5×10^6 cells/mL over four days, the resultant cells are pelleted at 3000 g for 3 minutes at room temperature and resuspended to produce 5×10^7 cells in 300 μ L of TAP media. 300 μ L of 0.6mm diameter glass beads, 0.6 μ g plasmid in 5 μ L and 100 μ L of 20% PEG MW8000 are added and the mix is vortexed at maximum speed for 30 seconds, then transferred to 10 mL of TAP and incubated for 16 hours with shaking in the dark. The cells are pelleted, resuspended in 200 μ L of TAP then plated on TAP plates containing 5mg/L zeocin and incubated in the dark for 3 weeks. Transformed colonies are subcultured to a fresh TAP + zeocin 5 mg/L plate after which they are grown up under standard media conditions with zeocin selection. After harvesting by centrifugation, the cell pellets are washed with water before being freeze-dried for lipid class fractionation and quantification analysis as described in Example 1. The 35S:MGAT1 promoter is expressed constitutively in the transformed algal cells. The oil content of the cells is significantly increased.

The coding region of the mouse MGAT2 gene, codon optimised for expression in plant cells, is substituted for the MGAT1 in the construct mentioned above, and introduced into *Chlamydomonas*. Oil content in the resultant transgenic cells is significantly increased.

Expression of MGAT in stably transformed *Lupinus angustifolius*

A chimeric vector including *M. musculus* MGAT1 is used to transform *Lupinus angustifolius*, a leguminous plant. Chimeric vectors 35S:MGAT1 and 35S:DGAT1 in *Agrobacterium* are used to transform *L. angustifolius* as described by Pigeaire et al. (1997). Briefly, shoot apex explants are co-cultivated with transgenic

Agrobacterium before being thoroughly wetted with PPT solution (2 mg/ml) and transferred onto a PPT-free regeneration medium. The multiple axillary shoots developing from the shoot apices are excised onto a medium containing 20 mg/L PPT and the surviving shoots transferred onto fresh medium containing 20 mg/L PPT.

- 5 Healthy shoots are then transferred to soil. The 35S promoter is expressed constitutively in cells of the transformed plants, increasing the oil content in the vegetative tissues and the seeds. A seed specific promoter is used to further increase the oil content in transgenic *Lupinus* seeds.

- 10 The coding region of the mouse MGAT2 gene, codon optimised for expression in plant cells, is substituted for the MGAT1 in the constructs mentioned above, and introduced into *Lupinus*. Seeds and vegetative tissues from the resultant transgenic plants are increased for oil content.

Expression of MGAT in stably transformed cells of *Sorghum bicolor*

- 15 A chimeric vector including *M. musculus* MGAT1 is used to stably transform *Sorghum bicolor*. Ubi:MGAT1 and Ubi:DGAT1 in *A. tumefaciens* strain AGL1 are used to transform *Sorghum bicolor* as described by Gurel et al. (2009). The *Agrobacterium* is first centrifuged at 5,000 rpm at 4°C for 5 minutes and diluted to OD550 = 0.4 with liquid co-culture medium. Previously isolated immature embryos
- 20 are then covered completely with the *Agrobacterium* suspension for 15 minutes and then cultured, scutellum side up, on co-cultivation medium in the dark for 2 days at 24°C. The immature embryos are then transferred to callus-induction medium (CIM) with 100 mg/L carbenicillin to inhibit the growth of the *Agrobacterium* and left for 4 weeks. Tissues are then transferred to regeneration medium to shoot and root. The
- 25 Ubi promoter is expressed constitutively in cells of the transformed plants, increasing the oil content in at least the vegetative tissues.

- The coding region of the mouse MGAT2 gene, codon optimised for expression in plant cells, is substituted for the MGAT1 in the constructs mentioned above, and introduced into *Sorghum*. Vegetative tissues from the resultant transgenic plants are
- 30 increased for oil content.

Expression of MGAT in stably transformed plants of *Glycine max*

- A chimeric gene encoding *M. musculus* MGAT1 is used to stably transform *Glycine max*, another legume which may be used for oil production. 35S:MGAT1 in
- 35 *Agrobacterium* is used to transform *G. max* as described by Zhang et al. (1999). The *Agrobacterium* is co-cultivated for three days with cotyledonary explants derived from five day old seedlings. Explants are then cultured on Gamborg's B5 medium supplemented with 1.67 mg/L BAP and 5.0 mg/L glufosinate for four weeks after

which explants are subcultured to medium containing MS major and minor salts and B5 vitamins (MS/B5) supplemented with 1.0 mg/L zeatin-riboside, 0.5 mg/L GA3 and 0.1 mg/L IAA amended with 1.7 mg/L or 2.0 mg/L glufosinate. Elongated shoots are rooted on a MS/B5 rooting medium supplemented with 0.5 mg/L NAA without further glufosinate selection. The 35S promoter is expressed constitutively in cells of the transformed plants, increasing the oil content in the vegetative tissues and the seeds.

The coding region of the mouse MGAT2 gene, codon optimised for expression in plant cells, is substituted for the MGAT1 in the constructs mentioned above, and introduced into *Glycine*. Vegetative tissues and seeds from the resultant transgenic plants are increased for oil content.

Expression of MGAT in stably transformed *Zea mays*

A chimeric gene encoding *M. musculus* MGAT1 is used to stably transform *Zea mays*. The vectors comprising 35S:MGAT1 and 35S:DGAT1 are used to transform *Zea mays* as described by Gould et al. (1991). Briefly, shoot apex explants are co-cultivated with transgenic *Agrobacterium* for two days before being transferred onto a MS salt media containing kanamycin and carbenicillin. After several rounds of sub-culture, transformed shoots and roots spontaneously form and are transplanted to soil. The 35S promoter is expressed in cells of the transformed plants, increasing the oil content in the vegetative tissues and the seeds.

The coding region of the mouse MGAT2 gene, codon optimised for expression in plant cells, is substituted for the MGAT1 in the constructs mentioned above, and introduced into *Zea mays*. Vegetative tissues and seeds from the resultant transgenic plants are increased for oil content. Alternatively, the MGAT coding regions are expressed under the control of an endosperm specific promoter such as the zein promoter, or an embryo specific promoter obtained from a monocotyledonous plant, for increased expression and increased oil content in the seeds. A further chimeric gene encoding a GPAT with phosphatase activity, such as *A. thaliana* GPAT4 or GPAT6 is introduced into *Zea mays* in combination with the MGAT, further increasing the oil content in corn seeds.

Expression of MGAT in stably transformed *Elaeis guineensis* (palm oil)

A chimeric gene encoding *M. musculus* MGAT1 is used to stably transform *Elaeis guineensis*. Chimeric vectors designated Ubi:MGAT1 and Ubi:DGAT1 in *Agrobacterium* are used. Following 48 hours vigorous culture, the cells are used to transform *Elaeis guineensis* as described by Izawati et al. (2009). The Ubi promoter

is expressed constitutively in cells of the transformed plants, increasing the oil content in at least the fruits and seeds, and may be used to obtain oil.

The coding region of the mouse MGAT2 gene, codon optimised for expression in plant cells, is substituted for the MGAT1 in the constructs mentioned above, and introduced into *Elaeis*. Seeds from the resultant transgenic plants are increased for oil content.

Expression of MGAT in stably transformed *Avena sativa* (oats)

A chimeric gene encoding *M. musculus* MGAT1 is used to stably transform *Avena sativa*, another monocotyledonous plant. Chimeric vectors designated Ubi:MGAT1 and Ubi: DGAT1, as described above and both containing a Ubi:BAR selectable marker, are used to transform *Avena sativa* as described by Zhang et al. (1999).

The coding region of the mouse MGAT2 gene, codon optimised for expression in plant cells, is substituted for the MGAT1 in the constructs mentioned above, and introduced into *Avena*. Seeds from the resultant transgenic plants are increased for oil content.

Example 7. Engineering a MGAT with DGAT activity

An MGAT with altered DGAT activity, especially increased DGAT activity and potentially increased MGAT activity may be engineered by performing random mutagenesis, targeted mutagenesis, or saturation mutagenesis on MGAT gene(s) of interest or by subjecting different MGAT and/or DGAT genes to DNA shuffling. DGAT function can be positively screened for by using, for example, a yeast strain that has an absolute requirement for TAG-synthesis complementation when fed free fatty acids, such as strain H1246 which contains mutations in four genes (*DGA1*, *LRO1*, *ARE1*, *ARE2*). Transforming the MGAT variants in such a strain and then supplying the transformed yeast with a concentration of free fatty acids that prevents complementation by the wildtype MGAT gene will only allow the growth of variants with increased TAG-synthesis capability due to improved DGAT activity. The MGAT activity of these mutated genes can be determined by feeding labelled *sn*-1 or *sn*-2 MAG and quantifying the production of labelled DAG. Several rounds of directed evolution in combination with rational protein design would result in the production of a novel MGAT gene with similar MGAT and DGAT activities.

The gene coding for the *M. musculus* MGAT1 acyltransferase was subjected to error prone PCR using *Taq* DNA polymerase in the presence of 0.15 mM MnCl₂ to introduce random mutations. The randomized coding regions were then used as megaprimers to amplify the entire yeast expression vector using high fidelity PCR

reaction conditions. Sequencing of 9099 bp of recovered, mutagenised DNA revealed a mutational frequency of about 0.8 %, corresponding to 8 mutations per gene or, on average, 5.3 amino acid substitutions per polypeptide. The entire mutagenised library was transformed into *E. coli* DH5 α for storage at -80 °C and plasmid preparation.

5 The size of the MGAT1 library was estimated at 3.8356E6 clones. A copy of the MGAT1 library was transformed into the yeast strain H1246, resulting in a library size of 3E6 clones. The MGAT1 library as well as a pYES2 negative control, transformed into *S. cerevisiae* H1246, were subjected to 8 selection rounds, each consisting of (re)diluting cultures in minimal induction medium (1 % raffinose + 2 %

10 galactose; diluted to OD₆₀₀ = 0.35-0.7) in the presence of C18:1 free fatty acid at a 1 mM final concentration. Negative controls consisted of identical cultures grown simultaneously in minimal medium containing glucose (2%) and in the absence of C18:1 free fatty acid. After 8 selection rounds, an aliquot of the selected MGAT1 library was plated on minimal medium containing glucose (2%). A total of 120

15 colonies were grown in 240 μ l minimal induction medium in 96 microtiter plates and assayed for neutral lipid yield using a Nile Red fluorescence assay as described by Siloto et al. (2009). Plasmid minipreps were prepared from 113 clones (= top 6 %) that displayed the highest TAG levels.

The entire MGAT1 coding region of the selected clones is sequenced to

20 identify the number of unique mutants and to identify the nature of the selected mutations. Unique MGAT1 mutants are retransformed into *S. cerevisiae* H1246 for *in vitro* MGAT and DGAT assays using labelled MAG and C18:1 substrates respectively (see Example 5). Selected MGAT1 variants are found to exhibit increased DGAT activity compared to the wild type acyltransferase, whilst MGAT

25 activity is possibly increased as well.

MGAT1 variants displaying increased MGAT and/or DGAT activities are used as parents in a DNA shuffling reaction. The resulting library is subjected to a similar selection system as described above resulting in further improvement of general acyltransferase activity. In addition, free fatty acids other than C18:1 are added to the

30 growth medium to select for MGAT1 variants displaying altered acyl-donor specificities.

Example 8. Constitutive expression of the *A. thaliana* diacylglycerol acyltransferase 2 in plants

35 Expression of the *A. thaliana* DGAT2 in yeast (Weselake et al., 2009) and insect cells (Lardizabal et al., 2001) did not demonstrate DGAT activity. Similarly, the DGAT2 was not able to complement an *A. thaliana* DGAT1 knockout (Weselake et al., 2009). The enzyme activity of the *A. thaliana* DGAT2 in leaf tissue was

determined using a *N. benthamiana* transient expression system as described in Example 1. The *A. thaliana* DGAT2 (accession Q9ASU1) was obtained by genomic PCR and cloned into a binary expression vector under the control of the 35S promoter to generate 35S:DGAT2. This chimeric vector was introduced into *A. tumefaciens* strain AGL1 and cells from cultures of these infiltrated into leaf tissue of *N. benthamiana* plants in a 24°C growth room using 35S:DGAT1 as a control. Several direct comparisons were infiltrated with the samples being compared located on either side of the same leaf. Experiments were performed in triplicate. Following infiltration the plants were grown for a further five days before leaf discs were taken and freeze-dried for lipid class fractionation and quantification analysis as described in Example 1. This analysis revealed that both DGAT1 and DGAT2 were functioning to increase leaf oil levels in *Nicotiana benthamiana* (Table 8).

Leaf tissue transformed with the 35S:p19 construct (negative control) contained an average of 25 µg TAG/100 mg dry leaf weight. Leaf tissue transformed with the 35S:p19 and 35S:DGAT1 constructs (positive control) contained an average of 241 µg TAG/100 mg dry leaf weight. Leaf tissue transformed with the 35S:p19 and 35S:DGAT2 constructs contained an average of 551 µg TAG/100 mg dry leaf weight.

The data described above demonstrates that the *A. thaliana* DGAT2 enzyme is more active than the *A. thaliana* DGAT1 enzyme in promoting TAG accumulation in leaf tissue. Expression of the DGAT2 gene resulted in 229% as much TAG accumulation in leaf tissue compared to when the TAG amount from DGAT1 over-expressed was set as relative 100% (Figure 9).

Transiently-transformed *N. benthamiana* leaf tissues expressing P19 alone (control), or P19 with either AtDGAT1 or AtDGAT2 were also used to prepare microsomes for *in vitro* assays of enzyme activity. A DGAT biochemical assay was performed using microsomes corresponding to 50 µg protein and adding 10 nmole [14]C6:0-DAG and 5 nmole acyl-CoA, in 50 mM Hepes buffer, pH 7.2, containing 5 mM MgCl₂, and 1% BSA in a final volume of 100 µL for each assay. The assays were conducted at 30°C for 30 minutes. Total lipid from each assay was extracted and samples loaded on TLC plates, which were developed using a hexane:DEE:Hac solvent (70:30:1 vol:vol:vol). The amount of radioactivity in DAG and TAG spots was quantified by PhosphorImage measurement. The percentage of DAG converted to TAG was calculated for each of the microsome preparations.

Some endogenous DGAT activity was detected in the *N. benthamiana* leaves, as the P19 control assay showed low levels of TAG production. The expression of AtDGAT1 yielded increased DGAT activity relative to the P19 control when the assays were supplemented with either C18:1-CoA or C18:2-CoA, but not when

supplemented with C18:3-CoA, where the levels of TAG for the P19 control and the AtDGAT1 were similar. However, in all of the microsomal assays when AtDGAT2 was expressed in the leaf tissues, greater levels of DGAT activity (TAG production) were observed compared to the AtDGAT1 microsomes. Greater levels of TAG production were observed when the microsomes were supplemented with either C18:2-CoA or C18:3-CoA relative to C18:1-CoA (Figure 10). This indicated that DGAT2 had a different substrate preference, in particular for C18:3-CoA (ALA), than DGAT1.

10 **Example 9. Co-expression of MGAT and GPAT in transgenic seed**

Yang et al. (2010) described two glycerol-3-phosphate acyltransferases (GPAT4 and GPAT6) from *A. thaliana* both having a *sn*-2 preference (i.e. preferentially forming *sn*-2 MAG rather than *sn*-1/3 MAG) and phosphatase activity, which were able to produce *sn*-2 MAG from G-3-P (Figure 1). These enzymes were proposed to be part of the cutin synthesis pathway. GPAT4 and GPAT6 were not expressed highly in seed tissue. Combining such a bifunctional GPAT/phosphatase with a MGAT yields a novel DAG synthesis pathway using G-3-P as a substrate that can replace or supplement the typical Kennedy Pathway for DAG synthesis in plants, particularly in oilseeds, or other cells, which results in increased oil content, in particular TAG levels.

Chimeric DNAs designated pJP3382 and pJP3383, encoding the *A. thaliana* GPAT4 and GPAT6, respectively, together with the *M. musculus* MGAT2 for expression in plant seeds were made by first inserting the entire MGAT2 coding region, contained within a *Swa*I fragment, into pJP3362 at the *Sma*I site to yield pJP3378. pJP3362 was a binary expression vector containing empty FAE1 and FP1 expression cassettes and a kanamycin resistance gene as a selectable marker. The *A. thaliana* GPAT4 was amplified from cDNA and cloned into pJP3378 at the *Not*I site to yield pJP3382 in which the GPAT4 was expressed by the truncated napin promoter, FP1, and the MGAT2 was expressed by the *A. thaliana* FAE1 promoter. Similarly, the *A. thaliana* GPAT6 was amplified from cDNA and cloned into pJP3378 at the *Not*I site to yield pJP3384 in which the GPAT6 was operably linked to the truncated napin promoter, FP1, and the MGAT2 was expressed by the *A. thaliana* FAE1 promoter. pJP3382 and pJP3383 were transformed into *A. thaliana* (ecotype Columbia) by the floral dip method. Seeds from the treated plants were plated onto media containing the antibiotic, kanamycin, to select for progeny plants (T1 plants) which were transformed. Transgenic seedlings were transferred to soil and grown in the greenhouse. Expression of the transgenes in the developing embryos was determined. Transgenic plants with the highest level of expression and which show a

3:1 ratio for transgenic:non-transgenic plants per line, indicative of a single locus of insertion of the transgenes, are selected and grown to maturity. Seeds were obtained from these plants (T2) which included some which were homozygous for the transgenes. 30 to 32 (T2 plants) from each line were grown in pots of soil in a random arrangement in the greenhouse with control plants, and the lipid content, TAG content and fatty acid compositions of the resultant seed was determined. The total fatty acid content (as determined from the total FAME), in particular the TAG content of the seeds comprising both a MGAT and a GPAT4 or GPAT6 was substantially and significantly increased by nearly 3% (absolute level) or by about 9% (relative increase) over the controls, and increased relative to seeds comprising the MGAT alone or the *A. thaliana* DGAT1 alone (Figure 11).

The coding region of the mouse MGAT2 gene, codon optimised for expression in plant cells, was introduced into *Brassica napus* together with a chimeric gene encoding *Arabidopsis* GPAT4. Seeds from the resultant transgenic plants were harvested and some were analysed. Data from these preliminary analyses showed variability in the oil content and fatty acid composition, probably due to the plants being grown at different times and under different environmental conditions. Seeds are planted to produce progeny plants, and progeny seeds are harvested.

Example 10. Testing the effect of GPAT4 and GPAT6 on MGAT-mediated TAG increase by GPAT silencing and mutation

The GPAT family is large and all known members contain two conserved domains, a plsC acyltransferase domain and a HAD-like hydrolase superfamily domain. In addition to this, *A. thaliana* GPAT4-8 all contain an N-terminal region homologous to a phosphoserine phosphatase domain. *A. thaliana* GPAT4 and GPAT6 both contain conserved residues that are known to be critical to phosphatase activity (Yang et al., 2010).

Degenerate primers based on the conserved amino acid sequence GDLVICPEGTTCREP (SEQ ID NO:228) were designed to amplify fragments on *N. benthamiana* GPATs expressed in leaf tissue. 3' RACE will be performed using these primers and oligo-dT reverse primers on RNA isolated from *N. benthamiana* leaf tissue. GPATs with phosphatase activity (i.e. GPAT4/6-like) will be identified by their homology with the N-terminal phosphoserine phosphatase domain region described above. 35S-driven RNAi constructs targeting these genes will be generated and transformed in *A. tumefaciens* strain AGL1. Similarly, a 35S:V2 construct containing the V2 viral silencing-suppressor protein will be transformed in *A. tumefaciens* strain AGL1. V2 is known to suppress the native plant silencing

mechanism to allow effective transient expression but also allow RNAi-based gene silencing to function.

TAG accumulation will then be compared between transiently-transformed leaf samples infiltrated with the following strain mixtures: 1) 35S:V2 (negative control); 2) 35S:V2 + 35S:MGAT2 (positive control); 3) 35S:V2 + GPAT-RNAi; 4) 35S:V2 + GPAT-RNAi + 35S:MGAT2. It is expected that the 35S:V2 + GPAT-RNAi + 35S:MGAT2 mixture will result in less TAG accumulation than the 35S:V2 + 35S:MGAT2 sample due to interrupted *sn-2* MAG synthesis resulting from the GPAT silencing.

A similar experiment will be performed using *A. thaliana* and *N. benthamiana* GPAT4/6-like sequences which are mutated to remove the conserved residues that are known to be critical to phosphatase activity (Yang et al., 2010). These mutated genes (known collectively as GPAT4/6-delta) will then be cloned into 35S-driven expression binary vectors and transformed in *A. tumefaciens*. TAG accumulation will then be compared between transiently-transformed leaf samples infiltrated with the following strain mixtures: 1) 35S:p19 (negative control); 2) 35S:p19 + 35S:MGAT2 (positive control); 3) 35S:p19 + GPAT4/6-delta; 4) 35S:p19 + GPAT4/6-delta + 35S:MGAT2. It is expected that the 35S:p19 + GPAT4/6-delta + 35S:MGAT2 mixture will result in less TAG accumulation than the 35S:p19 + 35S:MGAT2 sample due to interrupted *sn-2* MAG synthesis resulting from the GPAT mutation. Whilst the native *N. benthamiana* GPAT4/6-like genes will be present in this experiment it is expected that high-level expression of the GPAT4/6-delta constructs will outcompete the endogenous genes for access to the G-3-P substrate.

Example 11. Constitutive expression of a diacylglycerol acyltransferase and WRI1 transcription factor in plant cells

A vector designated 35S-pORE04 was made by inserting a *Pst*I fragment containing a 35S promoter into the *Sfo*I site of vector pORE04 after T4 DNA polymerase treatment to blunt the ends (Coutu et al., 2007). A genetic construct 35S:Arath-DGAT1 encoding the *A. thaliana* diacylglycerol acyltransferase DGAT1 (Bouvier-Nave et al., 2000) was made. Example 3 of WO 2009/129582 describes the construction of AtDGAT1 in pXZP163. A PCR amplified fragment with *Kpn*I and *Eco*RV ends was made from pXZP163 and inserted into pENTR11 to generate pXZP513E. The entire AtDGAT1 coding region of pXZP513E contained within a *Bam*HI-*Eco*RV fragment was inserted into 35S-pORE04 at the *Bam*HI-*Eco*RV site, generating pJP2078. A synthetic fragment, Arath-WRI1, coding for the *A. thaliana* WRI1 transcription factor (Cernac and Benning, 2004), flanked by *Eco*RI restriction sites and codon optimized for *B. napus*, was synthesized. A genetic construct

designated 35S:Arath-WRI1 was made by cloning the entire coding region of Arath-WRI1, flanked by *EcoRI* sites into 35S-pORE04 at the *EcoRI* site generating pJP3414. Expression of the genes in *N. benthamiana* leaf tissue was performed according to the transient expression system as described in Example 1.

5 Quantification of TAG levels of infiltrated *N. benthamiana* leaves by Iatroscan revealed that the combined expression of the *A. thaliana* DGAT1 and WRI1 genes resulted in 4.5-fold and 14.3-fold increased TAG content compared to expression of WRI1 and the V2 negative control respectively (Table 9). This corresponded to an average and maximum observed TAG yield per leaf dry weight of 5.7 % and 6.51 %
10 respectively (Table 9 and Figure 12). The increase in leaf oil was not solely due to the activity of the overexpressed DGAT1 acyltransferase as was apparent in the reduced TAG levels when WRI1 was left out of the combination. Furthermore, a synergistic effect was observed accounting for 48 % of the total TAG increase.

Both DGAT1 and WRI1 constructs also led to increased oleic acid levels at the
15 expense of linoleic acid in TAG fractions of infiltrated *N. benthamiana* leaves (Table 10). These results confirm recent findings by Andrianov et al. (2010) who reported similar shifts in the TAG, phospholipid and TFA lipid fractions of transgenic tobacco plants transformed with the *A. thaliana* DGAT1 acyltransferase. However, when DGAT1 and WRI1 genes were co-expressed, a synergistic effect was observed on the
20 accumulation of oleic acid in the *N. benthamiana* leaves – this synergism accounted for an estimated at 52 % of the total oleic acid content when both genes were expressed. The unexpected synergistic effects on both TAG accumulation and oleic acid levels in transgenic *N. benthamiana* leaves demonstrated the potential of simultaneously up-regulating fatty acid biosynthesis and acyl uptake into non-polar
25 lipid such as TAG in vegetative tissues, two metabolic processes that are highly active in developing oilseeds.

The transient expression experiment was repeated except that the P19 viral silencing suppressor was substituted for the V2 suppressor, and with careful comparison of samples on the same leaf to avoid any leaf-to-leaf variation. For this, a chimeric
30 35S:P19 construct for expression of the tomato bushy stunt virus P19 viral silencing suppressor protein (Wood et al., 2009) was separately introduced into *A. tumefaciens* GV3101 for co-infiltration.

Quantification of TAG levels of infiltrated *N. benthamiana* leaves by Iatroscan in this experiment revealed that the combined transient expression of the *A. thaliana*
35 DGAT1 and WRI1 genes resulted in 141-fold increased TAG content compared to P19 negative control (Figure 13). When compared to the expression of the DGAT1 and WRI1 genes separately on the same leaf, the combined infiltration increased TAG levels by 17- and 5- fold respectively. Once again, the co-expression of both genes

had a synergistic (larger than additive) effect on leaf oil accumulation with the synergistic component accounting for 73% of the total TAG increase. The greater extent of the increased TAG content in this experiment (141-fold) compared to the previous experiment (14.3-fold) may have been due to use of the P19 silencing suppressor rather than V2 and therefore increased gene expression from the transgenes.

Table 11 shows the fatty acid composition of the TAG. When DGAT1 and WRI1 genes were co-expressed in *N. benthamiana*, a synergistic effect was once again observed on the level of oleic acid accumulation in the leaf TAG fraction. This increase was largely at the expense of the medium chain unsaturated fatty acids palmitic acid and stearic acid (Table 11). Linoleic acid was also increased which can be explained by the higher oleic acid substrate levels available to the endogenous FAD2 $\Delta 12$ -desaturase. Individual expression of the DGAT1 and WRI1 genes in *N. benthamiana* led to intermediate changes in the TAG profile without as great an increase in oleic acid. In addition, but in contrast to the first experiment, higher levels of α -linolenic acid (ALA) were detected while this was observed to a lesser extent upon the DGAT1 and WRI1 co-expression in leaf tissue.

The observed synergistic effect of DGAT1 and WRI1 expression on TAG biosynthesis was confirmed in more detail by comparing the effect of introduction into *N. benthamiana* of both genes individually or in combination, compared to introduction of a P19 gene alone as a control, within the same leaf. This was beneficial in reducing leaf to leaf variation. In addition, the number of replicates was increased to 5 and samples were pooled across different leaves from the same plant to improve the quality of the data. Results are presented in Table 12.

25

Table 8. Fatty acid profile and quantification of TAG in triplicate *Nicotiana benthamiana* leaf tissue transiently transformed with the 35S:p19, 35S:DGAT1 and 35S:DGAT2 constructs.

Sample	C16:0	16:1w13t	C16:1d7	16:3w3	C18:0	C18:1	C18:1d11	C18:2	C18:3	C20:0	20:1d11	20:2	20:3n3	C22:0	C24:0	µg/100mg DW
P19	44.7	0.1	0.0	0.0	33.9	1.2	0.0	6.5	12.7	0.9	0.0	0.0	0.0	0.0	0.0	43.29
	44.1	1.7	0.0	0.0	15.3	2.0	0.0	15.2	19.5	2.2	0.0	0.0	0.0	0.0	0.0	23.12
	43.3	1.5	0.0	0.0	10.5	1.5	0.0	17.2	23.9	2.2	0.0	0.0	0.0	0.0	0.0	38.35
P19+AtDGAT1	36.3	0.5	0.1	0.4	11.6	2.3	0.3	17.8	24.5	3.6	0.0	0.0	0.2	1.5	0.2	144.77
	33.6	0.5	0.1	0.4	11.2	2.9	0.3	23.1	21.5	3.8	0.0	0.0	0.2	1.5	0.9	145.34
	36.8	0.5	0.0	0.0	12.4	2.9	0.4	21.3	19.3	3.9	0.0	0.0	0.0	1.5	1.0	90.0
P19+AtDGAT2	18.6	0.3	0.1	0.5	9.3	7.7	0.4	28.0	33.1	1.1	0.2	0.1	0.1	0.2	0.3	439.2
	17.5	0.3	0.1	0.3	10.2	9.9	0.5	32.7	26.5	1.2	0.1	0.0	0.1	0.2	0.4	282.50
	18.4	0.3	0.1	0.3	9.8	7.5	0.5	32.3	29.1	1.2	0.0	0.0	0.0	0.3	0.2	208.40

Table 9. TAG levels in triplicate *Nicotiana benthamiana* leaf tissue transiently transformed with 35S:V2, 35S:DGAT1 and 35S:WRI1.

Control combination	TAG (% dry weight) ¹	Genes expressed	TAG (% dry weight) ¹	Ratio ²
V2	0.41 ± 0.10	V2, WRI1, DGAT1	5.71 ± 0.63	14.28 ± 1.89
V2, WRI1	1.16 ± 0.60	V2, WRI1, DGAT1	4.25 ± 0.64	4.45 ± 2.24
V2, DGAT1	1.52 ± 0.34	V2, WRI1, DGAT1	4.76 ± 0.50	3.22 ± 0.75

¹ Average of three different infiltrated leaves as quantified by Iatroscan² Average ratio based on side-by-side comparisons on the same leaves**Table 10.** Fatty acid composition of TAG produced in *Nicotiana benthamiana* leaf tissue transiently transformed with 35S:V2, 35S:DGAT1 and 35S:WRI1 (data from triplicate infiltrations).

Fatty acid	V2	V2, WRI1, DGAT1	V2, WRI1	V2, WRI1, DGAT1	V2, DGAT1	V2, WRI1, DGAT1
C14:0	0	0	0	0	0	0
C14:1 ^{Δ9}	1.26 ± 2.18	0.05 ± 0.10	0	0.04 ± 0.08	0	0.04 ± 0.07
C16:0	46.12 ± 0.97	30.60 ± 0.41	50.09 ± 6.27	31.32 ± 3.31	35.44 ± 0.80	26.61 ± 1.41
C16:1 ^{Δ9}	0	0.13 ± 0.11	0	0.07 ± 0.11	0	0.13 ± 0.12
C18:0	13.44 ± 1.65	9.93 ± 1.19	9.28 ± 0.81	9.93 ± 0.53	12.20 ± 1.03	8.76 ± 0.91
C18:1 ^{Δ9}	5.09 ± 5.32	36.78 ± 2.23	9.72 ± 5.08	27.97 ± 4.19	8.77 ± 1.97	32.41 ± 1.39
C18:1 ^{Δ11}	0	0.56 ± 0.04	0	0.51 ± 0.04	0	0.55 ± 0.04
C18:2 ^{Δ9,12}	14.12 ± 0.75	11.83 ± 0.75	13.26 ± 1.95	16.45 ± 3.88	18.93 ± 0.77	17.03 ± 1.36

C18:3 ^{Δ9,12,15}	19.98 ± 6.33	4.77 ± 1.17	17.10 ± 4.31	8.75 ± 2.13	16.12 ± 3.36	9.57 ± 0.61
C20:0	0	2.63 ± 0.27	0.54 ± 0.93	2.53 ± 0.16	4.25 ± 0.33	2.43 ± 0.26
C22:0	0	1.56 ± 0.1	0	1.38 ± 0.03	2.37 ± 0.11	1.40 ± 0.13
C24:0	0	1.17 ± 0.15	0	1.05 ± 0.07	1.92 ± 0.16	1.07 ± 0.16

Table 11. Fatty acid composition of TAG produced in *N. benthamiana* leaf tissues transiently transformed with 35S:P19 (control), 35S:WRII and/or 35S:DGAT1 constructs.

Fatty acid	P19	P19+WRII	P19+DGAT1	P19+WRII+DGAT1
C14:0	3.0 ± 2.2	0.6 ± 0.1	0.2 ± 0.1	0.1 ± 0.0
C16:0	46.5 ± 4.1	48.7 ± 11.5	28.4 ± 0.3	28.1 ± 1.0
C16:1 ^{Δ3t}	1.3 ± 2.2	0.3 ± 0.3	0.5 ± 0.0	0.3 ± 0.0
C16:1 ^{Δ9}	0.0	0.9 ± 0.2	0.2 ± 0.0	0.4 ± 0.1
C16:3 ^{Δ7,12,15}	0.0	0.2 ± 0.2	0.5 ± 0.1	0.3 ± 0.0
C18:0	18.7 ± 4.7	7.9 ± 2.6	11.5 ± 0.6	7.2 ± 0.4
C18:1 ^{Δ9}	5.5 ± 1.3	3.9 ± 0.3	6.3 ± 0.2	19.4 ± 2.7
C18:1 ^{Δ11}	0.0	0.6 ± 0.1	0.2 ± 0.0	0.6 ± 0.1
C18:2 ^{Δ9,12}	11.3 ± 4.2	12.7 ± 3.6	25.2 ± 0.5	26.3 ± 1.0
C18:3 ^{Δ9,12,15}	9.3 ± 3.6	21.6 ± 10.3	18.1 ± 0.6	11.2 ± 0.7
C20:0	2.7 ± 0.2	1.4 ± 0.5	4.4 ± 0.1	2.7 ± 0.1
C20:1 ^{Δ11}	0.0	0.0	0.3 ± 0.0	0.3 ± 0.0
C20:2 ^{Δ11,14}	0.0	0.0	0.1 ± 0.1	0.2 ± 0.0
C20:3 ^{Δ11,14,17}	0.0	0.0	0.1 ± 0.0	0.1 ± 0.0
C22:0	1.5 ± 0.1	0.7 ± 0.1	2.3 ± 0.0	1.8 ± 0.1
C24:0	0.4 ± 0.6	0.5 ± 0.2	1.6 ± 0.1	1.0 ± 0.1

Table 12. Comparison of WRI1 + DGAT1 together with the single genes

Gene combination	TAG level (% dry weight)	Ratio (compared to P19)
P19 (control)	0.01 ± 0.00	1
P19+WRI1	0.08 ± 0.04	8
P19+DGAT1	0.27 ± 0.03	27
P19+WRI1+DGAT1	1.29 ± 0.26	129

Based on the individual effects of both DGAT1 and WRI1 genes upon expression in *N. benthamiana*, in the presence of merely an additive effect but the absence of any synergistic effect, the present inventors expected a TAG level of about 0.35 or a 35-fold increase compared to the P19 negative control. However, the introduction of both genes resulted in TAG levels that were 129-fold higher than the P19 control. Based on these results, the present inventors estimated the additive effect and the synergistic effect on TAG accumulation as 26.9 % and 73.1 %, respectively. In addition, when the fatty acid composition of the total lipid in the leaf samples was analysed by GC, a synergistic effect was observed on C18:1^{Δ9} levels in the TAG fraction of *N. benthamiana* leaves infiltrated with WRI1 and DGAT1 (3 repeats each). The data is shown in Table 11.

For seed-specific expression of the WRI1 + DGAT1 combination, *Arabidopsis thaliana* was transformed with a binary vector construct including a chimeric DNA having both pFAE1::WRI1 and pCln2::DGAT1 genes, or, for comparison, the single genes pFAE1::WRI1 or pCln2::DGAT1. T1 seeds were harvested from the plants. The oil content of the seeds is determined. The seeds have an increased oil content.

Example 12. Constitutive expression of a monoacylglycerol acyltransferase and WRI1 transcription factor in plant cells

A chimeric DNA encoding the *Mus musculus* MGAT2 (Cao et al., 2003; Yen and Farese, 2003) and codon-optimised for *B. napus* was synthesized by Genecart. A genetic construct designated 35S:Musmu-MGAT2 was made by inserting the entire coding region of 1022341_MusmuMGAT2, contained within an *EcoRI* fragment, into pJP3343 at the *EcoRI* site, generating pJP3347. Cloning of the 35S:Arath-WRI1 construct is described in Example 11. Transient expression in *N. benthamiana* leaf tissue was performed as described in Example 1.

When the mouse MGAT2 and the *A. thaliana* WRI1 transcription vector were coexpressed, average *N. benthamiana* leaf TAG levels were increased by 3.3-fold compared to the expression of WRI1 alone (Table 13). In addition, the expression of the two genes resulted in a small (29 %) synergistic effect on the accumulation of leaf

TAG. The TAG level obtained with the MGAT2 gene in the presence of WRI1 was 3.78 % as quantified by Iatroscan (Figure 12). The similar results obtained with the animal MGAT2 and plant DGAT1 acyltransferases in combination with the *A. thaliana* WRI1 suggests that a synergistic effect might be a general phenomenon when WRI and acyltransferases are overexpressed in non-oil accumulating vegetative plant tissues.

The experiment was repeated to introduce constructs for expressing V2+MGAT2 compared to V2+MGAT2+WRI1, such that infiltrated leaf samples were pooled across three leaves from the same plant, for two plants each. In total, each combination therefore had 6 replicate infiltrations. This yielded a smaller standard deviation than pooling leaf samples between different plants as was done in the first experiments. The data from this experiment is shown in Table 14. Earlier results (Table 13) were confirmed. Although absolute TAG levels are different (inherent to the Benth assay and also different pooling of samples), relative increase in TAG when WRI1 is co-expressed with V2+MGAT2 are similar (2.45- and 2.65- fold).

Example 13. Constitutive expression of a monoacylglycerol acyltransferase, diacylglycerol acyltransferase and WRI1 transcription factor in plant cells

The genes coding for the *A. thaliana* diacylglycerol acyltransferase DGAT1, the mouse monoacylglycerol acyltransferase MGAT2 and the *A. thaliana* WRI1 were expressed in different combinations in *N. benthamiana* leaf tissue according to the transient expression system as described in Example 1. A detailed description of the different constructs can be found in Examples 11 and 12.

The combined expression of the DGAT1, WRI1 and MGAT2 genes resulted in an almost 3-fold further average TAG increase when compared to the expression of the latter two (Table 15). The maximum observed TAG yield obtained was 7.28 % as quantified by Iatroscan (Figure 12). Leaf TAG levels were not significantly affected when the gene of the mouse MGAT2 acyltransferase was left out this combination. Results described in Example 16, however, clearly demonstrated the positive effect of the mouse MGAT2 on the biosynthesis of neutral lipids in *N. benthamiana* leaves when expressed in combination with WRI1, DGAT1 and the *Sesamum indicum* oleosin protein.

Table 13. TAG levels in triplicate *Nicotiana benthamiana* leaf tissue transiently transformed with 35S:V2, 35S:MGAT2 and 35S:WR11.

Control combination	TAG (% dry weight) ¹	Genes expressed	TAG (% dry weight) ¹	Ratio ²
V2, WR11	0.93 ± 0.37	V2, WR11, MGAT2	2.88 ± 0.56	3.30 ± 0.85
V2, MGAT2	1.56 ± 0.76	V2, WR11, MGAT2	3.15 ± 1.05	2.45 ± 1.73

¹ Average of three different infiltrated leaves as quantified by latroscan² Average ratio based on side-by-side comparisons on the same leaves

Table 14. TAG content of infiltrated *N. benthamiana* leaf samples.

Gene combination	TAG (% dry weight)	Ratio
V2+MGAT2	0.34 ± 0.04	2.65
V2+MGAT2+WRII	0.9 ± 0.19	

Table 15. TAG levels in triplicate *Nicotiana benthamiana* leaf tissue transiently transformed with 35S:V2, 35S:MGAT2, 35S:DGAT1 and 35S:WRII.

Control combination	TAG (% dry weight) ¹	Genes expressed	TAG (% dry weight) ¹	Ratio ²
V2, WRII, DGAT1	3.35 ± 0.29	V2, WRII, MGAT2, DGAT1	3.15 ± 0.49	0.94 ± 0.01
V2, WRII, MGAT2	1.72 ± 0.56	V2, WRII, MGAT2, DGAT1	4.62 ± 0.47	

¹ Average of three different infiltrated leaves as quantified by Iatroscan² Average ratio based on side-by-side comparisons on the same leaves

Additional data was obtained from a further experiment where leaf samples were pooling across leaves within same plant, 6 replicates of each. The data is shown in Table 16.

5 **Table 16.** TAG content of infiltrated *N. benthamiana* leaf samples.

Gene combination	TAG (% dry weight)	Ratio
V2+MGAT2+DGAT1	1.08 ± 0.1	2.06
V2+MGAT2+DGAT1+WRI1	2.22 ± 0.31	

Example 14. Constitutive expression of a monoacylglycerol acyltransferase, diacycerol acyltransferase, WRI1 transcription factor and glycerol-3-phosphate acyltransferase in plant cells

- 10 A 35S:GPAT4 genetic construct was made by cloning the *A. thaliana* GPAT4 gene (Zheng et al., 2003) from total RNA isolated from developing siliques, followed by insertion as an *EcoRI* fragment into pJP3343 resulting in pJP3344. Other constructs are described in Examples 11 and 12. Transient expression in *N. benthamiana* leaf tissue was performed as described in Example 1.
- 15 Transient expression of the *A. thaliana* GPAT4 acyltransferase in combination with MGAT2, DGAT1 and WRI1 led to a small decrease in the *N. benthamiana* leaf TAG content as quantified by Iatroscan (Table 17). The TAG level (5.78 %) was also found to be lower when GPAT4 was included in the infiltration mixture (Figure 12). However, this finding does not rule out the hypothesis of *sn2*-MAG synthesis from
- 20 G3P as catalysed by the GPAT4 acyltransferase. Rather, it suggests that this catalytic step is unlikely to be rate limiting in leaf tissue due to the high expression levels of the endogenous GPAT4 gene (Li et al., 2007). Moreover, the *A. thaliana* GPAT8 acyltransferase displays a similar expression profile as GPAT4 and has been shown to exhibit an overlapping function (Li et al., 2007). In developing seeds the expression
- 25 levels of GPAT4 and GPAT8 are low. As a result, coexpression of GPAT4 in a seed context might be crucial to ensure sufficient *sn2*-MAG substrate for a heterologous expressed MGAT acyltransferase.

Table 17. TAG levels in triplicate *Nicotiana benthamiana* leaf tissue transiently transformed with 35S:V2, 35S:MGAT2, 35S:DGAT1, 35S:WRI1 and 35S:GPAT4.

Control combination	TAG (% dry weight) ¹	Genes expressed	TAG (% dry weight) ¹	Ratio ²
V2, WRI1, MGAT2, DGAT1	4.14 ± 0.82	V2, WRI1, MGAT2, DGAT1, GPAT4	3.11 ± 0.20	0.77 ± 0.13
V2, WRI1, DGAT1, GPAT4	2.76 ± 0.74	V2, WRI1, MGAT2, DGAT1, GPAT4	4.05 ± 1.24	1.47 ± 0.22

Additional data was obtained from a further experiment where leaf samples were pooling across leaves within same plant, 6 replicates of each. The data is shown in Table 18.

5 **Table 18.** TAG content of infiltrated *N. benthamiana* leaf samples.

Gene combination	TAG (% dry weight)	Ratio
V2+MGAT2+DGAT1	1.54 ± 0.36	1.01
V2+MGAT2+DGAT1+GPAT4	1.56 ± 0.18	

Example 15. Constitutive expression of a monoacylglycerol acyltransferase, diacycerol acyltransferase, WRI1 transcription factor and AGPase-hpRNAi silencing construct in plant cells

10 A DNA fragment corresponding to nucleotides 595 to 1187 of the mRNA encoding the *Nicotiana tabacum* AGPase small subunit (DQ399915) (Kwak et al., 2007) was synthesized. The 593 bp 1118501_NtAGP fragment was first cut with *Nco*I, treated with DNA polymerase I large (Klenow) fragment to generate 5' blunt ends and finally digested with *Xho*I. Similarly, the pENTR11-*NCOI* entry vector was
 15 first digested with *Bam*HI, treated with DNA polymerase I large (Klenow) fragment and cut with *Xho*I. Ligation of the 1118501_NtAGP insert into pENTR11-*NCOI* generated the pENTR11-*NCOI*-NtAGP entry clone. LR recombination between the pENTR11-*NCOI*-NtAGP entry clone and the pHELLSGATE12 destination vector generated pTV35, a binary vector containing the NtAGPase RNAi cassette under the
 20 control of the 35S promotor. Other constructs are described in Examples 11 and 12. Transient expression in *N. benthamiana* leaf tissue was performed as described in Example 1.

Expression of the *N. tabacum* AGPase silencing construct together with the genes coding for MGAT2 and WRI resulted in a 1.7-fold increase in leaf TAG levels
 25 as quantified by Iatroscan (Table 19). In the absence of the MGAT2 acyltransferase TAG levels dropped almost 3-fold. Therefore the observed TAG increase cannot be attributed solely to the silencing of the endogenous *N. benthamiana* AGPase gene. Surprisingly, substituting MGAT2 for the *A. thaliana* DGAT1 did not alter TAG levels in infiltrated *N. benthamiana* leaves in combination with the *N. tabacum*
 30 AGPase silencing construct. Silencing of the *N. benthamiana* AGPase therefore appears to have a different metabolic effect on MGAT and DGAT acyltransferases. A similar difference is also observed in the maximum observed TAG levels with WRI1 and the AGPase silencing construct in combination with MGAT2 or DGAT1 yielding 6.16 % and 5.51 % leaf oil respectively (Figure 12).

Table 19. TAG levels in triplicate *Nicotiana benthamiana* leaf tissue transiently transformed with 35S:V2, 35S:MGAT2, 35S:DGAT1, 35S:WRI1 and 35S:AGPase-hpRNAi

Control combination	TAG (% dry weight) ¹	Genes expressed	TAG (% dry weight) ¹	Ratio ²
V2, WRI1, MGAT2	2.33 ± 1.23	V2, WRI1, MGAT2, AGPase-hpRNAi	3.60 ± 0.98	1.69 ± 0.40
V2, WRI1, AGPase-hpRNAi	1.86 ± 0.20	V2, WRI1, MGAT2, AGPase-hpRNAi	5.21 ± 1.48	2.87 ± 1.01
V2, WRI1, DGAT1	4.99 ± 0.95	V2, WRI1, DGAT1, AGPase-hpRNAi	4.77 ± 0.79	0.96 ± 0.07

¹ Average of three different infiltrated leaves as quantified by Iatroscan; ² Average ratio based on side-by-side comparisons on the same leaves

Overexpression of WRI1 and MGAT in combination with AGPase silencing is particularly promising to increase oil yields in starch accumulating tissues. Examples include tubers such as for potatoes, and the endosperm of cereals, potentially leading to cereals with increased grain oil content (Barthole et al., 2011). Although *N. tabacum* and *N. benthamiana* AGPase genes are likely to bear significant sequence identity, it is likely that a *N. benthamiana* AGPase-hpRNAi construct will further elevate TAG yields due to improved silencing efficiency.

Additional data was obtained from a further experiment where leaf samples were pooled across leaves within same plant, 6 replicates of each. The data is shown in Tables 19 and 20.

Table 20. TAG content of infiltrated *N. benthamiana* leaf samples.

Gene combination	TAG (% dry weight)	Ratio
V2+MGAT2+DGAT1+WRI1+Oleosin	1.93 ± 0.18	1.14
V2+MGAT2+DGAT1+WRI1+Oleosin+AGPase-hpRNAi	2.19 ± 0.19	

Example 16. Constitutive expression of a monoacylglycerol acyltransferase, diacycerol acyltransferase, WRI1 transcription factor and an oleosin protein in plant cells

A pRSh1 binary vector containing the gene coding for the *S. indicum* seed oleosin (Scott et al., 2010) under the control of the 35S promotor was provided by Dr. N. Roberts (AgResearch Limited, New Zealand). Other constructs are described in Examples 11 and 12. Transient expression in *N. benthamiana* leaf tissue was performed as described in Example 1.

When the sesame oleosin protein was expressed together with the *A. thaliana* WRI transcription factor and *M. musculus* MGAT2 acyltransferase, TAG levels in *N. benthamiana* leaves as quantified by Iatroscan were found to be 2.2-fold higher (Table 21). No significant changes in the leaf TAG fatty acid profiles were detected (Table 22). A small increase in TAG was also observed when the *A. thaliana* DGAT1 acyltransferase was included. Compared to the V2 negative control, the combined expression of WRI1, DGAT1 and the sesame oleosin protein resulted in a 3-fold TAG increase and a maximum observed TAG level of 7.72 % (Table 21 and Figure 12). Leaf TAG levels were further elevated by a factor of 2.5 upon including the MGAT2 acyltransferase. This corresponded to an average of 5.7 % and a maximum observed of 18.8 % TAG on a dry weight basis. The additional increase in leaf TAG when

MGAT2 was included clearly demonstrates the positive effect of this acyltransferase on the biosynthesis and accumulation of neutral lipids in transgenic leaf tissues.

The experiment was repeated with the combination of genes for expressing V2 and V2+MGAT2+DGAT1+WRI1+Oleosin, tested in different *N. benthamiana* plants, with samples pooled across leaves from the same plant and with 12 replicate infiltrations for each. The data is shown in Table 23. Replicate samples were also pooled across leaves from same plant, with 6 repeats for each infiltration: The data is shown in Tables 24 and 25.

Although infiltration of *N. benthamiana* leaves resulted in increased levels of leaf oil (TAG), no significant increase in the total lipid content was detected, suggesting that a redistribution of fatty acids from different lipids pools into TAG was occurring. In contrast, when the MGAT2 gene was coexpressed with the DGAT1, WRI1 and oleosin genes, total lipids were increased 2.21-fold, demonstrating a net increase in the synthesis of leaf lipids.

15

Example 17. Constitutive expression of a monoacylglycerol acyltransferase, diacylglycerol acyltransferase, WRI1 transcription factor and a FAD2-hpRNAi silencing construct in plant cells

A *N. benthamiana* FAD2 RNAi cassette under the control of a 35S promotor was obtained by LR recombination into the pHELLSGATE8 destination vector to generate vector pFN033. Other constructs are described in Examples 11 and 12.

Table 21. TAG levels in triplicate *Nicotiana benthamiana* leaf tissue transiently transformed with 35S:V2, 35S:MGAT2, 35S:DGAT1, 35S:WRII and 35S:Oleosin.

Control combination	TAG (% dry weight) ¹	Genes expressed	TAG (% dry weight) ¹	Ratio ²
V2, WRII, MGAT2	1.77 ± 0.75	V2, WRII, MGAT2, Oleosin	3.34 ± 0.19	2.20 ± 1.10
V2, WRII, Oleosin	1.31 ± 0.19	V2, WRII, MGAT2, Oleosin	2.36 ± 1.10	1.79 ± 0.84
V2, WRII, DGAT1	4.82 ± 1.67	V2, WRII, DGAT1, Oleosin	6.02 ± 1.57	1.32 ± 0.43
V2, WRII, MGAT2, DGAT1	5.17 ± 1.87	V2, WRII, MGAT2, DGAT1, Oleosin	6.34 ± 1.74	1.25 ± 0.11
V2, WRII, DGAT1, Oleosin	4.61 ± 1.83	V2, WRII, MGAT2, DGAT1, Oleosin	5.48 ± 1.39	1.24 ± 0.26
V2	1.46 ± 0.67	V2, WRII, DGAT1, Oleosin	3.71 ± 1.50	3.00 ± 1.63
V2	0.90 ± 0.43	V2, WRII, MGAT2, DGAT1, Oleosin	5.74 ± 0.22	7.45 ± 3.52

¹ Average of three different infiltrated leaves as quantified by Iatroscan

² Average ratio based on side-by-side comparisons on the same leaves

Table 22. TAG fatty acid profiles of triplicate *Nicotiana benthamiana* leaf tissue transiently transformed with 35S:V2, 35S:MGAT2, 35S:DGAT1, 35S:WRI1 and 35S:Oleosin.

Fatty acid	V2, WRI1, DGAT1	V2, WRI1, DGAT1, Oleosin	V2, WRI1, DGAT1, Oleosin	V2, WRI1, MGAT2, DGAT1, Oleosin	V2, WRI1, MGAT2, DGAT1, Oleosin
C14:0	0.05 ± 0.04	0.02 ± 0.04	0.05 ± 0.05	0.04 ± 0.04	0.04 ± 0.04
C14:1 ^{Δ9}	0.14 ± 0.03	0.10 ± 0.09	0	0	0
C16:0	30.64 ± 1.32	29.96 ± 1.23	25.53 ± 2.30	23.74 ± 1.83	23.74 ± 1.83
C16:1 ^{Δ9}	0.39 ± 0.17	0.32 ± 0.33	0.19 ± 0.02	0.36 ± 0.10	0.36 ± 0.10
C18:0	9.85 ± 0.34	10.23 ± 0.20	8.50 ± 1.42	8.30 ± 0.73	8.30 ± 0.73
C18:1 ^{Δ9}	38.17 ± 1.28	39.01 ± 1.87	35.14 ± 6.58	38.64 ± 5.12	38.64 ± 5.12
C18:1 ^{Δ11}	0.66 ± 0.04	0.74 ± 0.20	0.53 ± 0.06	0.56 ± 0.02	0.56 ± 0.02
C18:2 ^{Δ9,12}	11.58 ± 0.52	11.53 ± 0.86	16.48 ± 1.40	15.75 ± 1.58	15.75 ± 1.58
C18:3 ^{Δ9,12,15}	3.80 ± 0.32	3.97 ± 0.29	9.35 ± 0.74	9.61 ± 0.87	9.61 ± 0.87
C20:0	2.50 ± 0.20	2.41 ± 0.15	2.17 ± 0.42	1.64 ± 0.11	1.64 ± 0.11
C22:0	1.33 ± 0.21	1.16 ± 0.11	1.22 ± 0.22	0.80 ± 0.04	0.80 ± 0.04
C24:0	0.90 ± 0.19	0.55 ± 0.48	0.85 ± 0.23	0.55 ± 0.06	0.55 ± 0.06

Table 23. TAG content of infiltrated *N. benthamiana* leaf samples.

Gene combination	TAG (% dry weight)	Ratio
V2	0.19 ± 0.05	18.74
V2+MGAT2+DGAT1+WRI1+Oleosin	3.56 ± 0.86	

Table 24. TAG content of infiltrated *N. benthamiana* leaf samples.

Gene combination	TAG (% dry weight)	Ratio
V2+MGAT2+DGAT1+WRI1	2.17 ± 0.30	0.79
V2+MGAT2+DGAT1+WRI1+Oleosin	2.11 ± 0.20	
V2+MGAT2	0.32 ± 0.06	2.19
V2+MGAT2+Oleosin	0.70 ± 0.17	

5 **Table 25.** Total fatty acid content of infiltrated *N. benthamiana* leaf samples.

V2	3.12 ± 0.14	1
V2+MGAT2	3.28 ± 0.33	1.05
V2+MGAT2+DGAT1+WRI1+Oleosin	6.88 ± 0.37	2.21

The genes coding for the mouse monoacylglycerol acyltransferase MGAT2, *A. thaliana* diacylglycerol acyltransferase DGAT1, *A. thaliana* WRI1 and a *N. benthamiana* FAD2 Δ 12-fatty acid desaturase hairpin RNAi construct (Wood et al., manuscript in preparation) were expressed in combination in *N. benthamiana* leaf tissue using the transient expression system as described in Example 1.

Similar changes were observed in the fatty acid compositions of TAG, polar lipids and TFA of *N. benthamiana* leaves infiltrated with WRI1, MGAT2, DGAT1 and the Fad2 silencing construct (Tables 26-28). In all three lipid fractions, oleic acid levels were further increased and reached almost 20 % in polar lipids, 40 % in TFA and more than 55 % in TAG. This increase came mostly at the expense of linoleic acid reflecting the silencing effect on the endoplasmic reticulum FAD2 Δ 12-desaturase. Leaf TAG also contained less α -linolenic acid while levels in TFA and polar lipids were unaffected.

When these experiments were repeated and the fatty acid compositions determined for TAG, polar lipids and total lipids, the results (Table 29) were consistent with the first experiment.

Table 26. TAG fatty acid profiles of triplicate *Nicotiana benthamiana* leaf tissue transiently transformed with 35S:V2, 35S:MGAT2, 35S:DGAT1, 35S:WRI1 and 35S:FAD2-hpRNAi.

Fatty acid	V2	V2, WRI1, MGAT2, DGAT1	V2, WRI1, MGAT2, DGAT1, FAD2-hpRNAi
C14:1 ^{Δ9}	0.28 ± 0.48	0.14 ± 0.12	0.08 ± 0.13
C16:0	22.73 ± 0.40	22.63 ± 1.43	19.11 ± 1.62
C16:1 ^{Δ9}	0	0.28 ± 0.02	0.51 ± 0.11
C18:0	7.31 ± 1.44	5.27 ± 0.19	5.05 ± 0.11
C18:1 ^{Δ9}	29.87 ± 11.91	32.21 ± 4.73	55.21 ± 1.31
C18:1 ^{Δ11}	0	0.80 ± 0.04	0.89 ± 0.04
C18:2 ^{Δ9,12}	13.36 ± 3.22	20.23 ± 3.36	3.61 ± 0.18
C18:3 ^{Δ9,12,15}	25.03 ± 10.14	15.18 ± 0.89	12.03 ± 0.72
C20:0	0.99 ± 0.86	1.38 ± 0.07	1.41 ± 0.04
C20:1 ^{Δ11}	0	0.39 ± 0.05	0.62 ± 0.02
C22:0	0	0.85 ± 0.04	0.83 ± 0.05
C24:0	0.44 ± 0.76	0.64 ± 0.08	0.66 ± 0.05

Table 27. Fatty acid profiles of polar lipids isolated from triplicate *Nicotiana benthamiana* leaf tissue transiently transformed with 35S:V2, 35S:MGAT2, 35S:DGAT1, 35S:WRI1 and 35S:FAD2-hpRNAi.

Fatty acid	V2	V2, WRI1, MGAT2, DGAT1	V2, WRI1, MGAT2, DGAT1, FAD2-hpRNAi
C14:1 ^{Δ9}	0.13 ± 0.23	0.17 ± 0.15	0
C16:0	15.00 ± 0.30	15.99 ± 0.14	15.28 ± 0.31
C16:1 ^{Δ9}	2.66 ± 0.28	1.97 ± 0.40	2.09 ± 0.16
C18:0	2.47 ± 0.14	2.05 ± 0.18	1.95 ± 0.09
C18:1 ^{Δ9}	5.12 ± 2.22	10.57 ± 1.99	18.99 ± 0.76
C18:1 ^{Δ11}	0.28 ± 0.24	0.61 ± 0.01	0.67 ± 0.03
C18:2 ^{Δ9,12}	10.26 ± 0.96	12.39 ± 1.33	5.20 ± 0.32
C18:3 ^{Δ9,12,15}	63.90 ± 1.18	55.70 ± 1.26	55.65 ± 1.23
C20:0	0.09 ± 0.15	0.19 ± 0.16	0.08 ± 0.14
C20:1 ^{Δ11}	0	0	0
C22:0	0	0.17 ± 0.15	0
C24:0	0.09 ± 0.15	0.19 ± 0.16	0.09 ± 0.16

Table 28. Fatty acid profiles of total lipids isolated from triplicate *Nicotiana benhamiana* leaf tissue transiently transformed with 35S:V2, 35S:MGAT2, 35S:DGAT1, 35S:WRI1 and 35S:FAD2-hpRNAi.

Fatty acid	V2	V2, WRI1, MGAT2, DGAT1	V2, WRI1, MGAT2, DGAT1, FAD2-hpRNAi
C14:1 ^{Δ9}	0.53 ± 0.08	0.27 ± 0.02	0.26 ± 0.02
C16:0	16.00 ± 1.05	19.70 ± 0.63	17.30 ± 0.72
C16:1 ^{Δ9}	2.02 ± 0.62	0.24 ± 0.02	0.28 ± 0.02
C18:0	3.75 ± 0.25	4.33 ± 0.09	4.17 ± 0.03
C18:1 ^{Δ9}	11.12 ± 6.77	23.32 ± 4.09	40.37 ± 2.24
C18:1 ^{Δ11}	0.46 ± 0.08	0.69 ± 0.03	0.75 ± 0.01
C18:2 ^{Δ9,12}	11.14 ± 0.83	17.28 ± 2.34	4.56 ± 0.29
C18:3 ^{Δ9,12,15}	53.27 ± 7.34	32.43 ± 1.49	30.43 ± 1.34
C20:0	0.51 ± 0.16	0.93 ± 0.05	0.92 ± 0.03
C20:1 ^{Δ11}	0	0.26 ± 0.03	0.40 ± 0.02
C22:0	0.83 ± 0.24	0.36 ± 0.06	0.37 ± 0.04
C24:0	0.38 ± 0.09	0.19 ± 0.03	0.20 ± 0.02

Table 29. Fatty acid composition of TAG, Polar lipids and total lipids in infiltrated *N. benthamiana* a leaf samples.

	TAG		Polar lipids		Total lipids	
	V2+MGAT2+D GAT1+WRII+ Oleosin	V2+MGAT2+DG AT1+WRII+Oleo sin+FAD2- hpRNAi	V2+MGAT2+DGA T1+WRII+Oleosin	V2+MGAT2+DGAT1 +WRII+Oleosin+FAD 2-hpRNAi	V2+MGAT2+DG AT1+WRII+Oleos in	V2+MGAT2+DGAT1 +WRII+Oleosin+FAD 2-hpRNAi
C14:0	0.00	0.00	0.00	0.00	0.06 ± 0.03	0.02 ± 0.03
C14:1 ^{Δ9}	0.00	0.00	0.05 ± 0.12	0.00	0.24 ± 0.03	0.19 ± 0.10
C16:0	19.63 ± 0.53	16.95 ± 1.13	15.85 ± 1.17	16.42 ± 2.11	16.88 ± 0.92	15.45 ± 1.24
C16:1 ^{Δ13}	0.00	0.15 ± 0.16	2.31 ± 0.38	1.61 ± 0.85	1.03 ± 0.23	0.85 ± 0.23
C16:3 ^{Δ7,1} 2,15	0.00	0.00	7.54 ± 0.40	7.42 ± 0.59	3.25 ± 0.47	2.79 ± 0.43
C18:0	6.64 ± 1.35	6.99 ± 0.43	2.88 ± 0.13	2.41 ± 1.21	5.36 ± 0.15	5.40 ± 0.18
C18:1 ^{Δ9}	29.45 ± 3.65	53.97 ± 1.51	8.56 ± 2.04	19.68 ± 1.32	20.59 ± 3.52	39.27 ± 2.28
C18:1 ^{Δ11}	0.59 ± 0.29	0.68 ± 0.34	0.45 ± 0.22	0.40 ± 0.31	0.59 ± 0.03	0.66 ± 0.05
C18:2 ^{Δ9,1} 2	23.47 ± 1.18	5.29 ± 0.37	13.13 ± 0.32	4.72 ± 2.33	18.35 ± 0.70	5.56 ± 0.28
C18:3 ^{Δ9,1} 2,15	18.32 ± 5.48	12.84 ± 0.60	49.04 ± 2.78	47.22 ± 5.40	30.03 ± 3.51	26.28 ± 3.07
C20:0	1.12 ± 0.56	1.56 ± 0.04	0.19 ± 0.21	0.11 ± 0.17	0.88 ± 0.06	1.03 ± 0.07

C20:1 ^{Δ11}	0.00	0.20 ± 0.22	0.00	0.00	0.08 ± 0.08	0.22 ± 0.11
C22:0	0.57 ± 0.28	0.81 ± 0.07	0.00	0.00	0.54 ± 0.05	0.61 ± 0.07
C22:1	0.00	0.00	0.00	0.00	0.50 ± 0.03	0.22 ± 0.24
C22:2n6	0.00	0.00	0.00	0.00	0.80 ± 0.14	0.77 ± 0.13
C24:0	0.21 ± 0.32	0.57 ± 0.29	0.00	0.00	0.42 ± 0.01	0.48 ± 0.03
C24:1	0.00	0.00	0.00	0.00	0.42 ± 0.05	0.18 ± 0.2

Example 18. Expression of *Mus musculus* MGAT1 and MGAT2 in *Nicotiana benthamiana* cells by stable transformation

Constitutive expression in *N. benthamiana*

The enzyme activity of the *M. musculus* MGAT1 and MGAT2 was demonstrated in *Nicotiana benthamiana*. The chimeric vectors 35S:Musmu-MGAT1 and 35S:Musmu-MGAT2 were introduced into *A. tumefaciens* strain AGL1 via standard electroporation procedure and grown on solid LB media supplemented with kanamycin (50 mg/L) and rifampicin (25 mg/L) and incubated at 28°C for two days. A single colony was used to initiate fresh culture. Following 48 hours culturing with vigorous aeration, the cells were collected by centrifugation at 2,000x g and the supernatant were removed. The cells were resuspended in a new solution containing 50 % LB and 50 % MS medium at the density of OD₆₀₀ =0.5. Leaf samples of *Nicotiana benthamiana* plants grown aseptically *in vitro* were excised and cut into square sections around 0.5-1 cm² in size with a sharp scalpel while immersed in the *A. tumefaciens* solution. The wounded *N. benthamiana* leaf pieces submerged in *A. tumefaciens* were allowed to stand at room temperature for 10 min prior to being blotted dry on a sterile filter paper and transferred onto MS plates without supplement. Following a co-cultivation period of two days at 24°C, the explants were washed three times with sterile liquid MS medium, and finally blot dry with sterile filter paper and placed on the selective agar-solidified MS medium supplemented with 1.0 mg/L benzylaminopurine (BAP), 0.25 mg/L indoleacetic acid (IAA), 50 mg/L kanamycin and 250 mg/L cefotaxime and incubated at 24°C for two weeks to allow for shoot development from the transformed *N. benthamiana* leaf discs. To establish *in vitro* transgenic plants, healthy green shoots were cut off and transferred onto a new 200 mL tissue culture pots containing agar-solidified MS medium supplemented with 25 µg/L IAA and 50 mg/L kanamycin and 250 mg/L cefotaxime.

Expression of the MGAT1 and MGAT2 transgenes was determined by Real-Time PCR. Highly-expressing lines were selected and their seed harvested. This seed was planted directly onto soil and the segregating population of seedlings harvested after four weeks. Highly-expressing events were selected and seed produced by these planted out directly onto soil to result in a segregating population of 30 seedlings. After three weeks leaf discs were taken from each seedling for DNA extraction and subsequent PCR to determine which lines were transgenic and which were null for the transgene. The population was then harvested with the entire aerial tissue from each seedling cleaned of soil and freeze-dried. The dry weight of each sample was recorded and total lipids isolated. The TAG in these total lipid samples was quantified by TLC-FID and the ratio of TAG to an internal standard (DAGE) in each sample determined (Figure 14). The average level of TAG in the transgenic seedlings

of 35S:Musmu-MGAT2 line 3347-19 was found to be 4.1-fold higher than the average level of TAG in the null seedlings. The event with the largest increase in TAG had 7.3-fold higher TAG than the average of the null events.

Constitutive expression in *A. thaliana*

The enzyme activity of the *M. musculus* MGAT1 and MGAT2 was demonstrated in *A. thaliana*. The chimeric vectors 35S:Musmu-MGAT1 and 35S:Musmu-MGAT2 along with the empty vector control pORE04 were transformed in *A. thaliana* by the floral dip method and seed from primary transformants selected on kanamycin media. The T2 seed from these T1 plants was harvested and TFA of the seeds from each plant determined (Figure 15). The average mg TFA/g seed was found to be 139±13 for the control pORE04 lines with median 136.0, 152±14 for the 35S:MGAT1 lines with median 155.1 and 155±11 for the 35S:MGAT2 lines with median 154.7. This represented an average TFA increase compared to the control of 9.7 % for 35S:MGAT1 and 12.1 % for 35S:MGAT2.

Example 19. Additional genes

Further increases in oil

Additional genes are tested alongside the combinations described above to determine whether further oil increases can be achieved. These include the following *Arabidopsis* genes: AT4G02280, Sucrose synthase SUS3; AT2G36190, Invertase CWINV4; AT3G13790, Invertase CWINV1; AT1G61800, Glucose 6 phosphate:phosphate translocator GPT2; AT5G33320, Phosphoenolpyruvate transporter PPT1; AT4G15530, Pyruvate orthophosphate dikinase Plastid-PPDK; AT5G52920, Pyruvate kinase pPK-β1. The genes coding for these enzymes are synthesised and cloned into the constitutive binary expression vector pJP3343 as *EcoRI* fragments for testing in *N. benthamiana*.

When a number of genes were added to the combination of WRI1, DGAT1, MGAT2 and oleosin and expressed in *N. benthamiana* leaves, no additional increase in the level of TAG was observed, namely for: safflower PDAT, *Arabidopsis thaliana* PDAT1, *Arabidopsis thaliana* DGAT2, *Arabidopsis thaliana* caleosin, peanut oleosin, *Arabidopsis thaliana* haemoglobin 2, Homo sapiens iPLA α , *Arabidopsis thaliana* GPAT4, *E. coli* G3P dehydrogenase, yeast G3P dehydrogenase, castor LPAAT2, *Arabidopsis thaliana* beta-fructofuranosidase (ATBFRUCT1, NM_112232), *Arabidopsis thaliana* beta-fructofuranosidase (cwINV4, NM_129177), indicating that none of these enzyme activities were rate limiting in *N. benthamiana* leaves when expressed transiently. This does not indicate that they will have no effect in stably-transformed plants, such as in seed, or in other organisms.

Further additional genes are tested for additive or synergistic oil increase activity. These include the following *Arabidopsis thaliana* gene models or their encoded proteins, and homologues from other species, which are grouped by putative function and have previously been shown to be upregulated in tissues with increased oil content. Genes/proteins involved in sucrose degradation: AT1G73370, AT3G43190, AT4G02280, AT5G20830, AT5G37180, AT5G49190, AT2G36190, AT3G13784, AT3G13790, AT3G52600. Genes/proteins involved in the oxidative pentose phosphate pathway: AT3G27300, AT5G40760, AT1G09420, AT1G24280, AT5G13110, AT5G35790, AT3G02360, AT5G41670, AT1G64190, AT2G45290, AT3G60750, AT1G12230, AT5G13420, AT1G13700, AT5G24410, AT5G24420, AT5G24400, AT1G63290, AT3G01850, AT5G61410, AT1G71100, AT2G01290, AT3G04790, AT5G44520, AT4G26270, AT4G29220, AT4G32840, AT5G47810, AT5G56630, AT2G22480, AT5G61580, AT1G18270, AT2G36460, AT3G52930, AT4G26530, AT2G01140, AT2G21330, AT4G38970, AT3G55440, AT2G21170. Genes/proteins involved in glycolysis: AT1G13440, AT3G04120, AT1G16300, AT1G79530, AT1G79550, AT3G45090, AT5G60760, AT1G56190, AT3G12780, AT5G61450, AT1G09780, AT3G08590, AT3G30841, AT4G09520, AT1G22170, AT1G78050, AT2G36530, AT1G74030. Genes/proteins which function as plastid transporters: AT1G61800, AT5G16150, AT5G33320, AT5G46110, AT4G15530, AT2G36580, AT3G52990, AT3G55650, AT3G55810, AT4G26390, AT5G08570, AT5G56350, AT5G63680, AT1G32440, AT3G22960, AT3G49160, AT5G52920. Genes/proteins involved in malate and pyruvate metabolism: AT1G04410, AT5G43330, AT5G56720, AT1G53240, AT3G15020, AT2G22780, AT5G09660, AT3G47520, AT5G58330, AT2G19900, AT5G11670, AT5G25880, AT2G13560, AT4G00570.

Constructs are prepared which include sequences encoding these candidate proteins, which are infiltrated into *N. benthamiana* leaves as in previous experiments, and the fatty acid content and composition analysed. Genes which aid in increasing non-polar lipid content are combined with the other genes as described above, principally those encoding MGAT, Wri1, DGAT1 and an Oleosin, and used to transform plant cells.

Increases in unusual fatty acids

Additional genes are tested alongside the combinations described above to determine whether increases in unusual fatty acids can be achieved. These include the following genes (provided are the GenBank Accession Nos.) which are grouped by putative function and homologues from other species. Delta-12 acetylenases ABC00769, CAA76158, AAO38036, AAO38032; Delta-12 conjugases AAG42259,

AAG42260, AAN87574; Delta-12 desaturases P46313, ABS18716, AAS57577, AAL61825, AAF04093, AAF04094; Delta-12 epoxygenases XP_001840127, CAA76156, AAR23815; Delta-12 hydroxylases ACF37070, AAC32755, ABQ01458, AAC49010; and Delta-12 P450 enzymes such as AF406732.

Constructs are prepared which include sequences encoding these candidate proteins, which are infiltrated into *N. benthamiana* leaves as in previous experiments, and the fatty acid content and composition analysed. The nucleotide sequences of the coding regions may be codon-optimised for the host species of interest. Genes which aid in increasing unusual fatty acid content are combined with the other genes as described above, principally those encoding MGAT, WRI1, DGAT1 and an Oleosin, and used to transform plant cells.

Example 20. Stable transformation of plants including *Nicotiana tabacum* with combinations of oil increase genes

An existing binary expression vector, pORE04+11ABGBEC (US Provisional Patent Application No. 61/660392), which contained a double enhancer-region 35S promoter expressing the NPTII kanamycin resistance gene and three gene expression cassettes, was used as a starting vector to prepare several constructs each containing a combination of genes for stable transformation of plants. This vector was modified by exchanging the expressed genes with oil increase genes, as follows. pORE04+11ABGBEC was first modified by inserting an intron-interrupted sesame oleosin gene, flanked by *NotI* sites, from the vector pRSh1-PSP1 into the pORE04+11ABGBEC *NotI* sites to generate pJP3500. pJP3500 was then modified by inserting a codon-optimised DNA fragment encoding the *A. thaliana* WRL1 gene into the *EcoRI* sites to generate pJP3501. pJP3501 was further modified by inserting a DNA fragment encoding the wild-type *A. thaliana* DGAT1 coding region, flanked by *AsiSI* sites, into the *AsiSI* sites to generate pJP3502 (SEQ ID NO:409). A final modification was made by inserting another expression cassette, consisting of a double enhancer-region 35S promoter expressing a coding region encoding the *M. musculus* MGAT2, as a *StuI*-*ZraI* fragment into the *SfoI* site of pJP3502 to generate pJP3503 (SEQ ID NO:410). The MGAT2 expression cassette was excised from pJP3347 at the *StuI* + *ZraI* sites. pJP3502 and pJP3503 were both used to stably transform *N. tabacum* as described below. By these constructions, pJP3502 contained the *A. thaliana* WRL1 and DGAT1 coding regions driven by the *A. thaliana* Rubisco small subunit promoter (SSU) and double enhancer-region 35S promoter, respectively, as well as a SSU:sesame oleosin cassette. The T-DNA region of this construct is shown schematically in Figure 16. The vector pJP3503 additionally contained the e35S::MGAT2 cassette. This construct is shown schematically in Figure

17. The nucleotide sequence of the T-DNA region of the construct pJP3503 is given as SEQ ID NO:412.

Stable transformation of *Nicotiana tabacum* with combinations of genes

The binary vectors pJP3502 and pJP3503 were separately introduced into the *A. tumefaciens* strain AGL1 by a standard electroporation procedure. Transformed cells were selected and grown on LB-agar supplemented with kanamycin (50 mg/l) and rifampicin (25 mg/l) and incubated at 28°C for two days. A single colony of each was used to initiate fresh cultures in LB broth. Following 48 hours incubation with vigorous aeration, the cells were collected by centrifugation at 2,000 g and the supernatant was removed. The cells were resuspended at the density of $OD_{600} = 0.5$ in fresh medium consisting of 50% LB and 50% MS medium.

Leaf samples of *N. tabacum* cultivar W38 grown aseptically *in vitro* were excised with a scalpel and cut into pieces of about 0.5-1 cm² in size while immersed in the *A. tumefaciens* suspensions. The cut leaf pieces were left in the *A. tumefaciens* suspensions at room temperature for 15 minutes prior to being blotted dry on a sterile filter paper and transferred onto MS plates without antibiotic supplement. Following a co-cultivation period of two days at 24°C, the explants were washed three times with sterile, liquid MS medium, then blotted dry with sterile filter paper and placed on the selective MS agar supplemented with 1.0 mg/L benzylaminopurine (BAP), 0.5 mg/L indoleacetic acid (IAA), 100 mg/L kanamycin and 200 mg/L cefotaxime. The plates were incubated at 24°C for two weeks to allow for shoot development from the transformed *N. tabacum* leaf pieces.

To establish rooted transgenic plants *in vitro*, healthy green shoots were cut off and transferred to MS agar medium supplemented with 25 µg/L IAA, 100 mg/L kanamycin and 200 mg/L cefotaxime. After roots had developed, individual plants were transferred to soil and grown in the glasshouse. Leaf samples were harvested at different stages of plant development including before and during flowering. Total fatty acids, polar lipids and TAG were quantified and their fatty acid profiles determined by TLC/GC as described in Example 1.

Analysis of pJP3503 transformants

For the transformation with pJP3503 ("4-gene construct"), leaf samples of about 1 cm² were taken from 30 primary transformants prior to flower buds forming and TAG levels in the samples were quantified by Iatroscan. Seven plants were selected for further analysis, of which five displaying increased leaf oil levels and two exhibiting oil levels essentially the same as wild-type plants. Freeze-dried leaf samples from these plants were analysed for total lipid content and TAG content and

fatty acid composition by TLC and GC. Transformed plants numbered 4 and 29 were found to have considerably increased levels of leaf oil compared to the wild type, while plant number 21 exhibited the lowest TAG levels at essentially wild-type levels (Table 30). Plants numbered 11, 15 and 27 had intermediate levels of leaf oil. Oleic acid levels in TAG were found to be inversely correlated to the TAG yields, consistent with the results of the earlier transient expression experiments in *N. benthamiana*.

In the transformed plants numbered 4 and 29, leaf oil content (as a percentage of dry weight) was found to increase considerably at the time of flowering (Table 31). From the data in Table 31, the increase was at least 1.7- and 2.4-fold for plants 4 and 29, respectively. No such change was observed for plant 21 which had TAG levels similar to the wild-type control. Oleic acid levels in the TAG fractions isolated from each sample followed a similar pattern. This fatty acid accumulated up to 22.1% of the fatty acid in TAG from plants 4 and 29, a 17-18-fold increase compared to plant 21 and the wild-type. The increase in oleic acid was accompanied by increased linoleic acid and palmitic acid levels while α -linolenic acid levels dropped 8-fold compared to in plant 21 and the wild-type control. Unlike TAG, polar lipid levels decreased slightly at the flowering stage in the three lines (Table 32). Changes in C18 monounsaturated and polyunsaturated fatty acid levels in the polar lipid fractions of the three lines were similar to the shifts in their TAG composition although the changes in oleic acid and linoleic acid were less marked. Significant increases in total leaf lipids were observed for lines 4 and 29 during flowering with levels reaching more than 10% of dry weight (Table 33). Total leaf lipid levels in plant 21 before and during flowering were similar to levels observed in wild-type plants at similar stages (Tables 33 and 35). Changes in the total lipid fatty acid composition of all three plants were similar to the respective TAG fatty acid compositions. Leaf oil in plant 4 during seed setting was found to be further elevated at the onset of leaf chlorosis. The highest leaf TAG levels detected at this stage corresponded to a 65-fold increase compared to similar aged leaves in plant 21 during seed setting and a 130-fold increase compared to similar leaves of flowering wild-type plants (Table 34; Figure 18).

The increased TAG in this plant coincided with elevated oleic acid levels. Unlike plant 4, leaf TAG levels in the other two primary transformants and wild-type tobacco did not increase, or only marginally increased, after flowering and during chlorosis. The lower leaves of plants 4 and 29 exhibited reduced TAG levels upon senescence. In all plants, linoleic acid levels dropped while α -linolenic acid levels were increased with progressing leaf age.

Consistent with the increased TAG levels, total lipid levels in leaves of plants 4 and 29 during seed setting were further elevated compared to similar leaves of both plants during flowering (Tables 33 and 35). The highest total lipid level detected in plant 4 on a dry weight basis was 15.8 %, equivalent to a 7.6- and 11.2-fold increase compared to similar leaves of plant 21 and wild-type plants, respectively. Whilst the fatty acid composition of total lipid in the leaves of the wild type plant and plant 21 were similar, significant differences were observed in plants 4 and 29. These changes mirrored those found in TAG of both primary transformants.

Intriguingly, leaves of plants 4 and 29 before and during seed setting were characterized by a glossy surface, providing a phenotypic change that can serve as a phenotype that is easily scored visually, which could aid the timing of harvest for maximal oil content.

In summary, leaves of plants 4 and 29 rapidly accumulated TAG during flowering up till seed setting. At the latter stage, the majority of leaves exhibited TAG levels between 7 % and 13 % on a dry weight basis, compared to 0.1% - 0.2% for line 21. These observed TAG levels and total lipid levels far exceed the levels achieved by Andrianov et al., (2010) who reported up to a maximum of 5.8% and 6.8% TAG in leaves of *N. tabacum* upon constitutive expression of the *A. thaliana* DGAT1 and inducible expression of the *A. thaliana* LEC2 genes.

Table 30. Percentage TAG (% weight of leaf dry weight) and oleic acid levels (% of total fatty acids) in the TAG isolated from leaves of selected primary tobacco plants transformed with pJP3503

Plant No.	%TAG (DW)	% C18:1 ^{Δ9}	Development stage of plant
Wild type	0.06	2.3	Budding
Wild type	0.1	1.3	Flowering
3	0.05	1.5	Budding
4	1.29	10.2	Budding
11	0.21	7.4	Flowering
15	0.23	4.5	No buds
21	0.01	1.9	No buds
27	0.19	3.3	Budding
29	1.15	10.4	No buds

Table 31. TAG levels (% weight of leaf dry weight) and fatty acid composition of TAG isolated from wild-type and three selected tobacco plants transformed with pJP3503, before and during flowering. The data shown are averages and standard deviations of 2-3 independent repeats.

	Before flowering				Flowering			
	Wild type	Plant 21	Plant 4	Plant 29	Wild type	Plant 21	Plant 4	Plant 29
% /leaf DW	0.1	0.0	3.1 ± 0.3	4.1 ± 0.3	0.1	0.1 ± 0.0	7.3 ± 0.3	6.9 ± 0.5
C14:0	0.6	0.6 ± 0.2	0.1 ± 0.0	0.2 ± 0.0	0.5	0.4 ± 0.0	0.1 ± 0.0	0.1 ± 0.0
C16:0	9.1	15.9 ± 0.3	42.0 ± 0.5	34.9 ± 0.7	7.5	15.0 ± 0.7	33.1 ± 1.0	24.8 ± 1.4
C16:1 ^{Δ3t}	0.0	0.4 ± 0.1	0.1 ± 0.0	0.2 ± 0.0	0.4	0.2 ± 0.0	0.1 ± 0.0	0.2 ± 0.0
C16:1 ^{Δ9}	0.3	0.6 ± 0.0	3.3 ± 0.2	1.5 ± 0.1	0.3	0.6 ± 0.0	3.3 ± 0.2	1.2 ± 0.1
C16:3 ^{Δ7,12,15}	3.7	0.8 ± 0.1	0.1 ± 0.0	0.2 ± 0.0	3.6	0.8 ± 0.1	0.1 ± 0.0	0.2 ± 0.0
C18:0	3.2	4.6 ± 0.4	3.0 ± 0.1	5.6 ± 0.3	2.4	3.4 ± 0.0	3.3 ± 0.1	4.6 ± 0.1
C18:1 ^{Δ9}	2.3	1.2 ± 0.2	10.6 ± 0.3	10.6 ± 0.2	1.3	1.2 ± 0.2	19.1 ± 1.8	22.1 ± 2.7
C18:1 ^{Δ11}	0.1	0.1 ± 0.0	2.2 ± 0.2	1.2 ± 0.1	0.1	0.1 ± 0.0	2.1 ± 0.0	0.9 ± 0.0
C18:2 ^{Δ9,12}	26.9	20.4 ± 1.1	20.7 ± 0.7	30.0 ± 1.0	23.7	19.5 ± 0.8	25.3 ± 0.6	34.7 ± 1.0
C18:3 ^{Δ9,12,15}	52.6	54.0 ± 0.8	15.9 ± 0.5	11.5 ± 1.3	59.3	57.5 ± 0.4	10.8 ± 0.5	7.1 ± 0.6
C20:0	0.3	0.6 ± 0.0	1.0 ± 0.0	2.1 ± 0.2	0.3	0.4 ± 0.0	1.3 ± 0.0	2.0 ± 0.1
C20:1 ^{Δ11}	0.1	0.0	0.0	0.2 ± 0.1	0.1	0.1 ± 0.0	0.1 ± 0.0	0.3 ± 0.0
C20:2 ^{Δ11,14}	0.2	0.2 ± 0.0	0.0	0.1 ± 0.0	0.2	0.3 ± 0.0	0.0	0.1 ± 0.0
C20:3 ^{Δ11,14,17}	0.2	0.2 ± 0.0	0.0	0.0	0.1	0.2 ± 0.0	0.0	0.0
C22:0	0.1	0.2 ± 0.0	0.5 ± 0.0	1.0 ± 0.1	0.0	0.2 ± 0.0	0.7 ± 0.0	1.0 ± 0.0
C24:0	0.4	0.1 ± 0.1	0.4 ± 0.0	0.9 ± 0.1	0.3	0.1 ± 0.0	0.5 ± 0.0	0.7 ± 0.1

Table 32. Polar lipid levels (% weight of leaf dry weight) and fatty acid composition of polar lipids of leaf tissue from selected tobacco plants transformed with pJP3503, before and during flowering. Data shown are the average and standard deviations of 2-3 independent repeats, except for plant 21 (before flowering).

	Before flowering			Flowering		
	Plant 21	Plant 4	Plant 29	Plant 21	Plant 4	Plant 29
%/leaf DW	2.0	3.3 ± 0.4	2.5 ± 0.4	1.7 ± 0.0	2.4 ± 0.0	1.7 ± 0.1
C14:0	0.0	0.0	0.0	0.0	0.0	0.0
C16:0	12.0	18.7 ± 0.6	12.7 ± 0.4	12.7 ± 0.0	16.2 ± 0.2	11.8 ± 0.2
C16:1 ^{Δ3t}	1.6	0.5 ± 0.1	1.4 ± 0.1	1.4 ± 0.0	1.1 ± 0.2	1.4 ± 0.1
C16:1 ^{Δ9}	0.1	2.0 ± 0.1	0.7 ± 0.1	0.1 ± 0.0	1.8 ± 0.1	0.5 ± 0.0
C16:3 ^{Δ7,12,15}	7.5	2.4 ± 0.2	3.8 ± 0.1	6.3 ± 0.0	1.0 ± 0.1	3.7 ± 0.3
C18:0	2.1	1.1 ± 0.0	0.9 ± 0.1	2.5 ± 0.1	1.3 ± 0.0	1.4 ± 0.0
C18:1 ^{Δ9}	1.1	5.6 ± 0.5	4.1 ± 0.0	1.0 ± 0.1	11.9 ± 1.1	9.2 ± 0.9
C18:1 ^{Δ11}	0.1	2.2 ± 0.1	1.0 ± 0.1	0.1 ± 0.0	2.1 ± 0.1	0.6 ± 0.0
C18:2 ^{Δ9,12}	12.5	19.8 ± 1.0	19.7 ± 1.4	13.6 ± 0.1	27.6 ± 0.6	28.8 ± 0.5
C18:3 ^{Δ9,12,15}	62.3	46.8 ± 0.6	55.1 ± 0.8	61.3 ± 0.1	36.0 ± 1.1	41.6 ± 0.6
C20:0	0.3	0.3 ± 0.0	0.2 ± 0.0	0.3 ± 0.0	0.4 ± 0.0	0.4 ± 0.0
C20:1 ^{Δ11}	0.0	0.0	0.0	0.0	0.0	0.0
C20:2 ^{Δ11,14}	0.1	0.0	0.0	0.1 ± 0.0	0.0	0.0
C20:3 ^{Δ11,14,17}	0.1	0.0	0.0	0.1 ± 0.0	0.0	0.0
C22:0	0.2	0.2 ± 0.0	0.2 ± 0.0	0.2 ± 0.0	0.3 ± 0.0	0.3 ± 0.0
C24:0	0.2	0.2 ± 0.0	0.3 ± 0.0	0.2 ± 0.0	0.3 ± 0.0	0.3 ± 0.0

Table 33. Total lipid levels (% weight of leaf dry weight) and fatty acid composition of total lipids in leaves from tobacco plants transformed with pJP3503, just before and during flowering. Data shown are the average of 2-3 leaf samples.

	Before flowering			Flowering		
	Plant 21	Plant 4	Plant 29	Plant 21	Plant 4	Plant 29
%/leaf DW	2.4 ± 0.2	6.9 ± 0.5	4.9 ± 1.1	2.0 ± 0.1	9.8 ± 0.3	8.8 ± 0.3
C14:0	0.1 ± 0.0	0.1 ± 0.0	0.1 ± 0.0	0.1 ± 0.0	0.1 ± 0.0	0.1 ± 0.0
C16:0	12.1 ± 0.1	27.6 ± 1.2	20.7 ± 0.6	12.7 ± 0.1	26.9 ± 0.9	20.2 ± 1.2
C16:1 ^{Δ3t}	2.1 ± 0.2	0.4 ± 0.0	0.9 ± 0.1	2.7 ± 0.4	0.8 ± 0.1	0.7 ± 0.1
C16:1 ^{Δ9}	0.0	2.6 ± 0.2	1.0 ± 0.1	0.0	2.9 ± 0.1	1.0 ± 0.0
C16:3 ^{Δ7,12,15}	6.9 ± 0.3	1.3 ± 0.1	2.1 ± 0.1	5.6 ± 0.2	0.4 ± 0.0	0.9 ± 0.1
C18:0	2.2 ± 0.1	1.9 ± 0.0	2.8 ± 0.1	2.6 ± 0.1	2.6 ± 0.1	3.6 ± 0.1
C18:1 ^{Δ9}	1.1 ± 0.2	7.5 ± 0.5	6.6 ± 0.0	1.1 ± 0.2	16.2 ± 1.5	17.8 ± 2.0
C18:1 ^{Δ11}	0.2 ± 0.0	2.1 ± 0.1	1.1 ± 0.1	0.2 ± 0.0	2.0 ± 0.0	0.8 ± 0.0
C18:2 ^{Δ9,12}	13.8 ± 1.5	21.4 ± 0.7	26.5 ± 1.6	14.1 ± 0.5	27.3 ± 0.4	35.9 ± 0.5
C18:3 ^{Δ9,12,15}	60.6 ± 1.4	33.6 ± 1.5	36.0 ± 1.0	58.5 ± 1.2	18.3 ± 0.5	15.5 ± 0.6
C20:0	0.3 ± 0.0	0.7 ± 0.0	1.0 ± 0.1	0.3 ± 0.0	1.0 ± 0.0	1.5 ± 0.0
C20:1 ^{Δ11}	0.0	0.0	0.1 ± 0.0	0.0	0.1 ± 0.0	0.3 ± 0.0
C20:2 ^{Δ11,14}	0.1 ± 0.0	0.0	0.1 ± 0.0	0.2 ± 0.0	0.0	0.0
C20:3 ^{Δ11,14,17}	0.1 ± 0.0	0.0	0.0	0.3 ± 0.0	0.6 ± 0.0	0.8 ± 0.0
C22:0	0.2 ± 0.0	0.3 ± 0.0	0.5 ± 0.1	1.3 ± 0.1	0.3 ± 0.0	0.2 ± 0.0
C24:0	0.2 ± 0.0	0.3 ± 0.0	0.5 ± 0.0	0.3 ± 0.0	0.5 ± 0.0	0.6 ± 0.0

Table 34. TAG levels (% weight of leaf dry weight) and fatty acid composition of TAG isolated from different aged leaves, post flowering, of three selected tobacco plants transformed with pJP3503.

Plant: ^a	Wild type			Plant 21			Plant 4			Plant 29		
	G	YG		G	YG		G	YG	Y	G	YG	Y ^c
Leaf stage ^b												
%/leaf DW	0.1	0.1		0.1	0.2		9.5	13.0	10.7	7.0	7.1	2.1
C14:0	0.5	0.3		0.4	0.3		0.2	0.1	0.1	0.2	0.2	0.3
C16:0	7.5	14.8		8.9	14.9		31.1	33.3	38.0	25.7	33.0	38.5
C16:1 ^{A31}	0.4	0.3		0.3	0.2		0.1	0.1	0.1	0.2	0.2	0.3
C16:1 ^{A9}	0.3	0.2		0.2	0.2		3.0	3.1	2.4	1.2	1.1	0.7
C16:3 ^{A7,12,15}	3.6	0.6		3.2	1.2		0.1	0.1	0.1	0.3	0.3	0.2
C18:0	2.4	3.3		2.7	3.9		3.2	2.8	2.8	4.6	4.5	4.1
C18:1 ^{A9}	1.3	0.8		2.3	0.7		17.6	21.2	21.7	16.4	11.1	10.7
C18:1 ^{A11}	0.1	0.1		0.1	0.1		2.1	1.8	1.5	0.9	0.8	0.7
C18:2 ^{A9,12}	23.7	17.5		24.3	15.2		28.2	22.0	17.0	36.5	32.6	24.6
C18:3 ^{A9,12,15}	59.3	60.9		56.4	62.2		11.3	12.8	13.2	9.5	11.9	14.8
C20:0	0.3	0.4		0.3	0.5		1.4	1.2	1.3	2.1	2.1	2.1
C20:1 ^{A11}	0.1	0.1		0.1	0.0		0.1	0.1	0.1	0.3	0.2	0.1
C20:2 ^{A11,14}	0.2	0.2		0.2	0.1		0.1	0.0	0.0	0.1	0.1	0.0
C20:3 ^{A11,14,17}	0.1	0.2		0.1	0.1		0.0	0.0	0.0	0.0	0.0	0.0
C22:0	0.0	0.1		0.1	0.1		0.8	0.7	0.9	1.2	1.2	1.5
C24:0	0.3	0.2		0.2	0.2		0.7	0.5	0.7	0.9	0.8	1.4

^a: Leaf samples were taken from wild type at flowering stage and from the three pJP3503 primary transformants during seed setting.

^b: leaf stages by colour indicated by 'G', green; 'YG', yellow-green; 'Y', yellow

^c: very old leaf

Table 35. Total lipid yield (% weight of leaf dry weight) and fatty acid composition of total lipid isolated from differently aged leaves of wild-type and three selected tobacco plants transformed with pJP3503.

Plant:	Wild type			Plant 21			Plant 4			Plant 29			
Leaf stage ^a	G ^b	G ^c	YG	G	YG	G	YG	G	YG	Y	G	YG	Y
%/leaf DW	2.4	1.8	1.4	2.3	2.1	11.6	15.8	13.0	8.8	3.7			
C14:0	0.1	0.2	0.1	0.2	0.1	0.2	0.1	0.1	0.1	0.2			
C16:0	11.6	11.9	16.0	11.9	13.7	26.6	30.0	34.6	28.0	29.4			
C16:1 ^{Δ3t}	4.1	6.3	3.2	3.1	2.0	0.5	0.4	0.5	0.6	0.9			
C16:1 ^{Δ9}	0.0	0.0	0.0	0.0	0.0	3.0	2.9	2.3	1.1	0.7			
C16:3 ^{Δ7,12,15}	6.7	5.3	3.6	5.6	5.3	0.4	0.3	0.3	0.9	1.5			
C18:0	2.4	2.9	4.1	3.0	3.2	2.9	2.6	2.8	4.0	3.7			
C18:1 ^{Δ9}	1.4	1.2	0.9	1.4	0.5	15.8	20.2	20.8	10.2	11.3			
C18:1 ^{Δ11}	0.5	0.9	0.4	0.4	0.2	2.1	1.9	1.5	0.7	0.7			
C18:2 ^{Δ9,12}	16.0	15.5	16.5	15.6	12.4	28.9	23.3	18.6	33.4	26.8			
C18:3 ^{Δ9,12,15}	54.4	51.4	52.3	57.0	60.9	16.7	15.8	15.3	17.1	20.3			
C20:0	0.5	0.6	0.8	0.4	0.5	1.2	1.0	1.3	1.7	1.6			
C20:1 ^{Δ11}	0.1	0.0	0.0	0.0	0.0	0.1	0.1	0.1	0.2	0.1			
C20:2 ^{Δ11,14}	0.1	0.1	0.1	0.1	0.1	0.0	0.0	0.0	0.1	0.0			
C20:3 ^{Δ11,14,17}	0.1	0.1	0.1	0.1	0.1	0.0	0.0	0.0	0.0	0.0			
C22:0	0.3	0.4	0.5	0.2	0.3	0.7	0.7	0.9	1.1	1.3			
C24:0	0.3	0.3	0.4	0.2	0.3	0.6	0.5	0.7	0.8	1.2			

^a: samples taken from plants harbouring multiple seed pods unless indicated otherwise, ^b: before flowering, ^c: during flowering

Analysis of tobacco plants transformed with pJP3502

For the transformation with pJP3502 ("3 gene construct"), the sucrose level in the MS agar medium was reduced to half the standard level until sufficient calli were established, which aided the recovery of transformants expressing WRI1. Forty-one
5 primary transformants were obtained from the transformation with pJP3502 and transferred to the greenhouse. Leaf samples of different age were collected at either flowering or seed setting stages (Table 36). The plants look phenotypically normal except for three transformants, originating from the same callus in the transformation procedure and therefore likely to be from the same transformation event, which were
10 slightly smaller and displayed a glossy leaf phenotype similar to that observed for plant 4 with pJP3503 (above) but less in extent.

Leaf disk samples of primary transformants were harvested during flowering and TAG was quantified visualized by iodine staining after TLC. Selected transgenic plants displaying increased TAG levels compared to the wild type controls were
15 further analyzed in more detail by TLC and GC. The highest TAG level in young green leaves was detected in line 8.1 and corresponded to 8.3 % TAG on a dry weight basis or an approximate 83-fold increase compared to wild type leaves of the same age (Table 36). Yellow-green leaves typically contained a higher oil content compared to younger green leaves with maximum TAG levels observed in line 14.1
20 (17.3 % TAG on a dry weight basis). Total lipid content and fatty acid composition of total lipid in the leaves was also quantitated (Table 37).

Seed (T1 seed) was collected from the primary transformants at seed maturity and some were sown to produce T1 plants. These plants were predicted to be segregating for the transgene and therefore some null segregants were expected in the
25 T1 populations, which could serve as appropriate negative controls in addition to known wild-type plants which were grown at the same time and under the same conditions. 51 T1 plants, derived from primary transformant 14.1 which had a single-copy T-DNA insertion, which were 6-8 weeks of age and 10-25 cm in height were analysed together with 12 wild-type plants. The plants appeared phenotypically
30 normal, green and healthy, and did not appear smaller than the corresponding wild-type plants. Leaf samples of about 1 cm diameter were taken from fully expanded green leaves. 30 of the T1 plants showed elevated TAG levels in the leaves, of which 8 plants showed high levels of TAG, about double the level of TAG compared to the primary transformant 14.1 at the same stage of plant development. These latter plants
35 are likely to be homozygous for the transgenes. The level of TAG and the TAG fatty acid composition in leaves of selected T1 plants were measured by loading lipid isolated from about 5 mg dry weight of leaf tissue onto each TLC lane, the data is shown in Table 38.

Table 36. TAG levels (% weight of leaf dry weight) and fatty acid composition of TAG isolated from green leaves from ten selected tobacco plants transformed with pJP3502.

Line	2.1a	8.1	8.1	10.1	10.1	10.2	10.2	10.2	13.4	14.1	14.1	14.1	14.1	14.2	14.2	14.3	14.3	14.3	19.1	19.1	19.1	19.2	19.2
Stageb	Y	F	Y	Y	F	Y	S	Y	F	S	S	Y	Y	S	Y	F	Y	F	Y	F	Y	F	Y
TAG (% dry weight)	2.5	8.3	7.2	11.8	6.2	7.6	6.4	6.4	5.7	4.6	4.6	17.3	2.5	12.6	3	2	5.1	5.5	4.3	5.5	4.3	3.2	3.2
C14:0	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.2	0.2	0.2	0.1	0.2	0.1	0.1	0.1	0.2	0.1	0.1	0.1	0.2	0.1	0.1
C16:0	26.4	28.5	31.4	22.6	24.0	26.9	29.7	28	28.2	28.6	20.1	27.3	19.4	30.4	23.2	34.5	25.4	36.7					
C16:1Δ3	0.4	0.3	0.4	0.3	0.3	0.3	0.4	0.3	0.3	0.3	0.4	0.5	0.5	0.3	0.2	0.0	0.3	0.4	0.3	0.4	0.3	0.2	0.2
C16:1Δ9	0.6	1.8	1.0	1.4	1.1	1.4	1.0	2	1.7	1.6	1.1	1.3	2.2	2.9	1.4	0.9	1.3	0.8					
C16:3Δ7,12,15	0.3	0.4	0.4	0.3	0.4	0.3	0.4	0.2	0.3	0.2	0.3	0.2	0.5	0.3	0.3	0.1	0.2	0.3	0.4	0.3	0.4	0.3	0.3
C18:0	7.1	3.8	4.0	3.8	3.5	4	4.0	4.2	4.1	4.1	4.5	3.9	3.3	3.0	4	4.2	4.3	4.3					
C18:1Δ9	9.6	18.1	11.2	33.1	25.1	21.3	10.4	24.8	18.6	25.8	16.1	17.3	14	1.9	25.5	8.5	20.2	10.4					
C18:1Δ11	0.5	1.2	0.8	1.3	0.8	1	0.7	1.2	1.1	1.0	0.8	0.8	1	0.7	1.3	0.6	1.4	0.7					
C18:2Δ9,12	31.3	30.5	32.0	28	31.3	32.3	37.3	25.8	30.7	24.3	37.3	33.8	35.5	15.3	33.5	30.1	33.5	20.8					
C18:3Δ9,12,15	17.9	10.8	13.2	5.9	10.2	8.1	10.2	8.8	9.4	9.8	14.6	10.2	21.5	44.0	6.5	15.7	8.5	21.1					
C20:0	3	2	2.3	1.5	1.6	2	2.5	2	2.4	2.0	2.1	2.2	1.2	0.9	1.7	2.2	2	2.0					
C20:1Δ11	0.3	0.3	0.3	0.3	0.0	0.3	0.3	0.3	0.4	0.0	0.5	0.0	0.3	0.0	0.4	0.0	0.4	0.2					
C20:2Δ11,14	0.1	0.1	0.1	0	0.1	0.1	0.1	0	0.1	0.1	0	0.1	0.1	0.1	0.1	0.0	0.1	0.0					
C20:3Δ11,14,17	0	0	0.0	0	0.0	0	0.0	0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0					
C22:0	1.6	1.2	1.7	0.8	1.0	1.2	1.8	1.1	1.5	1.1	1	1.4	0.5	0.3	0.9	1.4	1.1	1.2					
C24:0	0.9	0.9	1.1	0.6	0.7	0.8	1.2	0.9	1.1	1.0	0.8	1.1	0.4	0.3	0.7	1.1	0.9	1.1					

^a very old plant containing only yellow leaves

^b seed setting ('S'), flowering ('F'), yellow or yellowing leaves ('Y')

Table 37. Total lipid and fatty acid composition of total lipid extracted from yellow leaves from selected tobacco plants transformed with pIP3502 (DGAT1+WR11+Oleosin)

Line	C14:0	C16:0	C16d3	16:1w13t	C16:1	C16:3	C18:0	C18:1	C18:1d11	C18:2	C18:3n3 DW	C20:0	C20:1d11	C20:2n6	C20:3n3	C22:0	C24:0	mg / 100 mg
14.1	0.1	26.6	0.0	0.6	1.6	0.3	3.8	25.6	1.1	25.2	11.0	1.8	0.3	0.1	0.0	1.0	0.9	23.4
14.2	0.1	25.2	0.0	0.6	1.3	0.4	3.6	16.5	0.8	35.1	11.6	1.9	0.3	0.1	0.0	1.5	1.0	15.5
10.1	0.1	22.4	0.0	0.7	1.3	0.6	3.2	24.3	0.8	31.9	11.4	1.5	0.3	0.0	0.0	0.9	0.6	9.7
13.4	0.2	19.6	0.0	3.6	0.0	1.7	4.1	3.9	0.6	34.6	27.7	1.6	0.3	0.0	0.0	1.5	0.8	2.0

Table 38. TAG levels (% weight of leaf dry weight) and fatty acid composition of TAG isolated from green leaves from selected tobacco T1 plants transformed with pJP3502.

Line No.	16:0	16:1	18:0	18:1d9	18:1d11	18:2	18:3n3	20:0	22:0	24:0	% TAG
4.1	23.3	0.9	3.3	11.3	1.5	44.8	10.7	1.9	1.4	0.9	3.7
4.2	22.4	0.9	3.1	13.3	1.5	44.3	10.5	1.8	1.3	0.9	6.1
4.3	24.2	1.0	3.4	14.9	1.6	41.8	9.0	1.9	1.3	0.9	4.0
6.1	24.6	0.9	3.9	16.5	1.4	33.4	14.4	2.1	1.5	1.3	3.5
6.2	22.8	0.8	3.8	18.6	1.4	32.7	15.1	2.1	1.5	1.1	3.4
6.3	26.6	1.0	4.2	16.2	1.5	31.0	15.6	2.3	1.6	0.0	2.3
8.1	27.0	1.0	4.8	18.3	1.5	27.1	15.9	2.6	1.9	0.0	1.5
8.2	24.2	1.0	4.3	19.0	1.4	27.6	16.7	2.5	1.8	1.5	2.2
8.3	26.5	1.3	4.8	22.0	1.7	24.9	16.2	2.6	0.0	0.0	1.2
13.1	33.7	1.3	5.0	7.4	1.3	34.2	14.2	2.8	0.0	0.0	1.2
13.2	29.1	1.1	4.4	7.2	1.3	37.2	17.1	2.6	0.0	0.0	1.6
13.3	34.3	0.0	5.5	6.9	0.0	36.4	13.7	3.2	0.0	0.0	0.8
21.1	27.4	0.7	4.2	8.5	1.2	37.1	15.4	2.4	1.6	1.4	2.1
21.2	29.7	0.9	4.5	9.1	1.3	36.3	15.9	2.3	0.0	0.0	1.6
21.3	27.1	0.8	4.3	12.7	1.4	37.1	13.0	2.2	1.4	0.0	2.4
29.1	27.2	0.8	4.3	12.8	1.1	34.9	14.4	2.1	1.4	1.0	3.9
29.2	26.9	1.0	4.1	14.7	1.2	35.3	13.0	2.0	1.2	0.8	3.8
29.3	25.7	1.4	4.1	18.4	1.2	35.7	11.0	1.6	1.0	0.0	3.7
23.1	29.9	0.9	4.3	8.1	1.2	35.0	18.4	2.1	0.0	0.0	1.6
23.2	30.8	1.0	4.8	9.3	1.3	33.5	17.2	2.3	0.0	0.0	1.5
23.3	29.2	0.9	4.3	9.1	1.3	36.2	15.2	2.3	1.5	0.0	2.3
49.1	27.0	0.9	3.9	5.8	1.4	43.8	11.0	2.5	2.1	1.5	2.4
49.2	27.5	0.9	3.8	7.1	1.5	44.7	10.2	2.4	2.0	0.0	2.2

Genetic constructs suitable for transformation of monocotyledonous plants are made by exchanging the Arath-SSU promoters in pJP3502 and pJP3503 for promoters more active in monocots. Suitable promoters include constitutive viral promoters from monocot viruses or promoters that have demonstrated to function in a transgenic
5 context in monocot species (e.g., the maize Ubi promoter described by Christensen et al., 1996). Similarly, the CaMV-35S promoters in pJP3502 and pJP3503 are exchanged for promoters that are more active in monocot species. These constructs are transformed into wheat, barley and maize using standard methods.

10 Miscanthus species

Genetic constructs for *Miscanthus* species transformation are made by exchanging the Arath-SSU promoters in pJP3502 and pJP3503 for promoters more active in *Miscanthus*. Suitable promoters include constitutive viral promoters, a ubiquitin promoter (Christensen et al., 1996) or promoters that have demonstrated to
15 function in a transgenic context in *Miscanthus*. Similarly, the CaMV-35S promoters in pJP3502 and pJP3503 are exchanged for promoters that are more active in *Miscanthus*. New constructs are transformed in *Miscanthus* by a microprojectile-mediated method similar to that described by Wang et al., 2011.

20 Switchgrass (*Panicum virgatum*)

Genetic constructs for switchgrass transformation are made by exchanging the Arath-SSU promoters in pJP3502 and pJP3503 for promoters more active in switchgrass. Suitable promoters include constitutive viral promoters or promoters that have demonstrated to function in a transgenic context in switchgrass (e.g., Mann et
25 al., 2011). Similarly, the CaMV-35S promoters in pJP3502 and pJP3503 are exchanged for promoters that are more active in switchgrass. New constructs are transformed in switchgrass by an *Agrobacterium*-mediated method similar to that described by Chen et al., 2010 and Ramamoorthy and Kumar, 2012.

30 Sugarcane

Genetic constructs for sugarcane transformation are made by exchanging the Arath-SSU promoters in pJP3502 and pJP3503 for promoters more active in sugarcane. Suitable promoters include constitutive viral promoters or promoters that have demonstrated to function in a transgenic context in sugarcane (e.g., the maize
35 Ubi promoter described by Christensen et al., 1996). Similarly, the CaMV-35S promoters in pJP3502 and pJP3503 are exchanged for promoters that are more active in sugarcane. New constructs are transformed in sugarcane by a microprojectile-mediated method similar to that described by Bower et al., 1996.

Elephant grass

Genetic constructs for *Pennisetum purpureum* transformation are made by exchanging the Arath-SSU promoters in pJP3502 and pJP3503 for promoters more active in elephant grass. Suitable promoters include constitutive viral promoters or promoters that have demonstrated to function in a transgenic context in *Pennisetum* species such as *P. glaucum* like (e.g. the maize Ubi promoter described by Christensen et al., 1996). Similarly, the CaMV-35S promoters in pJP3502 and pJP3503 are exchanged for promoters that are more active in *Pennisetum* species.

10 New constructs are transformed in *P. purpureum* by a microprojectile-mediated method similar to that described by Girgi et al., 2002.

Lolium

Genetic constructs for *Lolium perenne* and other *Lolium* species transformation are made by exchanging the Arath-SSU promoters in pJP3502 and pJP3503 for promoters more active in ryegrass. Suitable promoters include constitutive viral promoters or promoters that have demonstrated to function in a transgenic context in *Lolium* species (e.g. the maize Ubi promoter described by Christensen et al., 1996). Similarly, the CaMV-35S promoters in pJP3502 and pJP3503 are exchanged for promoters that are more active in *Pennisetum* species.

20 New constructs are transformed in *Lolium perenne* by a silicon carbide-mediated method similar to that described by Dalton et al., 2002 or an *Agrobacterium*-mediated method similar to that described by Bettany et al., 2003.

pJP3502 and pJP3503 are modified to seed-specific expression genetic constructs by exchanging the CaMV-35S and Arath-SSU promoters (except the selectable marker cassette) with seed-specific promoters active in the target species.

25

Canola

Genetic constructs for *Brassica napus* transformation are made by exchanging the CaMV-35S and Arath-SSU promoters in pJP3502 and pJP3503 for promoters more active in canola. Suitable promoters include promoters that have previously been demonstrated to function in a transgenic context in *Brassica napus* (e.g., the *A. thaliana* FAE1 promoter, *Brassica napus* napin promoter, *Linum usitatissimum* conlinin1 and conlinin2 promoters). New constructs are transformed in *B. napus* as previously described.

30

35

Soybean (*Glycine max*)

A genetic construct is made by cloning the *Psp*OMI fragment from a synthesised DNA fragment having the nucleotide sequence shown in SEQ ID NO:415 (Soybean synergy insert; Figure 19A) into a binary vector such as pORE04 at the *Not*I site. This fragment contains *Arath*-WRI1 expressed by a *Arath*-FAE1 promoter, *Arath*-DGAT1 expressed by a *Linus*-Cn12 promoter, *Musmu*-MGAT2 expressed by *Linus*-Cn11 and *Arath*-GPAT4 expressed by *Linus*-Cn11. A further genetic construct is made by exchanging the GPAT coding region for an oleosin coding region. A further genetic construct is made by deleting the MGAT expression cassette.

- 10 A genetic construct, pJP3569 (Figure 21), was generated by cloning the *Sbf*I-*Pst*I fragment from the DNA molecule having the nucleotide sequence shown in SEQ ID NO:415 into the *Pst*I site of pORE04. This construct contained (i) a coding region encoding the *A. thaliana* WRI1 transcription factor, codon optimised for *G. max* expression, and expressed from the *G. max* kunitz trypsin inhibitor 3 (*Glyma*-KTI3) promoter, (ii) a coding region encoding the *Umbelopsis ramanniana* DGAT2A (codon optimised as described by Lardizabal et al., 2008) and expressed from the *G. max* alpha-subunit beta-conglycinin (*Glyma*-b-conglycinin) promoter and (iii) a coding region encoding the *M. musculus* MGAT2, codon optimised for *G. max* expression. A second genetic construct, pJP3570, was generated by cloning the *Sbf*I-*Swa*I fragment of the DNA molecule having the nucleotide sequence shown in SEQ ID NO:415 into pORE04 at the *Eco*RV-*Pst*I sites to yield a binary vector containing genes expressing the *A. thaliana* WRI1 transcription factor and *U. ramanniana* DGAT2A enzyme. Similarly, a third genetic construct, pJP3571, was generated by cloning the *Asi*SI fragment of the DNA molecule having the nucleotide sequence shown in SEQ ID NO:415 into the *Asi*SI site of pORE04 to yield a binary vector containing a gene encoding the *U. ramanniana* DGAT2A enzyme. A fourth genetic construct, pJP3572, was generated by cloning the *Not*I fragment of the DNA molecule having the nucleotide sequence shown in SEQ ID NO:415 into pORE04 at the *Not*I site to yield a binary vector containing a gene expressing the *A. thaliana* WRI1 transcription factor.
- 25 A fifth genetic construct, pJP3573, was generated by cloning the *Swa*I fragment of the DNA molecule having the nucleotide sequence shown in SEQ ID NO:415 into pORE04 at the *Eco*RV site to yield a binary vector containing the gene encoding *M. musculus* MGAT2.

- 35 A sixth genetic construct, pJP3580, is generated by replacing the *M. musculus* MGAT2 with the *Sesamum indicum* oleosin gene.

Each of these six constructs are used to transform soybean, using the methods as described in Example 6. Transgenic plants produced by the transformation with each of the constructs, particularly pJP3569, produce seeds with increased oil content.

Sugarbeet

The vectors pJP3502 and pJP3503 (see above) as used for the transformation of tobacco are used to transform plants of sugarbeet (*Beta vulgaris*) by Agrobacterium-mediated transformation as described by Lindsey & Gallois (1990). The plants produce greatly increased levels of TAG in their leaves, similar in extent to the tobacco plants produced as described above. Transgenic sugarbeet plants are harvested while the leaves are still green or preferably green/yellow just prior to beginning of senescence or early in that developmental process, i.e., and while the sugar content of the beets is at a high level and after allowing accumulation of TAG in the leaves. This allows the production of dual-purpose sugarbeets which are suitable for production of both sugar from the beets and lipid from the leaves; the lipid may be converted directly to biodiesel fuel by crushing the leaves and centrifugation of the resultant material to separate the oil fraction, or the direct production of hydrocarbons by pyrolysis of the leaf material.

Promoters that are active in the root (tuber) of sugarbeet are also used to express transgenes in the tuber.

Example 21. Stable transformation of *Solanum tuberosum* with oil increase genes

pJP3502, the intermediate binary expression vector described in the previous example, was modified by first excising one SSU promoter by *AscI*+*NcoI* digestion and replacing it with the potato B33 promoter flanked by *AscI* and *NcoI* to generate pJP3504. The SSU promoter in pJP3504 along with a fragment of the *A. thaliana* WRL1 gene was replaced at the *PspOMI* sites by a potato B33 promoter with the same *A. thaliana* WRL1 gene fragment flanked by *NotI*-*PspOMI* to generate pJP3506. The pJP3347 was added to pJP3506 as described in the above example to generate pJP3507. This construct is shown schematically in Figure 20. Its sequence is given in SEQ ID NO:413. The construct is used to transform potato (*Solanum tuberosum*) to increase oil content in tubers.

Example 22. GPAT-MGAT fusion enzymes

The enzyme activity of GPAT-MGAT enzyme fusions is tested to determine whether this would increase the accessibility of the GPAT-produced MAG for MGAT activity. A suitable linker region was first synthesised and cloned into a cloning vector. This linker contained suitable sites for cloning the N-terminal (*EcoRI*-*ZraI*) and C-terminal coding regions (*NdeI*-*SmaI* or *NdeI*-*PstI*).

atttaaagcggccgcgaattcgctcgattgaggacgtccctactagacctgctggacctcctcctgctacttactacgattctct
cgctgtgcatatggcagtcagtcagccgggctgcaggcggccgcatttaa (SEQ ID NO:414)

A GPAT4-MGAT2 fusion (GPAT4 N-terminus and MGAT2 C-terminus) was
5 made by first cloning a DNA fragment encoding the *A. thaliana* GPAT4, flanked by
MfeI and *ZraI* sites and without a C-terminal stop codon, into the *EcoRI*-*ZraI* sites.
The DNA fragment encoding the *M. musculus* MGAT2, flanked by *NdeI*-*PstI* sites,
was then cloned into the *NdeI*-*PstI* sites to generate a single GPAT4-MGAT2 coding
sequence. The fused coding sequence was then cloned as a *NotI* fragment into pYES2
10 to generate pYES2::GPAT4-MGAT2 and the constitutive binary expression vector
pJP3343 to generate pJP3343::GPAT4-MGAT2.

Similarly, a MGAT2-GPAT4 fusion (MGAT2 N-terminus and GPAT4 C-
terminus) was made by first cloning the DNA fragment encoding *M. musculus*
MGAT2, flanked by *EcoRI* and *ZraI* sites without a C-terminal stop codon, into the
15 *EcoRI*-*ZraI* sites. The DNA fragment encoding the *A. thaliana* GPAT4, flanked by
NdeI-*PstI* sites, was then cloned into the *NdeI*-*PstI* sites to generate a single MGAT2-
GPAT4 coding sequence. The fused coding sequence was then cloned as a *NotI*
fragment into pYES2 to generate pYES2::MGAT2-GPAT4 and the constitutive
binary expression vector pJP3343 to generate pJP3343::MGAT2-GPAT4.

20 The yeast expression vectors are tested in yeast *S. cerevisiae* and the binary
vectors are tested in *N. benthamiana* and compared for oil content and composition
with single-coding region controls.

Example 23. Discovery of novel WRL1 sequences

25 Three novel WRL1 sequences are cloned into pJP3343 and other suitable
binary constitutive expression vectors and tested in *N. benthamiana*. These include
the genes encoding Sorbi-WRL1 (from *Sorghum bicolor*; SEQ ID NO:334), Lupan-
WRL1 (from *Lupinus angustifolius*; SEQ ID NO:335) and Ricco-WRL1 (from
Ricinus communis; SEQ ID NO:336). These constructs are tested in comparison with
30 the *Arabidopsis* WRI1-encoding gene in the *N. benthamiana* leaf assay.

As an initial step in the procedure, a partial cDNA fragment corresponding to
the WRL1 was identified in the developing seed EST database of *Lupinus*
angustifolius (NA-080818_Plate14f06.b1, SEQ ID NO:277). A full-length cDNA
(SEQ ID NO:278) was subsequently recovered by performing 5'- and 3'- RACE PCR
35 using nested primers and cDNAs isolated from developing seeds of *Lupinus*
angustifolius. The full length cDNA was 1729 bp long, including a 1284 bp protein
coding sequence encoding a predicted polypeptide of 428 amino acids (SEQ ID
NO:337). The entire coding region of the full length lupin WRL1 cDNA was then

PCR amplified using forward and reverse primers which both incorporated *EcoRI* restriction sites to facilitate the cloning into the pJP3343 vector under the control of a 35S promoter in the sense orientation. *A. tumefaciens* strain AGL1 harbouring the pJP3343-LuangWRL1 was infiltrated in *N. benthamiana* leave tissues as described in

- 5 Example 1. Leaf discs transiently expressing the pJP3343-LuangWRL1 were then harvested and analysed for oil content.

Example 24. Silencing of the CGI-58 homologue in *N. tabacum*

- James et al. (2010) have reported that the silencing of the *A. thaliana* CGI-58
10 homologue resulted in up to 10-fold TAG accumulation in leaves, mainly as lipid droplets in the cytosol. Galactolipid levels were also found to be higher, whereas levels of most major phospholipid species remained unchanged. Interestingly, TAG levels in seeds were unaffected and, unlike other TAG degradation mutants, no negative effect on seed germination was observed.

- 15 Three full length and two partial transcripts were found in the *N. benthamiana* transcriptome showing homology to the *A. thaliana* CGI-58 gene. A 434 bp region present in all five transcripts was amplified from *N. benthamiana* isolated leaf RNA and cloned via LR cloning (Gateway) into the pHELLSGATE12 destination vector. The resulting expression vector designated pTV46 encodes a hairpin RNA (dsRNA)
20 molecule for reducing expression of the tobacco gene encoding the homologue of CGI-58 and was used to transform *N. tabacum* as described in Example 1, yielding 52 primary transformants.

Primary transformants displaying increased TAG levels in their vegetative tissues are crossed with homozygous lines described in Example 20.

25

Example 25. Silencing of the *N. tabacum* ADP-glucose pyrophosphorylase (AGPase) small subunit

- Sanjaya et al. (2011) demonstrated that silencing of the AGPase small subunit in combination with WRI over-expression further increases TAG accumulation in *A.*
30 *thaliana* seedlings while starch levels were reduced. An AGPase small subunit has been cloned from flower buds (Kwak et al., 2007). The deduced amino acid sequence showed 87 % identity with the *A. thaliana* AGPase. A 593 bp fragment was synthesized and cloned into pHELLSGATE12 via LR cloning (Gateway) resulting in the binary vector pTV35. Transformation of *N. tabacum* was done as described in
35 Example 1 and yielded 43 primary transformants.

Primary transformants displaying a reduction in total leaf starch levels are crossed with homozygous lines described in Examples 20 and 21. In addition,

primary transformants are crossed with homozygous lines that are the result of a crossing of the lines described in 20 and 21.

Example 26. Production and use of constructs for gene combinations including an inducible promoter

Further genetic constructs are made using an inducible promoter system to drive expression of at least one of the genes in the combinations of genes as described above, particularly in pJP3503 and pJP3502. In the modified constructs, the WRI1 gene is expressed by an inducible promoter such as the *Aspergillus niger* alcA promoter in the presence of an expressed *Aspergillus niger* alcR gene. Alternatively, a DGAT is expressed using an inducible promoter. This is advantageous when maximal TAG accumulation is not desirable at all times during development. An inducible promoter system or a developmentally-controlled promoter system, preferably to drive the transcription factor such as WRI1, allows the induction of the high TAG phenotype at an appropriate time during development, and the subsequent accumulation of TAG to high levels.

TAG can be further increased by the co-expression of transcription factors including embryogenic transcription factors such as LEC2 or BABY BOOM (BBM, Srinivasan et al., 2007). These are expressed under control of inducible promoters are described above and super-transformed on transgenic lines or co-transformed with WRI and DGAT.

pJP3590 is generated by cloning a MAR spacer as a *AatII* fragment into the *AatII* site of pORE04. pJP3591 is generated by cloning a second MAR spacer as an *KpnI* fragment into the *KpnI* site of pJP3590. pJP3592 is generated by cloning the *AsiSI-SmaI* fragment of the DNA molecule having the nucleotide sequence shown in SEQ ID NO:416 (12ABFJYC_pJP3569_insert; Figure 19B) into the *AsiSI-EcoRV* sites of pJP3591. pJP3596 is generated by cloning a *PstI*-flanked inducible expression cassette containing the alcA promoter expressing the *M. musculus* MGAT2 and a *Glycine max* lectin polyadenylation signal into an introduced *SbfI* site in pJP3592. Hygromycin-resistant versions of both pJP3592 and pJP3596 (pJP3598 and pJP3597, respectively) are generated by replacing the NPTII selectable marker gene with the HPH flanked gene at the *FseI-AscI* sites.

These constructs are used to transform the same plant species as described in Example 20. Expression from the inducible promoter is increased by treatment with the inducer of the transgenic plants after they have grown substantially, so that they accumulate increased levels of TAG. These constructs are also super-transformed in stably transformed constructs already containing an oil-increase construct including the three-gene or four-gene TDNA region (SEQ ID NO:411 and SEQ ID NO:412,

respectively). Alternatively, the gene expression cassettes from the three-gene and four-gene constructs are cloned into the *NotI* sites of pJP3597 and pJP3598 to yield a combined constitutive and inducible vector system for high fatty acid and TAG synthesis, accumulation and storage.

- 5 In addition to other inducible promoters, an alternative is that gene expression can be temporally and spatially restricted by using promoters that are only active during specific developmental periods or in specific tissues. Endogenous chemically inducible promoters are also used to limit expression to specific developmental windows.

- 10 It will be appreciated by persons skilled in the art that numerous variations and/or modifications may be made to the invention as shown in the specific embodiments without departing from the spirit or scope of the invention as broadly described. The present embodiments are, therefore, to be considered in all respects as illustrative and not restrictive.

- 15 The present application claims priority from US 61/580590 filed 27 December 2011 and US 61/718,563 filed 25 October 2012, the entire contents of both of which are incorporated herein by reference.

All publications discussed and/or referenced herein are incorporated herein in their entirety.

- 20 Any discussion of documents, acts, materials, devices, articles or the like which has been included in the present specification is solely for the purpose of providing a context for the present invention. It is not to be taken as an admission that any or all of these matters form part of the prior art base or were common general knowledge in the field relevant to the present invention as it existed before the priority date of each claim of this application.

25

REFERENCES

- Abdullah et al. (1986) *Biotech.* 4:1087.
- Al-Mariri et al. (2002) *Infect. Immun.* 70:1915-1923.
- 5 Alemanno et al. (2008) *Planta* 227:853-866.
- Almeida and Allshire (2005) *TRENDS Cell Biol.* 15:251-258.
- Alonso et al. (2010) *Green Chem.* 12:1493-1513.
- Alvarez et al. (2000) *Theor. Appl. Genet.* 100:319-327.
- Andrianov et al. (2010) *Plant Biotech. J.* 8:277-287.
- 10 Barthole et al. (2011) *Plant Sci.* 185-186:33-39.
- Bartlett et al. (2008) *Plant Methods* 4:22.
- Baud et al. (2007) *Plant J.* 50:825-838.
- Baumlein et al. (1991) *Mol. Gen. Genet.* 225:459-467.
- Baumlein et al. (1992) *Plant J.* 2:233-239.
- 15 Benghezal et al. (2007) *J. Biol. Chem.* 282:30845-30855.
- Bettany et al. (2003) *Plant Cell Rep.* 21:437-444.
- Bligh and Dyer (1959) *Canadian Journal of Biochemistry and Physiology* 37:911-917.
- Bohner et al. (1999) *Plant J.* 19:87-95.
- Bourque (1995) *Plant Sci.* 105:125-149.
- 20 Boutilier et al. (2002) *Plant Cell* 14:1737-1749.
- Bouvier-Nave et al. (2000) *European Journal of Biochemistry / FEBS* 267:85-96.
- Bower et al. (1996) *Mol. Breed.* 2: 239-249.
- Bradford (1976) *Anal. Biochem.* 72:248-254.
- Broothaerts et al. (2005) *Nature* 433:629-633.
- 25 Broun et al. (1998) *Plant J.* 13:201-210.
- Bruce et al. (2000) *P. Plant Cell* 12:65-80.
- Buchanan-Wollaston (1994) *Plant Physiol.* 105:839-846.
- Busk et al. (1997) *Plant J.* 11:1285-1295.
- Cadwell and Joyce (1992) *PCR Methods Appl.* 2:28-33.
- 30 Cao et al. (2003) *J. Biol. Chem.* 278:13860-13866.
- Cao et al. (2007) *J. Lipid Res.* 48:583-591.
- Capuano et al. (2007) *Biotechnol. Adv.* 25:203-206.
- Cernac and Benning (2004) *Plant J.* 40:575-585
- Chen et al. (2010) *Biotechnol. for Biofuels.* 3:9.
- 35 Cheng et al. (1996) *Plant Cell Rep.* 15:653-657.
- Cheng et al. (2003) *J. Biol. Chem.* 278, 13611-13614.
- Chikwamba et al. (2003) *Proc. Natl. Acad. Sci. U.S.A.* 100:11127-11132.
- Christensen and Quail (1996) *Transgenic Res.* 5:213-218.

- Christie (1993) *Advances in Lipid Methodology* – Two, Oily Press, Dundee, pages 195-213.
- Chung et al. (2006) *BMC Genomics* 7:120.
- Clough and Bent (1998) *Plant J.* 16:735-743.
- 5 Coco et al. (2001) *Nature Biotechnology* 19:354–359.
- Coco et al. (2002) *Nature Biotechnology* 20:1246–1250.
- Comai et al. (2004) *Plant J* 37: 778-786.
- Corrado and Karali (2009) *Biotechnol. Adv.* 27:733-743.
- Courvalin et al. (1995) *Life Sci.* 318:1207-1212.
- 10 Coutu et al. (2007) *Transgenic Res.* 16:771-781.
- Cramer et al. (1998) *Nature* 391:288–291.
- Dalton et al. (1998) *Plant Science* 132:31–43.
- Dandik and Aksoy (1998) *Fuel Process Technol.* 57: 81-92.
- Darji et al. (1997) *Cell* 91:765-775.
- 15 Dauk et al (2007) *Plant Sci.* 173:43-49.
- Deshpande (1992) *Appl. Biochem. Biotechnol.* 36:227-234.
- Dhadialla et al. (1998) *Annu. Rev. Entomol.* 43:545-569.
- Dietrich et al. (1998) *Nature Biotech.* 18:181-185.
- Dulermo and Nicaud (2011) *Metab. Eng.* 13:482-491.
- 20 Durrett et al. (2008) *Plant J.* 54:593-607.
- Dyer et al. (2002) *Plant Physiol.* 130:2027-2038.
- Eggert et al. (2005) *Chembiochem* 6:1062–1067.
- Ellerstrom et al. (1996) *Plant Mol. Biol.* 32:1019-1027.
- Endalew et al. (2011) *Biomass and Bioenergy* 35:3787-3809.
- 25 Felenbok (1991) *J. Biotechnol.* 17:11-17.
- Fennelly et al. (1999) *J. Immunol.* 162:1603-1610.
- Froissard et al. (2009) *FEMS Yeast Res* 9:428-438.
- Fujimura et al. (1985) *Plant Tissue Culture Lett.* 2:74.
- Gan (1995) *Molecular characterization and genetic manipulation of plant senescence.* PhD thesis. University of Wisconsin, Madison.
- 30 Gan and Amasino (1995) *Science* 270:1986-1988
- Gatz (1997) *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 48:89-108.
- Ghosal et al. (2007) *Biochimica et Biophysica Acta* 1771:1457-1463.
- Ghosh et al. (2009) *Plant Physiol.* 151:869-881.
- 35 Girgi et al. (2002) *Molecular Breeding* 10:243-252.
- Glevin et al. (2003) *Microbiol. Mol. Biol. Rev.* 67:16-37.
- Goffman et al. (2005) *Plant Physiol.* 138:2269-2279.
- Gong and Jiang (2011) *Biotechnol. Lett.* 33:1269-1284.

- Gould et al. (1991) *Plant Physiol.* 95:426-434.
- Grant et al. (1995) *Plant Cell Rep.* 15:254-258.
- Greenwell et al. (2010) *J. R. Soc. Interface* 7:703-726.
- Grillot-Courvalin et al. (1999) *Curr. Opin. Biotech.* 10(5) :477-481.
- 5 Grillot-Courvalin et al. (1998) *Nature Biotech.* 16:862-866.
- Gurel et al. (2009) *Plant Cell Rep.* 28:429-444.
- Harayama (1998) *Trends Biotechnol.* 16: 76-82.
- Haseloff and Gerlach (1988) *Nature* 334:585-591.
- Hellings (1997) *Proc. Natl. Acad. Sci.* 94:10015-10017.
- 10 Henikoff et al. (2004) *Plant Physiol* 135: 630-636.
- Hense et al. (2001) *Cell Microbiol.* 3:599-609.
- Hershey and Stoner (1991) *Plant Mol. Biol.* 17:679-690.
- Hinchey et al. (1988) *Biotechnology* 6:915-922.
- Hirayama and Hujii (1965) *Agricultural and Biological Chemistry* 29:1-6.
- 15 Horvath et al. (2000) *Proc. Natl. Acad. Sci. U.S.A.* 97:1914-1919.
- Huang (1996) *Plant Physiol.* 110:1055-1061.
- Iwabuchi et al. (2003) *J. Biol. Chem.* 278:4603-4610.
- Izawati et al. (2009) *J. Oil Palm Res.* 21:643-652.
- James et al. (2010) *Proc. Natl. Acad. Sci. USA* 107:17833-17838.
- 20 Jepson et al. (1994) *Plant Mol. Biol.* 26:1855-1866.
- Jézéquel et al. (2008) *Biotechniques* 45:523-532.
- Joviet et al. (2004) *Plant Physiol. Biochem.* 42:501-509.
- Karmakar et al. (2010) *Bioresource Technology* 101:7201-7210.
- Kindle (1990) *Proc. Nat. Acad. Sci. USA* 87: 1228-1232.
- 25 Klein et al. (1998) *Exp. Neurol.* 150:183-194.
- Knothe (2005) *Fuel Process. Technol.* 86:1059-1070.
- Knothe and Steidley (2005) *Fuel* 84:1059-1065.
- Kozziel et al. (1996) *Plant Mol. Biol.* 32:393-405.
- Kuhn et al. (2009) *J. Biol. Chem.* 284:34092-102.
- 30 Kunik et al. (2001) *Proc. Natl. Acad. Sci. U.S.A.* 98:1871-1876.
- Kwak et al. (2007) *Plant Physiol.* 145:277-289.
- Lacroix et al. (2008) *Proc. Natl. Acad. Sci. U.S.A.* 105: 15429-15434.
- Lakshminarayana et al. (1984) *JAOCs* 61:1249-1253.
- Lam et al. (2010) *Biotechnol. Adv.* 28:500-18.
- 35 Lardizabal et al. (2001) *J. Biol. Chem.* 276:38862-38869.
- Lardizabal et al. (2008) *Plant Physiol.* 148: 89-96.
- Larkin et al. (1996) *Transgenic Res.* 5:325-335.
- Lee et al. (1998) *Science* 280:915-918.

- Leung et al. (1989) *Technique* 1:11-15.
- Li et al. (1996) *FEBS Lett.* 379:117-121.
- Li et al. (2006) *Phytochemistry* 67: 904-915.
- Li et al. (2007) *Proc. Natl. Acad. Sci. U S A.* 104:18339-44.
- 5 Lin et al. (2005) *Plant Physiol. Biochem.* 43:770-776.
- Lindsey and Gallois (1990) *Journal of Experimental Botany* 41:529-536.
- Liu et al. (2010a) *Fuel* 89:2735-2740.
- Liu et al. (2010b) *Plant Physiol. Biochem.* 48: 9-15.
- Lui et al. (2009) *J. Agric. Food Chem.* 57: 2308-2313.
- 10 Maher and Bressler (2007) *Bioresource Technology* 98:2351-2368.
- Mann et al. (2011) *BMC Biotechnol.* 11:74.
- Martinez and Jepson, Ecdysteroid agonist-inducible control of gene expression in plants. In *Inducible Gene Expression in Plants*. Edited by Reynolds PHS. New York: CABI Publishing; 1999:23-41.
- 15 Martinez et al. (1999) *Plant J.* 19:97-106.
- Matsuoka et al. (1994) *Plant J.* 6:311-319.
- Meier et al. (1997) *FEBS Lett.* 415:91-95.
- Mett et al. (1993) *Proc. Natl. Acad. Sci. USA* 90:4567-4571.
- Millar and Waterhouse (2005). *Funct Integr Genomics* 5:129-135.
- 20 Murashige and Skoog (1962) *Physiologia Plantarum* 15:473-497.
- Needleman and Wunsch (1970) *J. Mol Biol.* 45: 443-453
- Ness et al. (2002) *Nature Biotechnology* 20:1251-1255.
- Niedz et al. (1995) *Plant Cell Reports* 14:403-406.
- Ostermeier et al. (1999) *Nature Biotechnology* 17:1205-1209.
- 25 Ow et al. (1986) *Science* 234:856-859.
- Padidam (2003) *Curr. Opin. Plant Biol.* 6:169-77.
- Padidam et al. (2003) *Transgenic Res.* 12:101-9.
- Panekina (1978) *Chemistry of Natural Compounds* 14:33-36.
- Parthibane et al. (2012) *J. Biol. Chem.* 287:1946-54.
- 30 Pasquinelli et al. (2005). *Curr Opin Genet Develop* 15:200-205.
- Perez-Vich et al. (1998) *JAOCS* 75:547-555
- Perriman et al. (1992) *Gene* 113:157-163.
- Perrin et al. (2000) *Mol Breed* 6:345-352.
- Perry and Harwood (1993) *Phytochemistry* 33:329-333.
- 35 Phillips et al. (2002) *Journal of Food Composition and Analysis* 12:123-142.
- Picard (1994) *Cur. Top. Biotech.* 5:511-515.
- Pigeaire et al. (1997) *Mol. Breed.* 3:341-349.
- Potenza et al. (2004) *In Vitro Cell Dev. Biol. Plant* 40:1-22.

- Powell et al. (1996) *Vaccines* 183, Abstract.
- Prasher et al. (1985) *Biochem. Biophys. Res. Commun.* 127:31-36.
- Qiu et al. (2001) *J. Biol. Chem.* 276:31561-3156
- Ramamoorthy and Kumar (2012) *Plant Cell Rep.* 31:1923-1931.
- 5 Riddiford et al. (2000) *Vitam. Horm.* 60:1-73.
- Rossell and Pritchard (1991) *Analysis of Oilseeds, Fats and Fatty Foods*. Elsevier
- Ruuska et al. (2002) *Plant Cell* 14:1191-1206.
- Ryan et al. (1984) *J. Am. Oil Chem. Soc.* 61:1610-1619.
- Saha et al. (2006) *Plant Physiol.* 141:1533-1543.
- 10 Sanjaya et al. (2011) *Plant Biotech. J.* 9:874-883.
- Schaffner (1980) *Proc. Natl. Acad. Sci. U.S.A.* 77:2163-2167.
- Science Publishers Ltd: London (Chapter 2, pp.48-53).
- Scott et al. (2010) *Plant Biotechnol. J.* 8(8):912-27
- Semwal et al. (2011) *Bioresource Technology* 102:2151-2161.
- 15 Senior (1998) *Biotech. Genet. Engin. Revs.* 15:79-119.
- Shiau et al. (2001) *Vaccine* 19:3947-3956.
- Shiina et al. (1997) *Plant Physiol.* 115:477-483.
- Shimada and Hara-Nishimura (2010) *Biol. Pharm. Bull.* 33:360-363.
- Shippy et al., (1999) *Mol. Biotech.* 12:117-129.
- 20 Sieber et al. (2001) *Nature Biotechnology* 19:456-460.
- Siloto et al. (2009) *Lipids* 44:963-973.
- Sizemore et al. (1995) *Science* 270:299-302.
- Slade and Knauf (2005) *Transgenic Res.* 14: 109-115.
- Smith et al. (2000) *Nature* 407:319-320.
- 25 Srinivasan et al. (2007) *Planta* 225:341-51.
- Stalker et al. 1988 *Science* 242: 419-423.
- Stemmer (1994a) *Proc. Natl. Acad. Sci. USA* 91:10747-10751.
- Stemmer (1994b) *Nature* 370(6488):389-391.
- Stobart et al. (1997) *Planta* 203:58-66.
- 30 Tan et al. (2011) *Plant Physiol.* 156:1577-1588.
- Taylor (1997) *The Plant Cell* 9:1245-1249.
- Thillet et al. (1988) *J. Biol. Chem* 263:12500-12508.
- Tingay et al. (1997) *Plant J.* 11:1369-1376.
- Toriyama et al. (1986) *Theor. Appl. Genet.* 73:16-19.
- 35 Tumaney et al. (2001) *J. Biol. Chem.* 276:10847-10852.
- Tzfira and Citovsky (2006) *Curr. Opin. Biotech.* 17:147-154.
- Ulmasov et al. (1995) *Plant Physiol.* 108:919-927.
- Unger et al. (2002) *Transgenic Res.* 11:455-465.

- van de Loo et al. (1995) *Proc Natl Acad Sci U S A.* 92:6743-6747.
- Voinnet et al. (2003) *Plant J.* 33:949-956.
- Volkov et al. (1999) *Nucleic Acids Research* 27(18):e18.
- Wang et al. (2011) *J. Agric. Food Chem.* 60:144-52.
- 5 Waterhouse et al. (1998). *Proc. Natl. Acad. Sci. USA* 95:13959-13964.
- Weinmann et al. (1994) *Plant J.* 5:559-569.
- Weiss (2003) *Int. J. Med. Microbiol.* 293:95-106.
- Weselake et al. (2009) *Biotechnology Advances* 27:866-878.
- Winans et al. (1988) *Journal of Bacteriology* 170:4047-4054.
- 10 Wood et al. (2009). *Plant Biotech. J.* 7: 914-924.
- Wood et al., manuscript in preparation
- Yang et al. (2003) *Planta* 216:597-603.
- Yang et al. (2010) *PNAS* 107:12040-12045.
- Yen and Farese (2003) *J. Biol. Chem.* 278:18532-18537.
- 15 Yen et al. (2002) *Proc. Natl. Acad. Sci. U.S.A.* 99:8512-8517.
- Yen et al. (2005) *J. Lipid Res.* 46: 1502-1511.
- Zhang et al. (1999a) *Plant Cell Reports* 18:959-966.
- Zhang et al. (1999b) *Plant Cell, Tissue and Organ Culture* 56:37-46.
- Zhao et al. (1998) *Nature Biotechnology* 16:258-261.
- 20 Zheng et al. (2003) *The Plant Cell* 15:1872-1887.
- Zolotukhin et al. (1996) *J. Virol.* 70:4646-4654.

CLAIMS

1. A non-human organism or a part thereof comprising one or more exogenous polynucleotide(s) and an increased level of one or more non-polar lipid(s) relative to a
5 corresponding non-human organism or a part thereof lacking the one or more exogenous polynucleotide(s), wherein each of the one or more exogenous polynucleotides is operably linked to a promoter which is capable of directing expression of the polynucleotide in a non-human organism or part thereof, and wherein:
- (i) the one or more exogenous polynucleotide(s) comprise a first exogenous
10 polynucleotide which encodes a Wrinkled 1 (WRI1) transcription factor polypeptide, and a second exogenous polynucleotide which encodes a fatty acid acyltransferase, and
- (ii) the first and second exogenous polynucleotides have an effect on non-polar lipid accumulation in the non-human organism that is larger than an additive effect of the individual effects of each polynucleotide alone.
- 15
2. The non-human organism or a part thereof of claim 1, wherein one or more or all of the features (i) to (ix) apply:
- (i) the total fatty acid content in the non-polar lipid of the non-human organism or part thereof comprises at least 2% more oleic acid and/or at least 2% less palmitic
20 acid than the non-polar lipid in the corresponding non-human organism or part thereof lacking the one or more exogenous polynucleotides,
- (ii) the non-polar lipid(s) comprise a fatty acid which comprises a hydroxyl group, an epoxy group, a cyclopropane group, a double carbon-carbon bond, a triple carbon-carbon bond, conjugated double bonds, a branched chain such as a methylated
25 or hydroxylated branched chain, or a combination of two or more thereof, or any of two, three, four, five or six of the aforementioned groups, bonds or branched chains,
- (iii) the non-human organism or part thereof comprises oleic acid in an esterified or non-esterified form in its lipid, wherein at least 20% (mol%), at least 22% (mol%), at least 30% (mol%), at least 40% (mol%), at least 50% (mol%), or at least 60%
30 (mol%), at least 65% (mol%) or at least 66% (mol%) of the total fatty acids in the lipid of the non-human organism or part thereof is oleic acid,
- (iv) the non-human organism or part thereof comprises oleic acid in an esterified form in its non-polar lipid, wherein at least 20% (mol%), at least 22% (mol%), at least 30% (mol%), at least 40% (mol%), at least 50% (mol%), or at least
35 60% (mol%), at least 65% (mol%) or at least 66% (mol%) of the total fatty acids in the non-polar lipid of the plant or vegetative plant part is oleic acid,

(v) the total fatty acid content in the lipid of the non-human organism or part thereof comprises at least 2% more oleic acid and/or at least 2% less palmitic acid than the lipid in the corresponding non-human organism or part thereof lacking the one or more exogenous polynucleotides,

5 (vi) the non-human organism is an alga selected from the group consisting of diatoms (bacillariophytes), green algae (chlorophytes), blue-green algae (cyanophytes), golden-brown algae (chrysophytes), haptophytes, brown algae and heterokont algae,

(vii) the non-polar lipid(s) comprise a modified level of total sterols, free sterols, steroyl esters and/or steroyl glycosides,

10 (viii) the non-polar lipid(s) comprise waxes and/or wax esters,

(ix) the non-human organism or part thereof is one member of a population or collection of at least about 1000 such non-human organisms or parts thereof, and

(x) if the non-human organism is a plant, a vegetative part of the plant has a total non-polar lipid content of at least 10%.

15

3. The non-human organism or part thereof according to claim 1 or claim 2, wherein:

(i) the non-human organism is a plant, an alga or an organism suitable for fermentation such as a fungus, or

20 (ii) the non-human organism is a fungus such as an oleaginous fungus such as an oleaginous yeast, or

(iii) the part of the non-human organism is a seed, fruit, or a vegetative part of a plant such as an aerial plant part or a green part such as a leaf or stem.

25 4. The non-human organism or part thereof according to any one of claims 1 to 3, wherein the total sterol content and/or composition in the non-polar lipid is significantly different to the total sterol content and/or composition in the non-polar lipid in the corresponding non-human organism or part thereof lacking the one or more exogenous polynucleotides.

30

5. The non-human organism or part thereof according to any one of claims 1 to 4, wherein the increased level of one or more non-polar lipid(s) is such that one or more or all of the following features apply:

35 (i) the level is at least 0.5% greater on a weight basis than the level in a corresponding non-human organism or part thereof,

(ii) the level is at least 1% greater on a relative basis than in a corresponding non-human organism or part thereof,

(iii) the total non-polar lipid content of the non-human organism or part thereof is at least 0.5% greater on a weight basis than the content of a corresponding non-human organism or part thereof, respectively,

(iv) the total non-polar lipid content of the non-human organism or part thereof is at least 1% greater on a relative basis than the content of a corresponding non-human organism or part thereof, respectively, and

(v) the level of one or more non-polar lipid(s) and/or the total non-polar lipid content of the non-human organism or part thereof is at least 0.5% greater on a weight basis and/or at least 1% greater on a relative basis than in a corresponding non-human organism or part thereof, respectively, which is lacking the one or more exogenous polynucleotides and which comprises an exogenous polynucleotide encoding an *Arabidopsis thaliana* DGAT1.

15

6. The non-human organism or part thereof according to any one of claims 1 to 5, which comprises:

(i) a TAG, DAG, TAG and DAG, or MAG content which is at least 10% greater on a relative basis than the TAG, DAG, TAG and DAG, or MAG content of a corresponding non-human organism or part thereof, and/or

(ii) a total polyunsaturated fatty acid (PUFA) content which is increased or decreased relative to the total PUFA content of a corresponding non-human organism or part thereof.

7. The non-human organism or part thereof according to any one of claims 1 to 6, wherein the second exogenous polynucleotide encodes a polypeptide having diacylglycerol acyltransferase (DGAT) activity, monoacylglycerol acyltransferase (MGAT) activity, or glycerol-3-phosphate acyltransferase (GPAT) activity.

8. The non-human organism or part thereof according to any one of claims 1 to 7, wherein the non-human organism or part thereof further comprises a third, or more, exogenous polynucleotide(s) which encode one or more or any combination of:

(i) a further transcription factor polypeptide that increases the expression of one or more glycolytic or fatty acid biosynthetic genes in a non-human organism or a part thereof,

35

(ii) a further polypeptide involved in biosynthesis of one or more non-polar lipids,

(iii) a polypeptide that stabilizes the one or more non-polar lipids, an oleosin, such as a polyoleosin or a caleosin,

5 (iv) an RNA molecule which inhibits expression of a gene encoding a polypeptide involved in starch biosynthesis such as a AGPase polypeptide,

(v) an RNA molecule which inhibits expression of a gene encoding a polypeptide involved in the degradation of lipid and/or which reduces lipid content such as a lipase such as a CGi58 polypeptide, and

10 (vi) a silencing suppressor polypeptide,

wherein the third, or more, exogenous polynucleotide(s) is operably linked to a promoter which is capable of directing expression of the polynucleotide(s) in a non-human organism or a part thereof.

15 9. The non-human organism or part thereof according to any one of claims 1 to 8, wherein the one or more exogenous polynucleotides encode:

(i) a WRI1 transcription factor and a DGAT and an Oleosin,

(ii) a WRI1 transcription factor, a DGAT, a MGAT and an Oleosin,

20 (iii) a WRI1 transcription factor and a diacylglycerol acyltransferase 2 (DGAT2),

(iv) a WRI1 transcription factor, a MGAT and a glycerol-3-phosphate acyltransferase (GPAT),

(v) a WRI1 transcription factor, a MGAT, a GPAT and a DGAT,

(vi) a WRI1 transcription factor and a MGAT,

25 (vii) a WRI1 transcription factor, a DGAT and a MGAT,

(viii) a WRI1 transcription factor, a DGAT, a MGAT, an Oleosin and a GPAT,

or

(xiii) a WRI1 transcription factor, a MGAT and an Oleosin, and

(ix) optionally, a silencing suppressor polypeptide.

30

10. The non-human organism or part thereof according to any one of claims 1 to 9, wherein the part thereof is a vegetative plant part that comprises:

35 (i) a total non-polar lipid content of at least about 11%, at least about 12%, at least about 13%, at least about 14%, or at least about 15% (w/w dry weight or seed weight), and/or

(ii) a total TAG content of at least about 7%, at least about 10%, at least about 11%, at least about 12%, at least about 13%, at least about 14%, at least about 15%, or at least about 17% (w/w dry weight or seed weight).

5 11. The non-human organism or part thereof according to any one of claims 1 to 10, wherein the second exogenous polynucleotide encodes a DGAT, such as a DGAT1, and the non-human organism or part thereof further comprises a third exogenous polynucleotide encoding an oleosin, wherein the non-human organism or part thereof has one or more or all of the following features:

10 (i) a total TAG content of at least 10%, at least 12.5%, at least 15% or at least 17% (% weight of dry weight),

(ii) at least a 40 fold, at least a 50 fold, at least a 60 fold, or at least a 70 fold, or at least a 100 fold, higher total TAG content in the non-human organism or part thereof relative to a corresponding non-human organism or part thereof lacking the exogenous polynucleotides,

15 (iii) oleic acid comprises at least 19%, at least 22%, or at least 25% (% weight) of the fatty acids in TAG,

(iv) at least a 10 fold, at least a 15 fold, at least a 17 fold, or at least a 19 fold, higher level of oleic acid in TAG in the non-human organism or part thereof relative to a corresponding non-human organism or part thereof lacking the exogenous polynucleotides,

20 (v) palmitic acid comprises at least 20%, at least 25%, or at least 28% (% weight) of the fatty acids in TAG,

(vi) at least a 1.25 fold higher level of palmitic acid in TAG in the non-human organism or part thereof relative to a corresponding non-human organism or part thereof lacking the exogenous polynucleotides,

25 (vii) linoleic acid comprises at least 15%, or at least 20%, (% weight) of the fatty acids in TAG,

(viii) α -linolenic acid comprises less than 15%, less than 11% or less than 8% (% weight) of the fatty acids in TAG, and

30 (ix) at least a 5 fold, or at least an 8 fold, lower level of α -linolenic acid in TAG in the non-human organism or part thereof relative to a corresponding non-human organism or part thereof lacking the exogenous polynucleotides.

35 12. Use of a first polynucleotide encoding a Wrinkled 1 (WRI1) transcription factor polypeptide, together with a second polynucleotide that encodes a fatty acid

acyltransferase, for producing a transgenic non-human organism or part thereof having an enhanced ability to produce one or more non-polar lipids relative to a corresponding non-human organism or part thereof lacking the first and second polynucleotides, wherein the first and second polynucleotides are each exogenous to the non-human organism or part thereof and are each operably linked to a promoter which is capable of directing expression of the polynucleotide in the transgenic non-human organism or part thereof, wherein the first and second exogenous polynucleotides have an effect on non-polar lipid accumulation in the non-human organism that is larger than an additive effect of the individual effects of each polynucleotide alone.

13. Use of the non-human organism or a part thereof according to any one of claims 1 to 11 for the manufacture of an industrial product, such as a hydrocarbon product such as fatty acid esters, fatty acid methyl esters and/or a fatty acid ethyl esters, an alkane such as methane, ethane or a longer-chain alkane, a mixture of longer chain alkanes, an alkene, a biofuel, or carbon monoxide and/or hydrogen gas, a bioalcohol such as ethanol, propanol, or butanol, biochar, or a combination of carbon monoxide, hydrogen and biochar.

14. A process for producing an industrial product, the process comprising the steps of:

(i) obtaining a non-human organism or a part thereof according to any one of claims 1 to 11, and

(ii) optionally, physically processing the non-human organism or part thereof of step (i), and

(iii) converting at least some of the lipid in the non-human organism or part thereof of step (i), or in the processed non-human organism or part thereof obtained by step (ii), to the industrial product by applying heat, chemical, or enzymatic means, or any combination thereof, to the lipid, *in situ* in the non-human organism or part thereof of step (i), or in the processed non-human organism or part thereof obtained by step (ii), and

(iv) recovering the industrial product, thereby producing the industrial product.

15. A process for producing extracted lipid, the process comprising the steps of:

(i) obtaining a non-human organism or a part thereof according to any one of claims 1 to 11,

- (ii) extracting lipid from the non-human organism or part thereof, and
- (iii) recovering the extracted lipid.

16. A process for producing fuel, the process comprising:

- (i) reacting lipid recovered or extracted from the non-human organism or part thereof according to any one of claims 1 to 11 with an alcohol, optionally, in the presence of a catalyst, to produce alkyl esters, and
- (ii) optionally, blending the alkyl esters with petroleum based fuel.

17. A process for producing a synthetic diesel fuel, the process comprising:

- (i) converting the lipid in the non-human organism or part thereof according to any one of claims 1 to 11 to a syngas by gasification, and
- (ii) converting the syngas to a biofuel using a metal catalyst or a microbial catalyst.

18. A process for producing a biofuel, the process comprising converting the lipid in the non-human organism or part thereof according to any one of claims 1 to 11 to bio-oil by pyrolysis, a bioalcohol by fermentation, or a biogas by gasification or anaerobic digestion.

19. Feedstuffs, cosmetics or chemicals comprising the non-human organism or a part thereof according to any one of claims 1 to 11.

20. A process for feeding an animal comprising providing an animal with the feedstuff of claim 19.

1/20

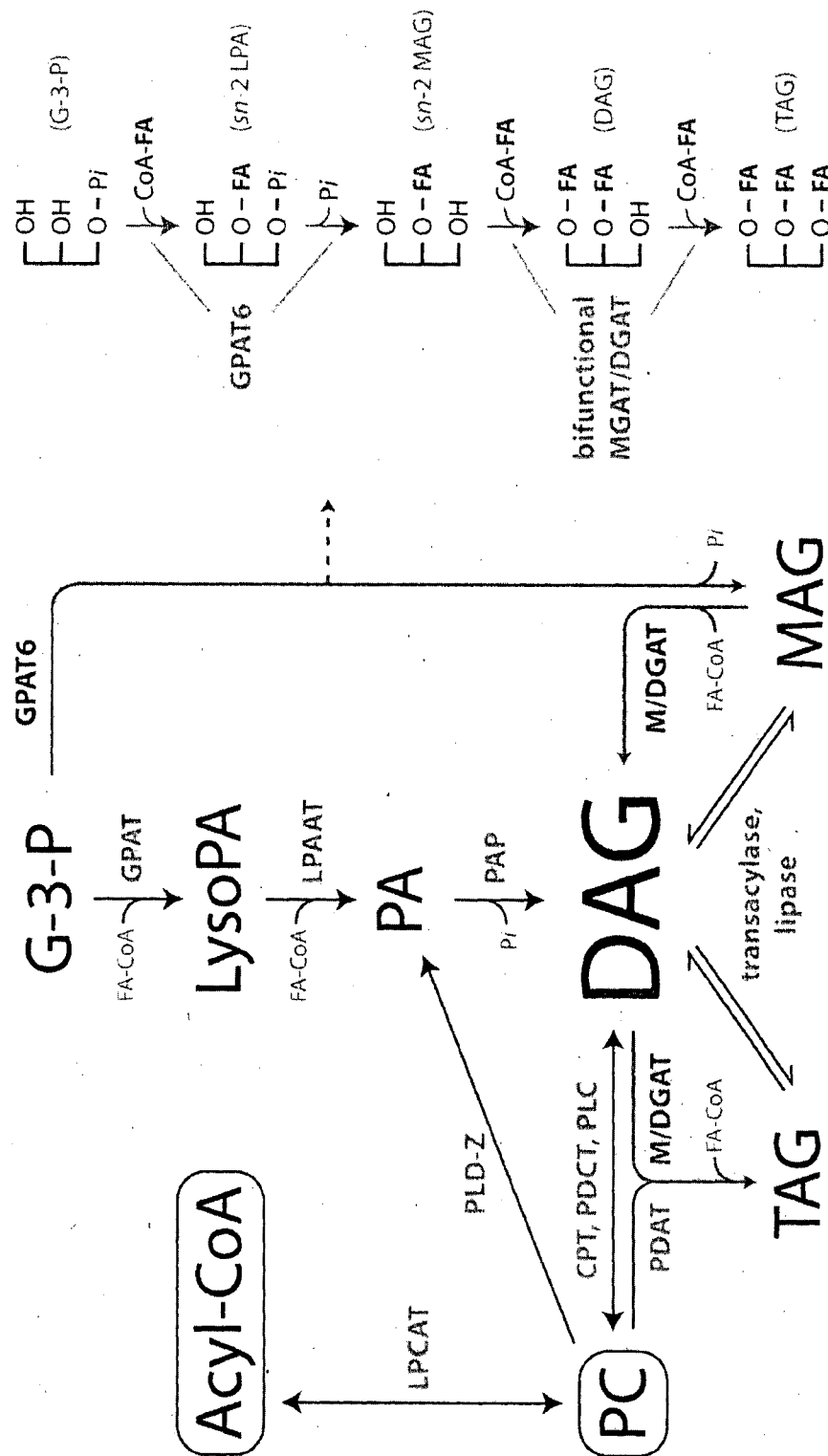


Figure 1

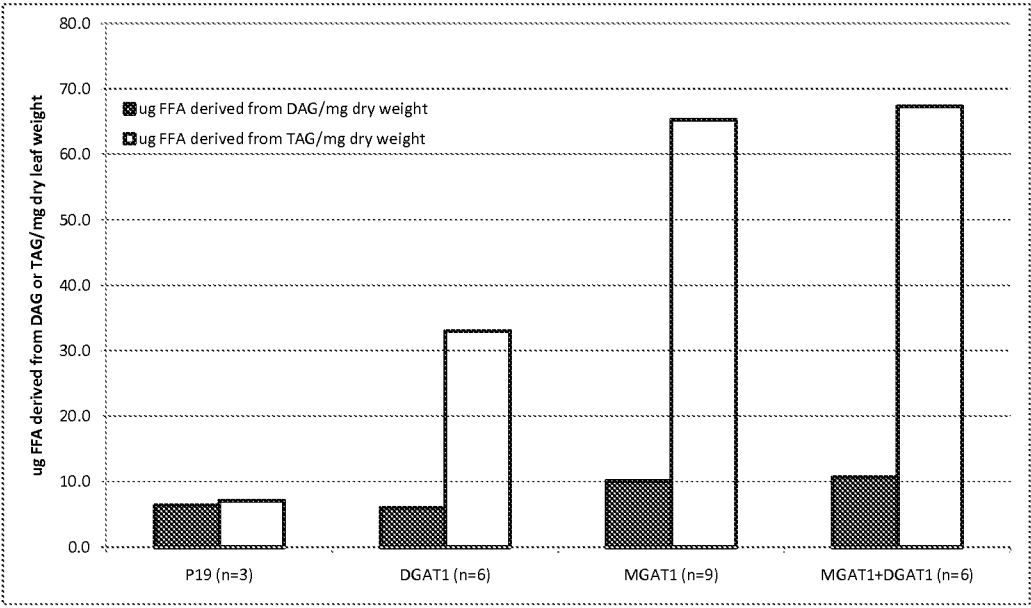
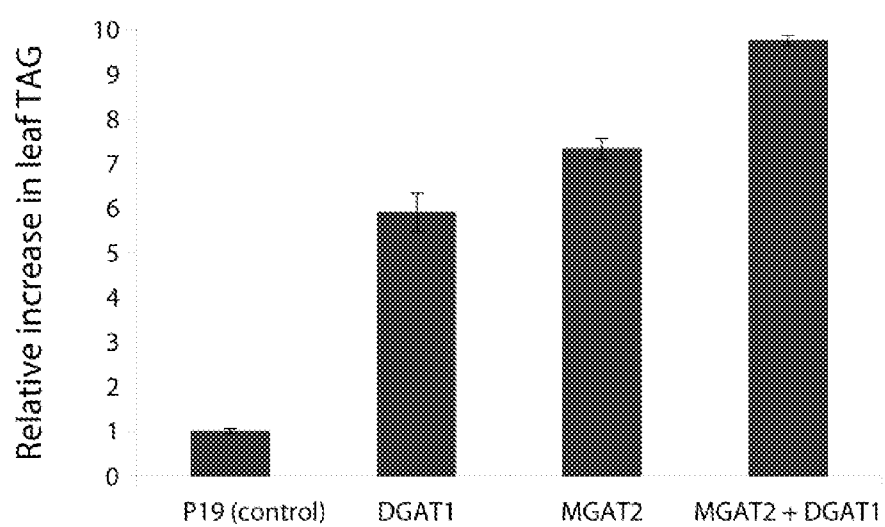


Figure 2

3/20

**Figure 3**

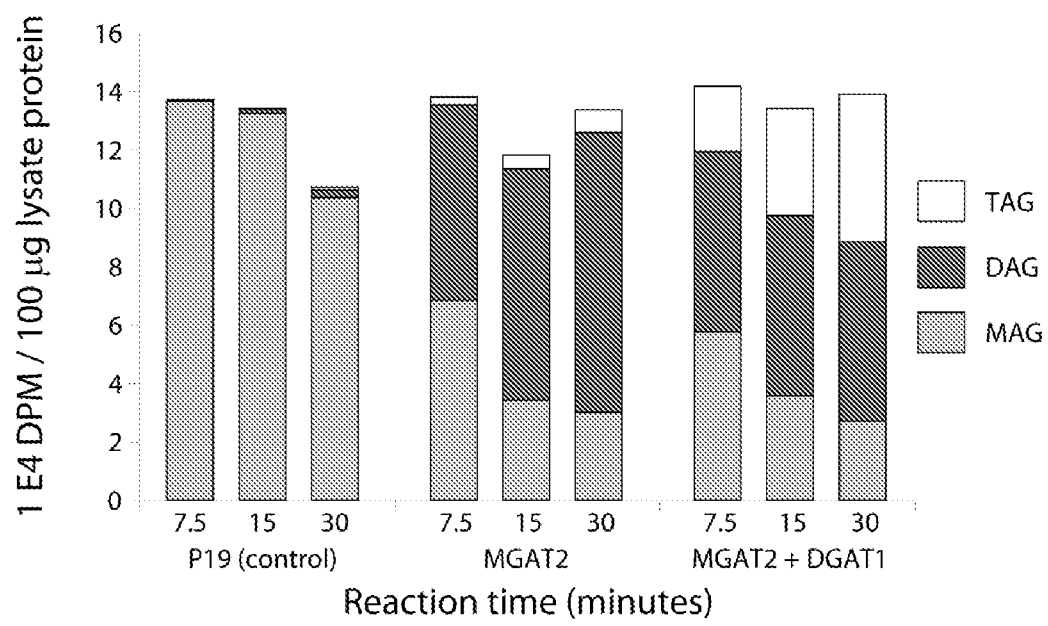


Figure 4

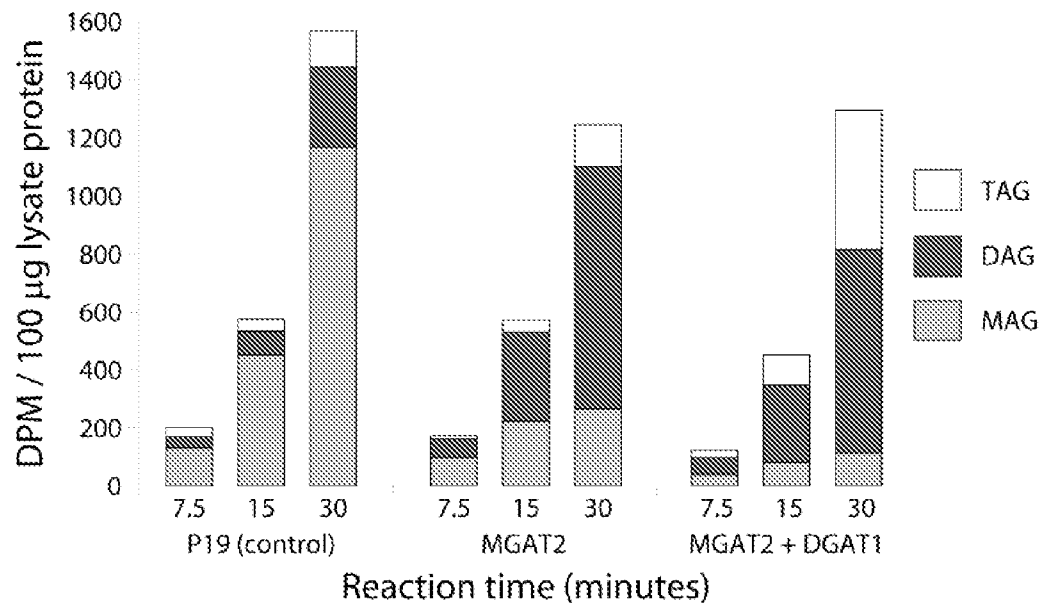


Figure 5

6/20

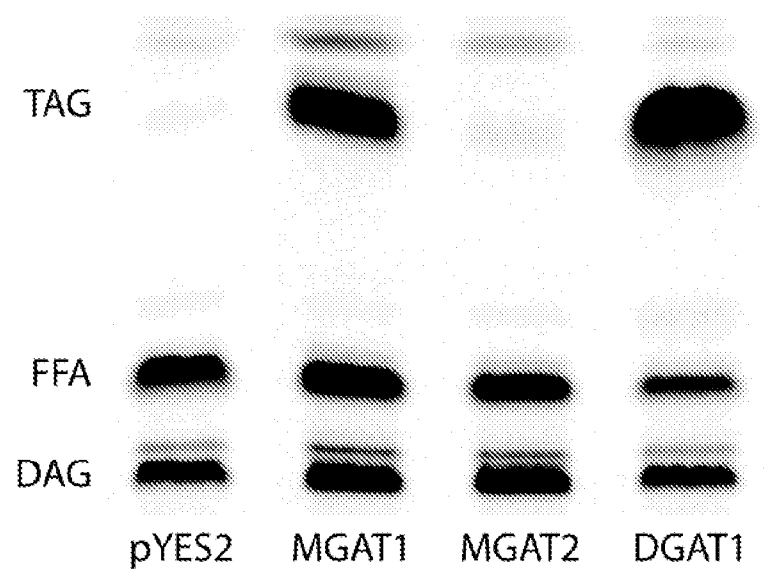


Figure 6

7/20

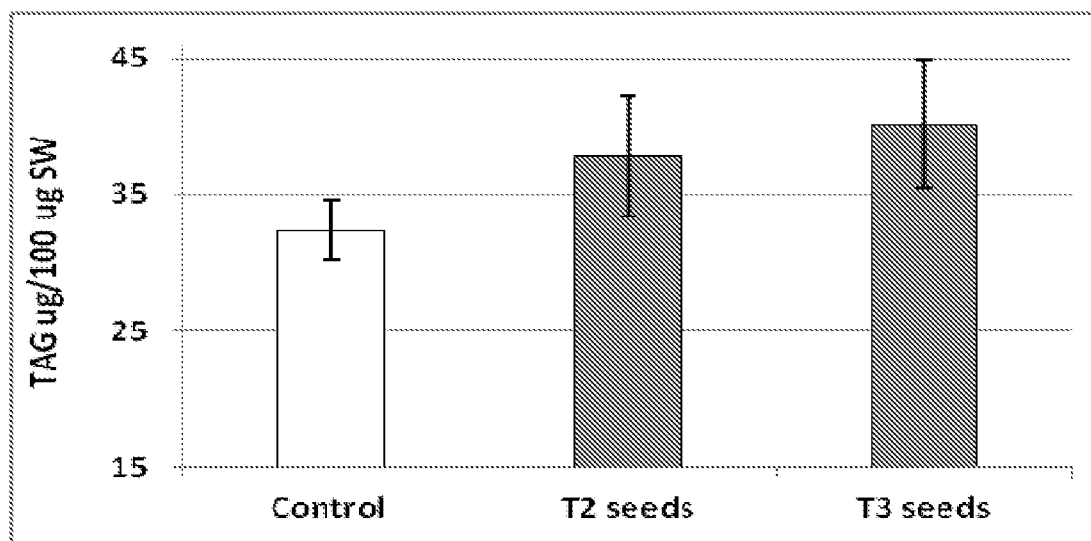


Figure 7

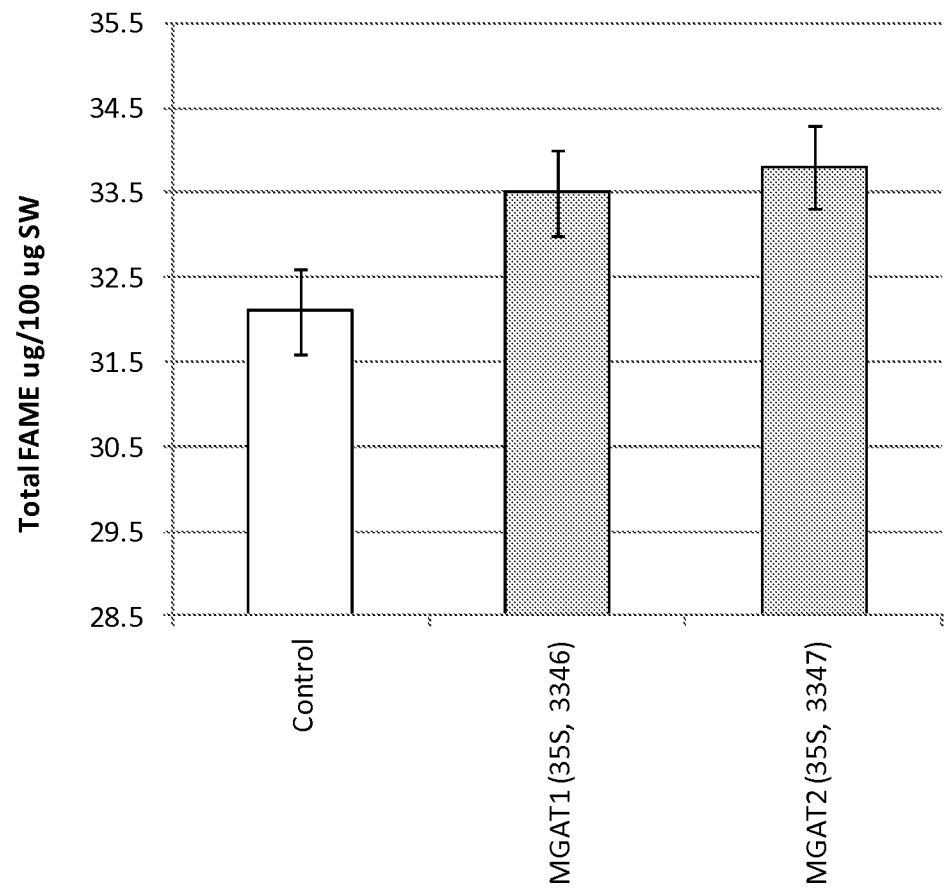


Figure 8

9/20

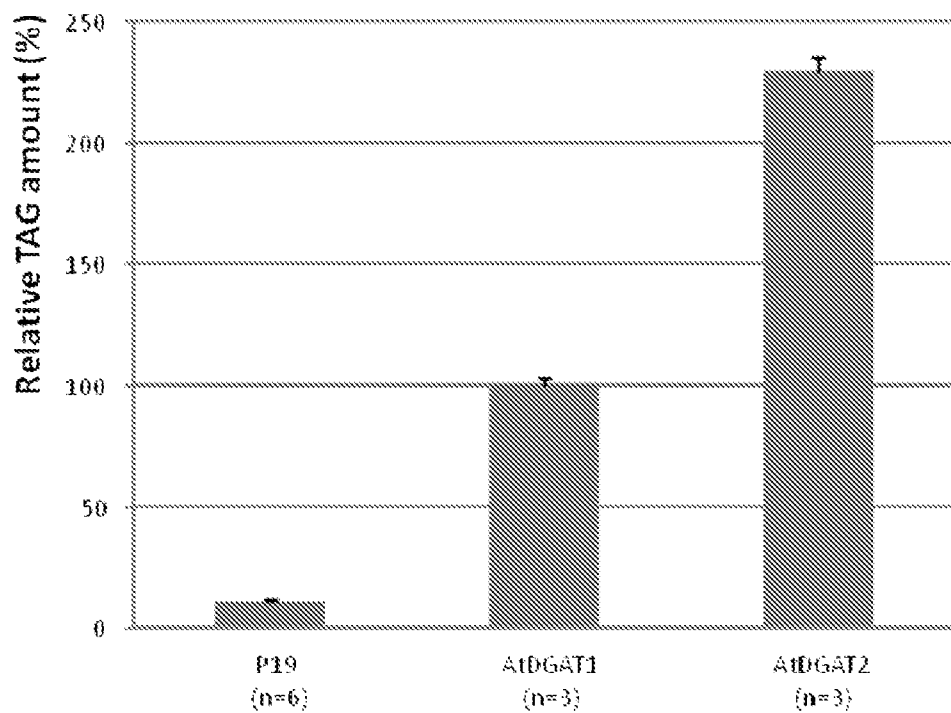


Figure 9

10/20

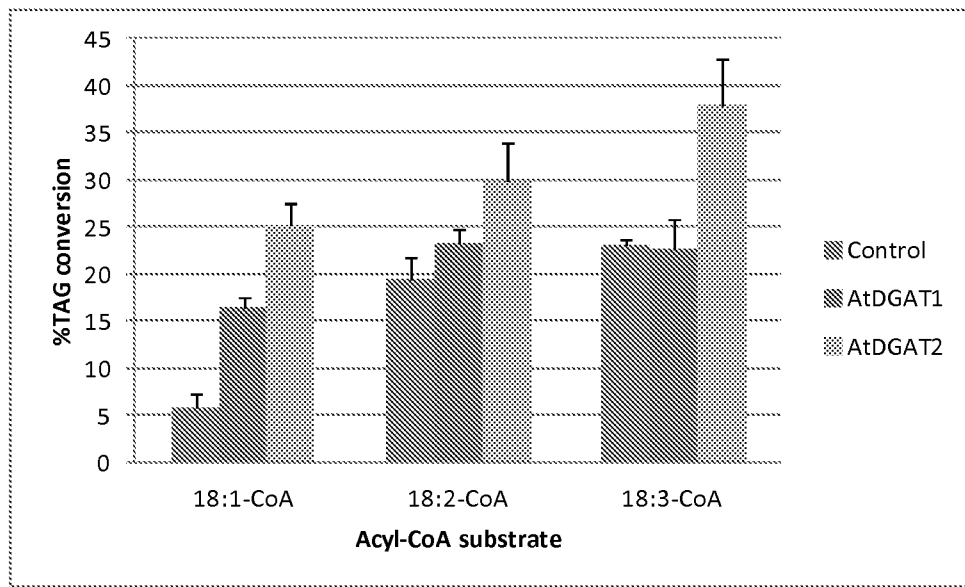


Figure 10

11/20

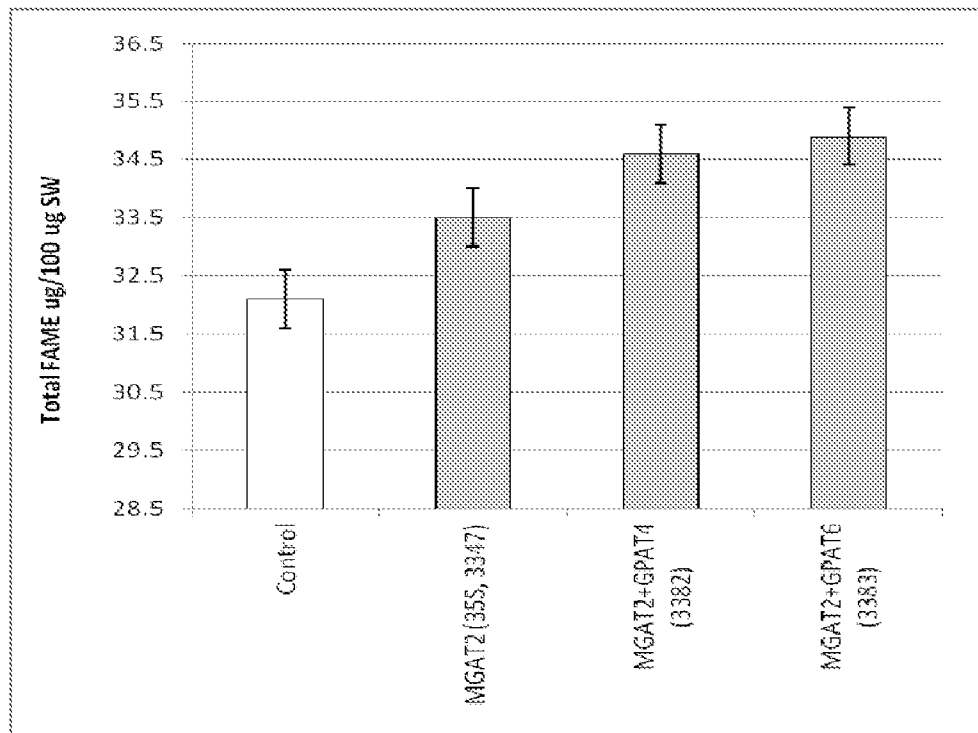


Figure 11

12/20

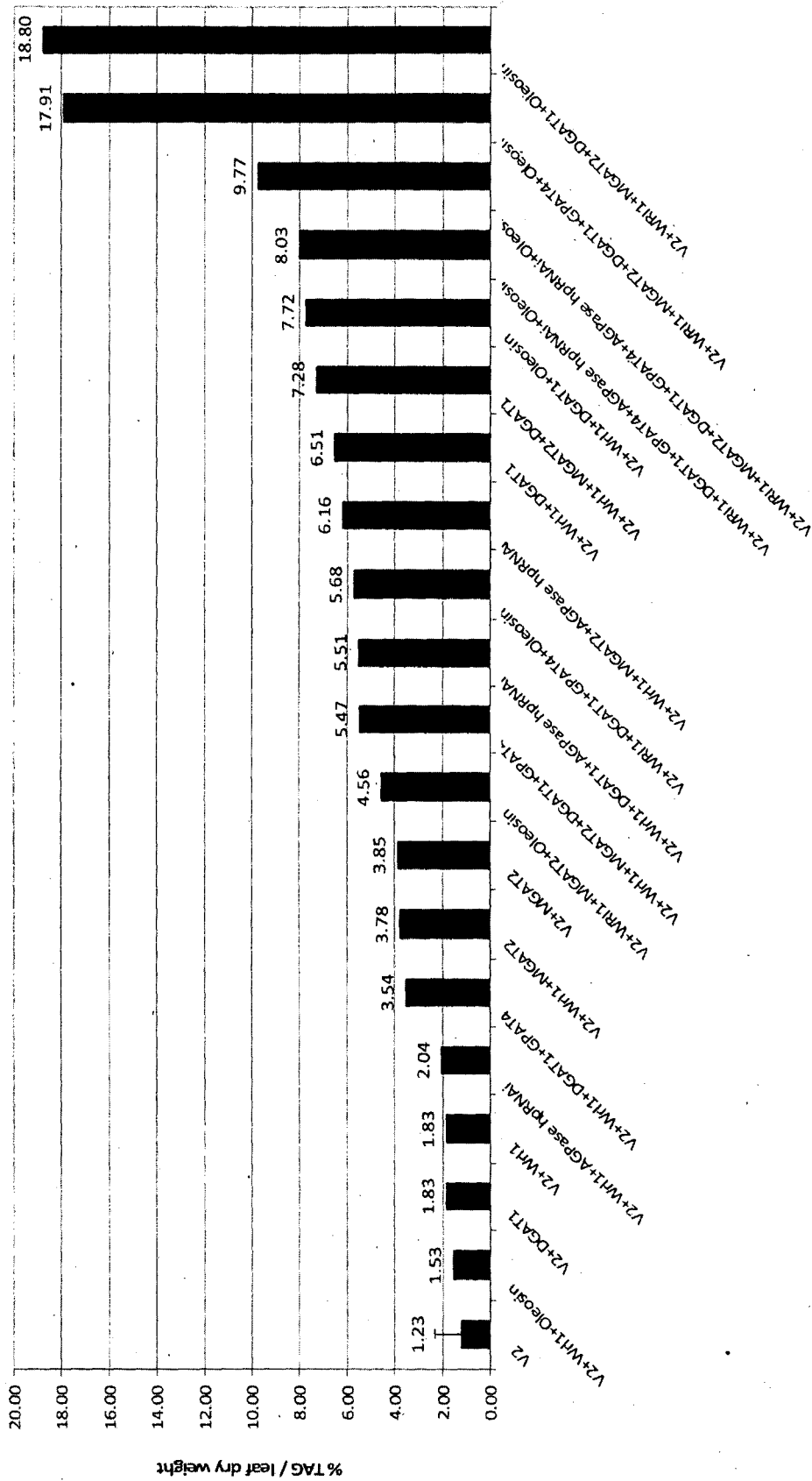


Figure 12

13/20

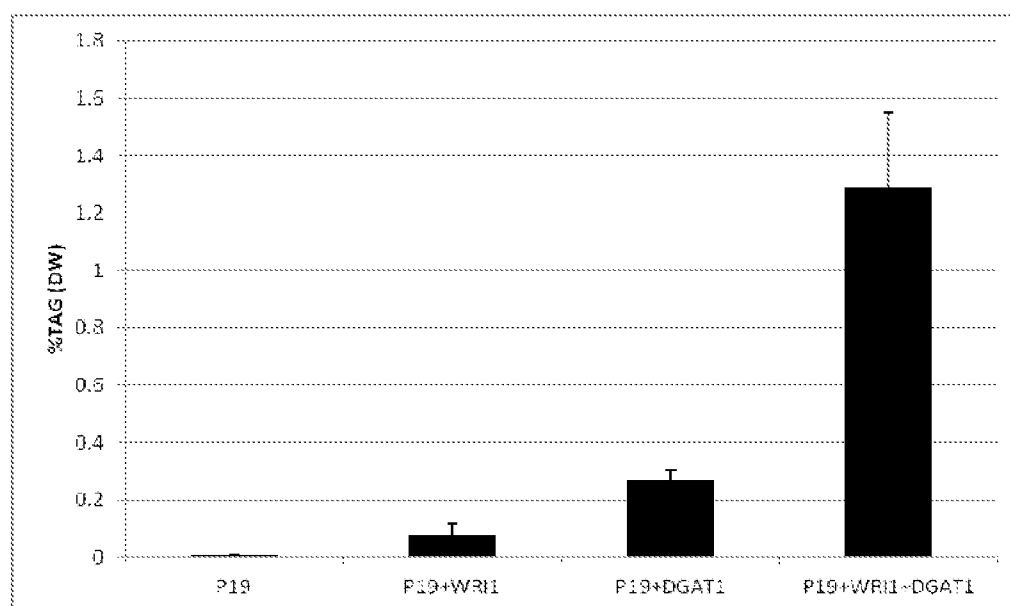


Figure 13

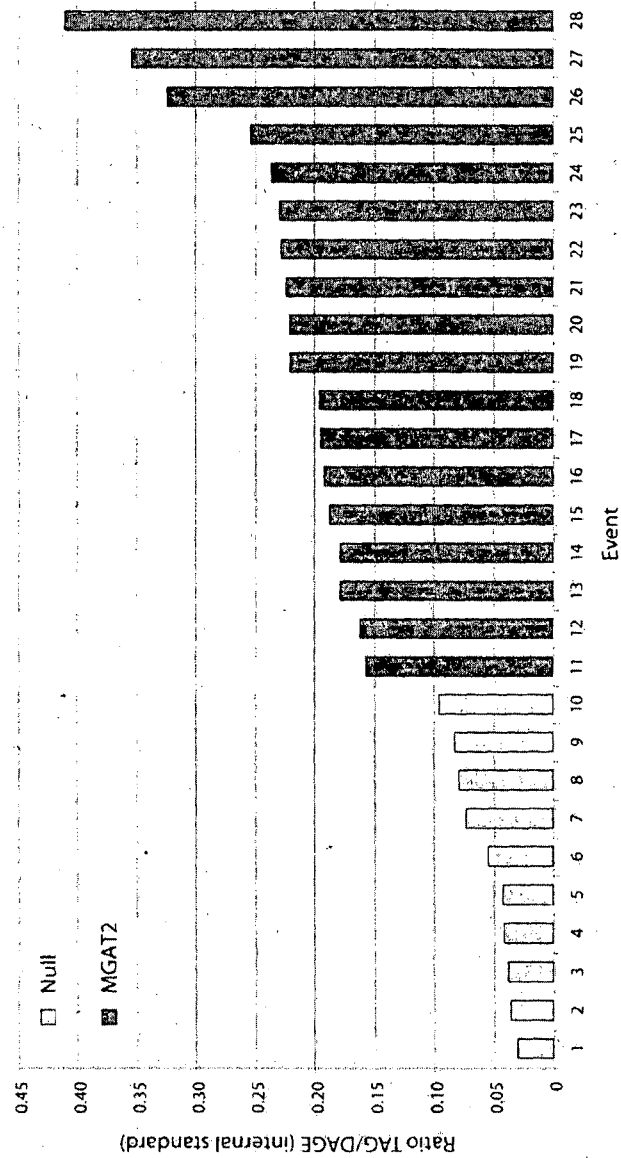


Figure 14

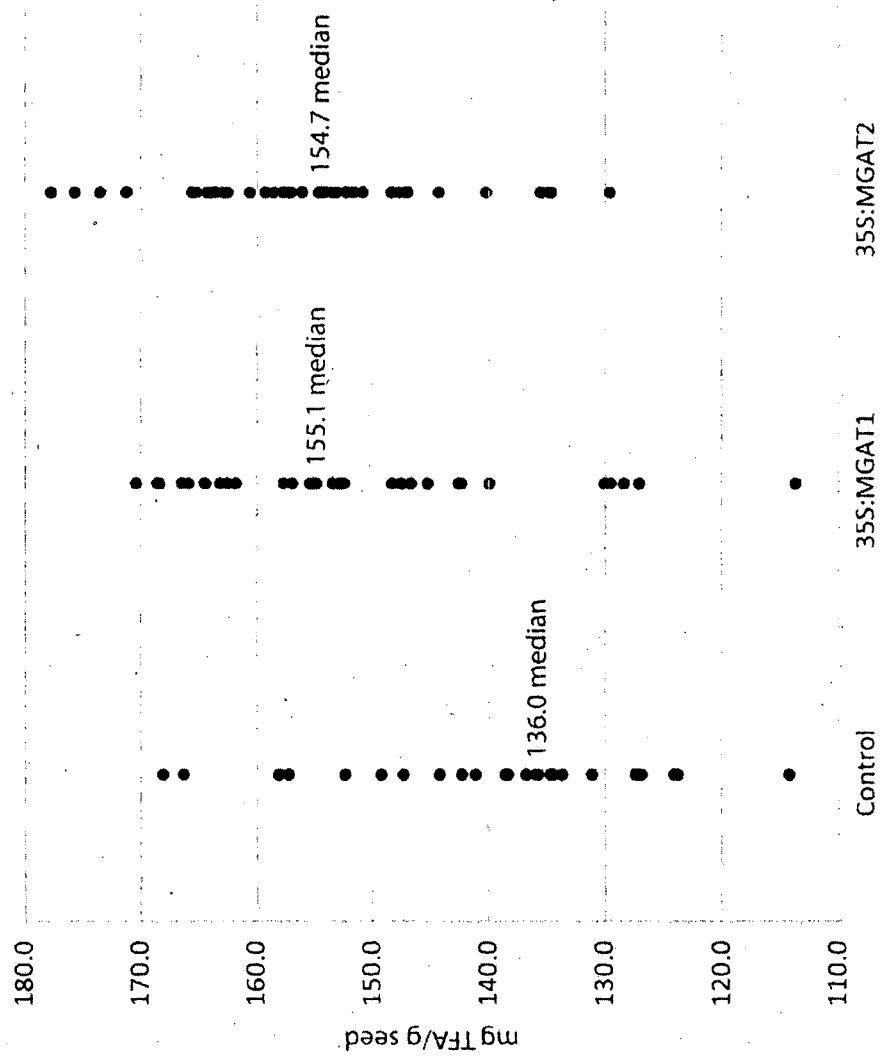


Figure 15

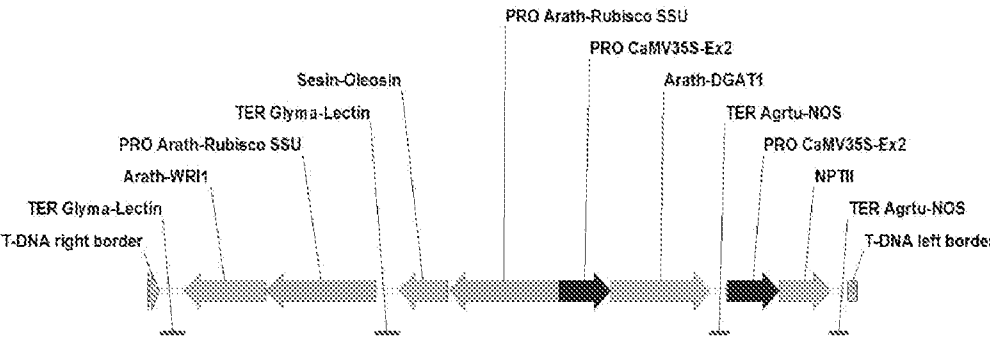


Figure 16

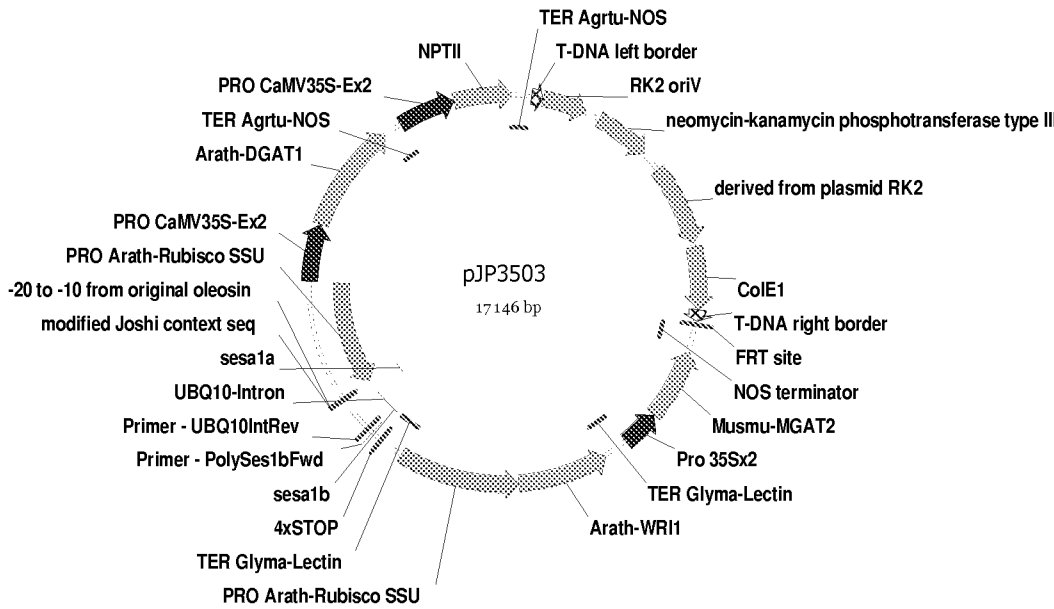


Figure 17

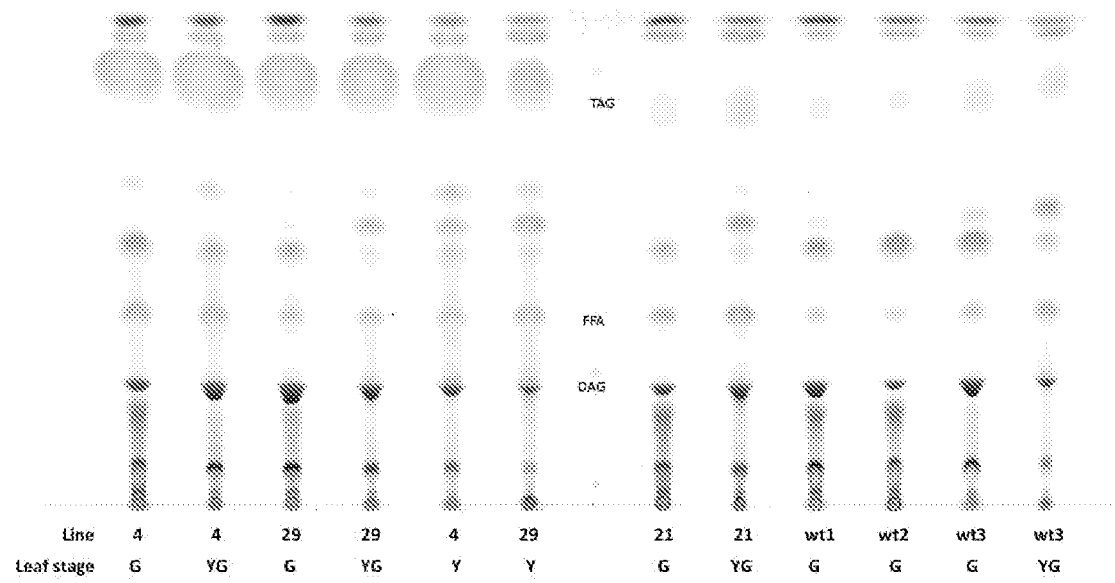


Figure 18

18/20

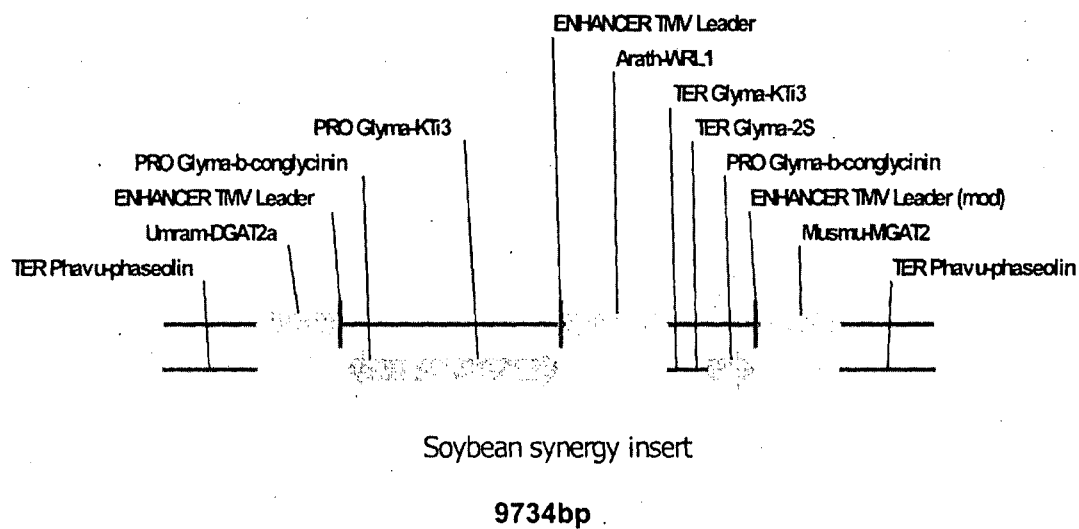


Figure 19A

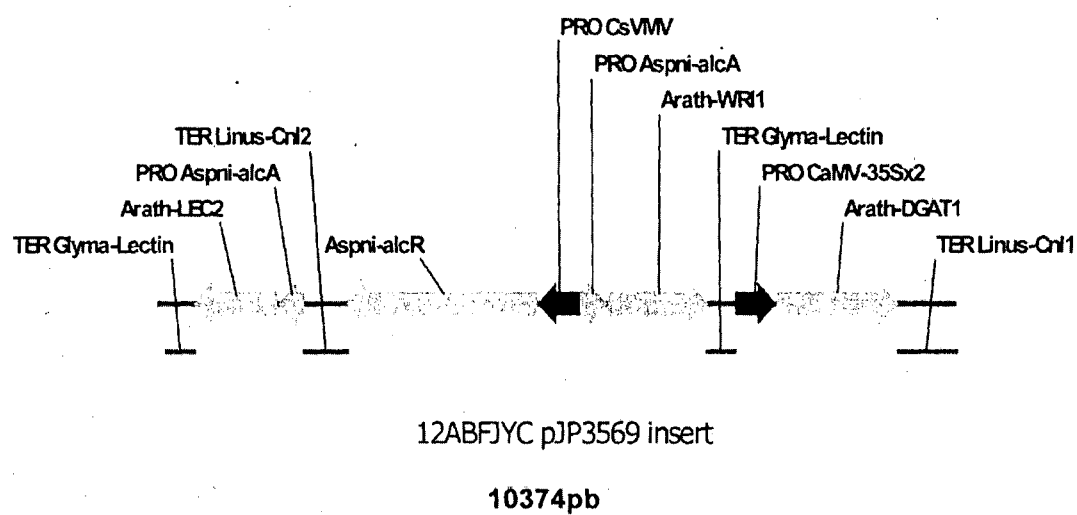


Figure 19B

19/20

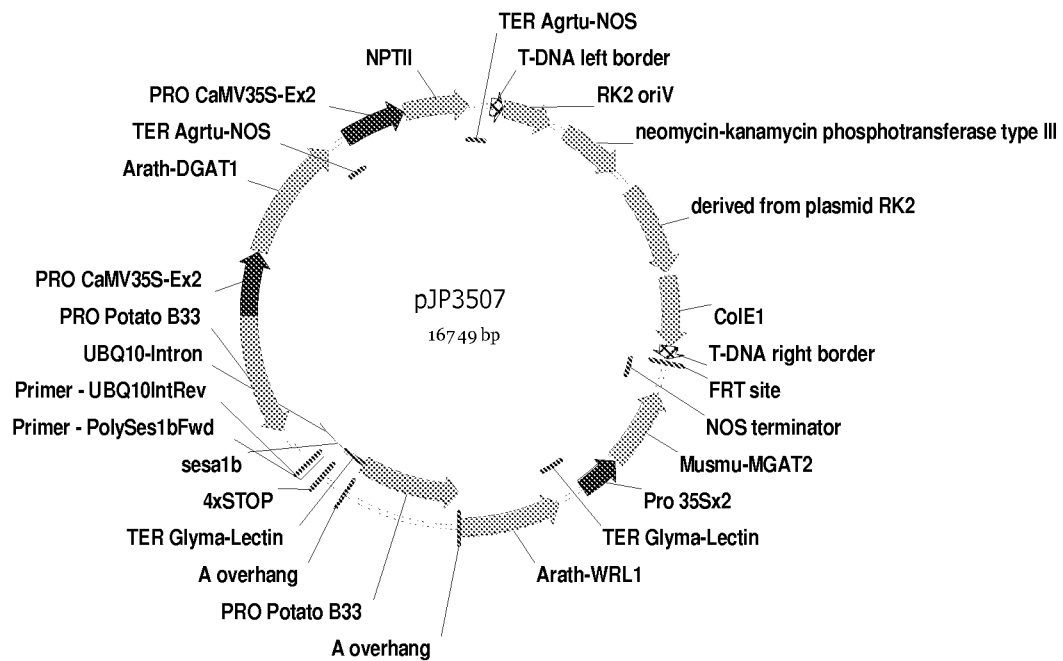


Figure 20

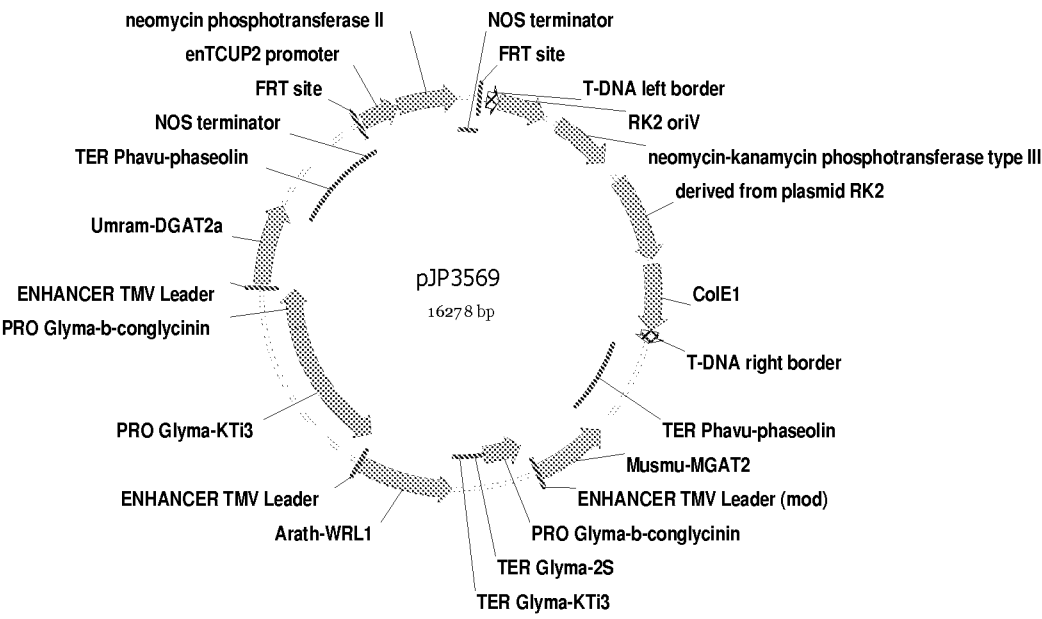


Figure 21

pctau2012001598-seql
SEQUENCE LISTING

<110> Commonwealth Scientific and Industrial Research
Organisation

<120> Processes of producing lipid

<130> 512813

<150> US61/580590

<151> 2011-12-27

<150> US61/718563

<151> 2012-10-25

<160> 448

<170> PatentIn version 3.5

<210> 1

<211> 1011

<212> DNA

<213> Mus musculus

<400> 1

atgatggttg agttcgtcc tcttaacact cctcttgcta gatgccttca aactgctgct	60
gttcttcaat gggttctctc tttcttgctc cttgttcaag tgtgcatcgg aatcatggtt	120
atgctcgttc tctacaacta ctggttcctc tacatccctt acctcgtttg gttctactac	180
gattggagaa ctctgagca aggtggaaga agatggaact gggttcaatc ttggcctggt	240
tggaagtact tcaaagagta cttccctatc tgcctcgtta agactcagga tcttgatcct	300
ggacacaact acatcttcgg attccatcct catggaatct tcgtgcctgg tgctttcgga	360
aacttctgta ctaagtacag cgacttcaag aagcttttcc ctggattcac ctcttacctc	420
catgttgcta agatctgggt ctgcttccca ctcttcagag aatacctcat gtctaacgga	480
cctgtgtctg tgtctaaaga atctctctct cacgttctct ccaaggatgg tggaggaaac	540
gtttcaatca tcgttcttgg aggtgctaaa gaggtccttg aggctcatcc tggaactttc	600
actttgtgta tcagacagcg taagggattc gttaagatgg ctcttactca cgggtgcttct	660
cttgttccag ttttctcatt cggagagaac gatctctaca agcagatcaa caaccctaag	720
ggatcttggc ttagaactat ccaggatgct atgtacgact ctatgggagt tgctttgcct	780
cttatctacg ctagaggaat cttccagcac tacttcggaa tcatgcctta cagaaagctc	840
atctacaccg ttgttggtag acctatccct gttcaacaga tccttaacct tacttctgag	900
cagatcgagg aacttcatca gacttacctt gaggaactca agaagttggt caacgagcac	960
aagggaaagt acggaatccc tgagcatgag actctcgttt tcaagtgatg a	1011

<210> 2

<211> 1008

<212> DNA

<213> Mus musculus

<400> 2

atggtcgagt tcgctccact tcttgttcct tgggagagaa gattgcagac tttcgctggt	60

pctau2012001598-seq1

cttcagtggg	ttttctcatt	ccttgctctt	gctcaactct	gcatcgttat	cttcgttgga	120
ctcttggtca	ccaggttctg	gcttttctct	gttctttacg	ctacctggtg	gtacttggat	180
tgggataagc	ctagacaagg	tggtagacct	atccaattct	tcagaagatt	ggccatctgg	240
aagtacatga	aggattactt	ccctgtgtct	ctcgttaaga	ctgctgagct	tgatcctagc	300
cgtaactaca	ttgctggatt	ccatcctcat	ggtgttcttg	ctgctggtgc	tttccttaac	360
ctctgtactg	agtctactgg	attcacctct	ttgttccctg	gaatcagatc	ctaccttatg	420
atgcttaccg	tttggttcag	ggctccattc	ttcagagatt	acatcatgtc	tggtggactc	480
gtgtctagcg	agaaagtttc	tgctgatcac	atccttagca	gaaaggggtg	tggaacacct	540
cttgctatca	tcgttggagg	tgctcaagag	gctcttgatg	ctagacctgg	tgcttacaga	600
cttctcttga	agaaccgtaa	gggattcatc	agacttgctc	ttatgcatgg	tgctgctctt	660
gttcctatct	tctcattcgg	agagaacaac	ctcttcaacc	aggttgagaa	cactcctgga	720
acttggctta	gatggatcca	gaacagactc	caaaagatta	tggaatcag	ccttccactt	780
ttccatggaa	ggggagtttt	ccaatactct	ttcggactca	tgcccttcag	acagcctatc	840
actactatcg	ttggaaagcc	tatcgagggt	caaagtactc	ctcagccttc	tagagaggaa	900
gttgatagac	ttcaccagcg	ttacatcaaa	gagttgtgca	agcttttcga	ggaacacaag	960
ctcaagttca	acgttccaga	ggatcagcat	cttgagttct	gctaataga		1008

<210> 3
 <211> 1008
 <212> DNA
 <213> *Ciona intestinalis*

<400> 3		
atgtgggagt	tcgcccctct	tagaatccct atgatcagaa gattgcagac ccttgctggt 60
tctcacttcg	tgttctgttt	cttggctctt gctgctactt gcactcttct tactatctat 120
atcaccctct	tcaccagcta	ctggcctatc actactgctt acctcgtttt cttgttcctc 180
gatcgtaca	ctcctgagtc	tggtggtaga agatctgagt gggttagaaa ctggtctttg 240
tggaagtgga	tgtctgatta	cttcccttgt acccttcaca agaccgttga tcttgatcct 300
aagaggaact	acatcttcgg	aatccatcct catggtgttc tctgtatcgg atctttcact 360
cacttctcta	ccaacggatc	tggattctct catgttttcc ctggattcac ctctacctt 420
actatgcttc	cattctgggt	caagatgcct ttcttcaggg attacgtgat gtctggtgga 480
cttactcctg	ctactaggaa	ggctatcaag cacactatca ctagacctgg tggtggacat 540
atctgttgta	ttatccctgg	tggtgctcct gagtctctta acgctagacc aggtgatggt 600
gttcttttgc	ttaagcagag	actcggattc cttaagcttg ctatcactaa cgggtgttcct 660
cttgttcctg	tgttctcttt	cgggtgatcat gctctttggg agcaaaagcc taaccctcct 720
ggatctctta	tcagaagatt	ccaggactct agccaaaagt ggatgcaagt tgctctccct 780
gttttccatg	ctagaggaat	cttccagtac aacttcggac tcatccctta cagaagatct 840
gttcacactg	ttgttggaga	gcctatcgaa gttcctcaaa actctaacc tacctctgag 900

pctau2012001598-seq1

gatctcatgt ctcttcaaga ggactacatc aacagactca gggctatctt cgatgagcac	960
aagtctaagt acctccctga ggattgcaag ctcatcatca actgatga	1008

<210> 4
 <211> 1020
 <212> DNA
 <213> Tribolium castaneum

<400> 4	
atgaagatcc tcggaatcaa gttcgcctcct cttcatatcc ctcttgagag aagattgcag	60
actcttgctg ctggatgttg gttcactact cttgctttcg gaactttcat cggaaccttc	120
atctggggtt acgtgttctt caccagggtt tggtacttgt ctgttcttta cgccaccatc	180
atctacctcg agaagtctaa gtgtgagaag ggtggtagac ctatcgagtg gatcagacat	240
tggggatggg ggtactacct caagaactac ttcccttgca agctcgattt cgttcctgga	300
cttactttcg atcctaagag gaactacctc ttcgcttggt accctcatgg aatcctccct	360
gctggacctt tcaacactat cggatctcca tactctgagt tcagcaagct tttcccaaag	420
ttcagagtga gactcgttat ccttcaccag catttcttca tccatttcct tagagagatc	480
gcttacggaa ctggtggaat ctctgcttct gctaagtctc ttaaccacgt tctcggatct	540
ccagagggag gttacattgc tgttcttatg cctggtggtg ctggtgaggc ttacaactct	600
agacctggac agtacagaat catcctcaag aacagaaagg gattcggtta gctcgctctc	660
agaaacggat ctctcttgtt tcctgttatc tctttcggag agcctgagct tttcgatcaa	720
gttgagggaa gaacctcag aaagatccaa gagtctatca gaaagtacct tggactcgct	780
cctgttatct tctctggaag aggattcttc cagtactctt tcggagttaa ccctcaaaga	840
aggcctatca ctactgttgt tggacaccct atcgagggtta caaagatcga gaagcctacc	900
aacgaggaag ttgacgagct tcacaaaaag tacatgcaag agcttgagaa cttgttcgaa	960
gagtacaagt tcaagtacct cgagaaccct aaggatatcc atcttgagtt cgagtgatga	1020

<210> 5
 <211> 867
 <212> DNA
 <213> Danio rerio

<400> 5	
atgctcggag ccctttacgc tggatggctt taccttgata gagatacacc ttcttgcgga	60
ggtagaagat ctcaatgggt tagatcttgg aggatctgga tgcacttcag agattacttc	120
cctatcacc cgtttaagac cgttgatctt gatcctaggc acaactacct tctcggattc	180
catcctcatg gtgttcttgt tgctggtgga ttcggaaact tctgtactga ggcttctgga	240
ttctctcaaa tgttccttg acttaccctt taccttctta tgcttccatt ctgggttcagg	300
gtgccattct tccgtgagta catcatgtgc ggaggacttg tttcttctga gaaggcttct	360
gcttcttacc ttcttgaca tcctggtggt ggacaagctg ctgttatcgc tgttgagggt	420
gctcctgagt ctcttgaggc tagacctggt gctcttactc ttcaacttct tcagaggaag	480

pctau2012001598-seql

ggattcatca agcttgctct taagcacggt gcttggcttg ttccagtttt ctcattcgga	540
gagaacgagc ttttcgacca aatggaaaac cctgctggat ctgctctcag aagaatgcaa	600
gaaagactcc agaggattat gggagttgct ctcccacttt tccatgctag gggagttttc	660
caatactctt tcggacttct cccatacaga aagcctatcc ataccgttgt tggtagacct	720
atccctgttt ctcaaactcc ttgcccttct aaagaggata tcgacgctct ccacactctc	780
tacatgcaag gacttaccca ggtgttcgag gaaaacaaga agcactacgg aatcgctgat	840
gataagcacc ttaagttcac ctgatga	867

<210> 6
 <211> 762
 <212> DNA
 <213> Danio rerio

<400> 6	
atgagggact acttcctat cagacttatc aagaccgctg atctcgacac cagaaagaac	60
tacgttatgg gattccatcc tcacggaatc cttgttgctg gtgcttttac taacttctgc	120
actgaggcta ccggattctc taagttgttc cctggaatca agagcaacct tcttatgctt	180
cctctttggt tccgtgctcc attcttcaga gactacatca tgtctgctgg actcgttcct	240
tctgacaaaag agtctgcttc ttacctcctt agaagaaagg gtggtggaaa cgctgttggt	300
atcgctgttg gaggtgctcc tgaggctctt gatgctcatc ctggtgatta cactgttcac	360
cttgctaaca agaagggatt catcaagctt gctatcgagc atggtgctga tcttgttcct	420
atctactctt tcggagagaa cgaggttttc gatcaagttc aaaaccctag aggaacctgg	480
cttcgttaca tccaagaaag actccagagg attatgggag tttctcttcc acttttccac	540
gctaggggag ttttccagta cactttcgga cttatgcctt accgtaagcc tatcaacacc	600
gttgttggta gacctatccc tgttgagaag aacgagaagc ctactgctga ggaacttgat	660
gtttaccatc agcgttatcat ggatgagctt gctagacttt tcgaggatca caagggaaac	720
tacggtgttc ctgaggatac tcatctcggt ttccagtgat ga	762

<210> 7
 <211> 1093
 <212> DNA
 <213> Homo sapiens

<400> 7	
cgtagggtagc ggctgcagt gctggcgccg tcctcgccc gccaggccat gaaggtagag	60
tttgaccgc tcaacatcca gctggcgccg cggctgcaga cggtaggcccgt gctgcagtgg	120
gtcctttctt ttcttacagg gccgatgtcc attggaatca ctgtgatgct gatcatacac	180
aactatattgt tcctttacat cccttatttg atgtggcttt actttgactg gcatacccca	240
gagcgaggag gcaggagatc cagctggatc aaaaattgga ctctttggaa acactttaag	300
gactattttc caattcatct tatcaaaact caagatttgg atccaagtca caactatata	360
tttgggtttc acccccatgg aataatggca gttggagcct ttgggaattt ttctgtaaat	420

pctau2012001598-seql

tattctgact	tcaaggacct	gtttcctggc	tttacttcat	atcttcacgt	gctgccactt	480
tggttctggg	gtcctgtctt	tcgagaatat	gtgatgagtg	ttgggctggg	ttcagtttcc	540
aagaaaagtg	tgtcctacat	ggtaagcaag	gagggaggtg	gaaacatctc	tgtcattgtc	600
cttgggggtg	caaaagaatc	actggatgct	catcctggaa	agttcactct	gttcatccgc	660
cagcggaaag	gatttggtta	aattgctttg	acccatggcg	cctctctggg	cccagtgggt	720
tcttttggg	aaaatgaact	gtttaaacaa	actgacaacc	ctgaaggatc	atggattaga	780
actgttcaga	ataaactgca	gaagatcatg	gggtttgctt	tgcccctggt	tcatgccagg	840
ggagtttttc	agtacaattt	tggcctaata	acctatagga	aagccatcca	cactgttggt	900
ggccgcccga	tccctgttcg	tcagactctg	aacccgaccc	aggagcagat	tgaggagtta	960
catcagacct	atatggagga	acttaggaaa	ttgtttgagg	aacacaaagg	aaagtatggc	1020
attccagagc	acgagactct	tgttttaaaa	tgacttgact	ataaaaaaaaa	attaaaaaat	1080
aaaaataaat	gac					1093

<210> 8
 <211> 1008
 <212> DNA
 <213> Mus musculus

<400> 8	
atgatgggtcg	agttcgcgcc actcaacacc ccgctggcac ggtgcctaca gaccgctgcg 60
gtgctgcagt	gggtcctgtc cttcctcctg ctcgtgcagg tgtgcattgg aattatgggtg 120
atgctgggtcc	tgtacaacta ttgggttcctt tacatcccat atctgggtctg gttttactat 180
gactggagaa	ccccagagca aggaggcaga agatggaact ggggtccaaag ctggcctgtg 240
tggaagtatt	ttaaggagta ttttccaatc tgtcttgtca aaacgcagga tttggatccg 300
ggtcacaatt	atatatttgg gtttcaccct catggaatat tcgtgcctgg agcctttgga 360
aattttttgta	caaaatactc ggacttcaag aagctatttc ctggcctttac atcgtatctc 420
cacgtggcca	agatctgggt ctgtttcccg ttgttccgag aatatctgat gagtaacggg 480
ccggtttcag	tgtctaagga gagtttgtct catgtgctga gcaaggatgg aggtggcaat 540
gtctcaatca	ttgtcctcgg aggtgcaaag gaggcgtgg aggctcacc aggaacattc 600
accctgtgca	tccgccagcg caaagggttt gttaagatgg ccttgacca tggtgccagt 660
ttgggtccag	tattttcttt tggtgaaaat gatctatata agcaaattaa caaccccaaa 720
ggctcctggc	tacgaactat acaagacgca atgtatgatt caatgggagt agccttgcca 780
ctgatatatg	ccagaggaat tttccagcac tactttggca taatgcccta tcggaagctg 840
atctacactg	ttgttggccg ccctatccct gttcagcaga ttctgaaccc gacctcagag 900
cagattgaag	agctgcatca gacataccta gaggagctaa agaaactatt caatgaacac 960
aaagggaaat	atgggattcc ggagcacgaa actctggtat ttaaataa 1008

<210> 9

pctau2012001598-seq1

<211> 1093

<212> DNA

<213> Pan troglodytes

<400> 9

```

cgtgggtgca ggctgcagtg gctggcgccg tcctcgcccc gccaggccat gaaggtagag      60
tttgcaccgc tcaacatcca gccggcgcgg cggctgcaga cgggtggccgt gctgcagtg      120
gtccttttctt ttcttacagg gccgatgtcc attggaatca ctgtgatgct gatcatacac      180
agctattcgt tcctttacat cccttatttg atgtggcttt actttgactg gcatacccca      240
gagcgaggag gcaggagatc cagctggatc aaaaattgga ctctttggaa acactttaag      300
gactattttc caattcatct catcaaaact caagatttgg atccaagtca caactatata      360
tttgggtttc acccccatgg aataatggca gttggagcct ttgggaattt ttctgtaa      420
tattctgact tcaaggacct gtttcctggc ttacttcat atcttcacgt gctgccactt      480
tggttctggt gtctgtctt tcgagaatat gtgctgagtg ttggcctggt ttcagtttcc      540
aagaaaagtg tgtcctacat ggtaagcaag gagggaggtg gaaacatctc tgtcattgtc      600
cttgggggtg caaaagaatc actggatgct catcctggaa agttcactct gttcatccac      660
cagcggaag gatttgtaa aactgctttg acccatggcg cctctctggt cccagtggtt      720
tcttttggtg aaaatgaact gtttaacaa actgacaacc ctgaaggatc atggattaga      780
actgttcaga ataaactaca gaagatcatg gggtttgctt tgcccctggt tcatgccagg      840
ggagtttttc agtacaattt tggcctaata ccctatagga aagccatcca cactgttggt      900
ggccgccccg tccctgttcg tcagactctg aaccggaccc aggagcagat tgaggagtta      960
catcagacct atatggagga acttaggaaa ttgtttgagg agcacaaagg aaagtatggc     1020
attccagagc acgagactct tgttttaaaa tgacttgact ataaaaaaaa atttttaaat     1080
aaaaataaat gac                                                         1093

```

<210> 10

<211> 1008

<212> DNA

<213> Pan troglodytes

<400> 10

```

atgaaggtag agtttgcacc gctcaacatc cagccggcgc ggcggtgca gacggtggcc      60
gtgctgcagt gggtcctgaa atacctgctg ctcgggccga tgtccattgg aatcactgtg     120
atgctgatca tacacagcta ttcgttcctt tacatccctt atttgatgtg gctttacttt     180
gactggcata cccagagcg aggaggcagg agatccagct ggatcaaaaa ttggactctt     240
tggaacact ttaaggacta ttttccaatt catctcatca aaactcaaga tttggatcca     300
agtcacaact atatatttgg gtttcacccc catggaataa tggcagttgg agcctttggg     360
aatTTTTctg taaattattc tgacttcaag gacctgtttc ctggctttac ttcatatctt     420
cacgtgctgc cactttgggt ctggtgtcct gtctttcgag aatatgtgct gagtgttggc     480
ctggtttcag tttccaagaa aagtgtgtcc tacatggtaa gcaaggaggg aggtggaaac     540

```

pctau2012001598-seq1

atctctgtca ttgtccttgg ggggtgcaaaa gaatcactgg atgctcatcc tggaaagtcc	600
actctgttca tccaccagcg gaaaggattt gttaaaactg ctttgacca tggcgctct	660
ctgggtcccag tggtttcttt tggtgaaaat gaactgttta aacaaactga caaccctgaa	720
ggatcatgga ttagaactgt tcagaataaa ctacagaaga tcatgggggt tgccttgccc	780
ctgtttcatg ccaggggagt ttttcagtac aattttggcc taatgcccta taggaaagcc	840
atccacactg ttgttggccg cccgatccct gttcgtcaga ctctgaaccc gacccaggag	900
cagattgagg agttacatca gacctatatg gaggaactta ggaaattgtt tgaggagcac	960
aaaggaaagt atggcattcc agagcacgag actcttgttt taaaatga	1008

<210> 11
 <211> 1008
 <212> DNA
 <213> Canis familiaris

<400> 11	
atgaaggtcg agtttgcgcc gctccacatc ccgctggcgc ggcggctgca gacggccgcg	60
gtgctgcagt gggtcctgtc cttcctgctg ctgcgcagcagg tgtgcatcgg aatcatcgtg	120
gtactgatcg tgcacaacta ttggttcctc tacgtccctt atctgacatg gctttgcctt	180
gactggcggg cccagagca gggaggcagg agatccaact gggtcagaag ctggaccgtt	240
tggagggtatt ttaaggacta ttttccaatt cacctcatca aaacttgga tttggatcca	300
agtcacaact atatatttgg gtttcacccc caggagtgct ttgttgctgg agcctttgga	360
aatttttgta cgaatcattc agacttcgag gagctgtttc ctggcttcac tgcgtatctt	420
cacgtgctcc cattttgggt ccggtgtcct ctctttcgag aatatctgat gactagcggg	480
tcagtctcag tttctaagaa aagtgtgtcc tatgtgtcga gcaaggaggg aggtggaaac	540
atttcagtca ttgttcttgg ggggtgcagaa gaatcactgg atgcccattc tggaaaattc	600
actctgttca tccgccagcg gaaaggattt gttaaagtgg ctttgacca tgggtgcctcc	660
ttgggtcccg tgttttcttt tggagaaaat gaactattta agcaagttaa caaccctgaa	720
ggctcatggc ttcgaactgt gcaggagaag ctacagaaga tcatgggatt tgccttgctt	780
ctgttccatg ccagagggat ttttcaatac aattttggcc tgatgcccta taggaaacct	840
atccacactg tggttggtcg cccaatccct gttcatcgga ctccgcaccc aagcccggga	900
cagattgagg agttgcatca gacctatatg gaggagctaa ggaaattatt tgaagcacac	960
aaaaggaagt acggtattcc cgagcatgaa actctcattt ttaaataa	1008

<210> 12
 <211> 1118
 <212> DNA
 <213> Bos taurus

<400> 12	
atgaaggtcg agtttgcgcc actcaacatc ccgctggcgc ggcggctgca gacggctgcg	60
gtgctgcatt ggttgctgtc cttcctgctg ttgcacagcagg tatgcctcgg aattatcgtg	120

pctau2012001598-seq1

ttcctgatca tatacaacta ctgggttcctc taccttcctt acttgacatg gctttacttt	180
gactggcaaa cccagagca aggaggcaga agatctgaat gggtcagaaa ctgggccatt	240
tggaggtact ttaaggacta ttttccaatt catctcatca aaacctggga tttggatccg	300
agtcacaact acatatttgg gtttcacccc catggagtgc ttgtggttgg agcctttgga	360
aacttctgta caaattattc ggccttcaag gagctgtttc ccggctttac ctcctatctt	420
cacgtgctgc cgtattgggt ccggtgtccg ctcttcgagg aatatctgat gagcagtggg	480
ccagtctcag tttccaagaa aagtgtgtgc cacgtgttaa gcaaagaggg aggtggaaac	540
atctcagtca ttgtgctcgg ggggtgcagag gaatcactgg atgcccatcc tggaaagtgc	600
actctgttca tccgccagcg gaagggattc gtgaaaattg ctttgacca tgggtgcttat	660
ttggtgccag tgttttcttt tgggtgaaaat gaactattta aacaagttag caatcccga	720
ggatcatggc ttcgaaatgt gcaggagaaa ctacagaaga tcatgggatt tgctttgcc	780
ctgttccacg ccagaggaat ttttcaatac aattttggcc taatccccta taggaagccc	840
attcacactg ttgttgccg cccaatccct gttcgtcaga ccctgaaccc aacctcagag	900
cagattgagg agttgcatca gacctacatg gaggagctga gaaaattatt tgaagaacac	960
aaagggaagt atggtattcc agaaaatgaa actctcatct ttagataact gttgtcttag	1020
aagtcatgca agagagcgag agaacatccc actgaaagaa aagatattct aaataaattt	1080
ttttctttga caatggaatc aaatttgcaa ctgtaaaa	1118

<210> 13
 <211> 749
 <212> DNA
 <213> Rattus norvegicus

<400> 13	
ttgactcact aagggccagt ttcagtttct aagaggagtt tgtctcatgt gatgagcaag	60
gaaggaggtg ggaacatctc agtcattgtc atcgggggtg cgaaagagct actggagtcc	120
ttccctggaa gatattctct gtgccttctt cagcgcaaag gatttgtaa gatcgcttg	180
acccatggtg ccatttgggt tccagtgttt tcttttggtg aaaatgaact atatagccaa	240
gtcgacaacc ctaaaggctc atggcttcgg actgccaag acaaagtgtg taacctaacg	300
gggttgccct tgccactatt ctatgccaga ggaattttcc agaacagctt tggcctaattg	360
ccctatcgga aattgatcta cactgttgct gcggttgga cagcagctcg tttgactcag	420
aagtcaaacc atcagaacga cagctcagtg tctgctgaga agggaaactc ttttatttgt	480
tgtcaagtgg gatggtgtga ccgagatagg cagccattgg ccgccccatc cctgttcaga	540
agactctgca cccgaccca gagcagattg atgagctgca tcagacgtac atggaggagc	600
taaggaaact cttcaatgaa cataaaggga aatacggat tccggagcat gaaactctgg	660
tttttaaata accgtcttcc tccaagccat actgaaggaa gacaagaaaa cattgttgag	720
aaaataaata ttttaaataa cttctctgt	749

pctau2012001598-seq1

<210> 14
 <211> 1547
 <212> DNA
 <213> Danio rerio

<400> 14
 tttttttttt ttttttcctg tttgagctgt ctttattatc gtcctcgcca ttgagttcgc 60
 tcctctgcgc gtgcctctgc gcaggcgcat tcagacggcc gcggtggtcc agtgggtctt 120
 cagcttcttg gctctggctc aggtgtgtct ggccgcgttc gtgctgctgt gtctcagtga 180
 ttggtggatg ctcgagcggc tgtatgctgg ttggctgtat ctggaccggg acacgccctc 240
 atgtggaggt cggaggtcac agtgggtcag aagctggagg atctggatgc atttcagggg 300
 ttatttccca atcacactgg tcaagacggt ggatctggac ccgaggcaca attacctgct 360
 gggtttccat ccgcacggtg ttctggtggc cgggtggattc ggggaacttct gcacggaggc 420
 gtccggattc tcccagatgt tcccgggact cagcccgta ctgctgatgc tgccgttctg 480
 gttcaggggt cccttcttca gagaatacat catgtgcgga gggctggtgt ccagtgagaa 540
 ggccagtgcg agttacctgc tgggtcatcc aggcgggggt caagcggccg tcattgcagt 600
 gggcggagct ccggagtctc tggaggcgcg gcctggagca ctgacactgc aactgctgca 660
 gcgcaaaggc ttcatcaaac tcgccctcaa acacggagcg tggctgggtc cagtgttctc 720
 gtttggggag aacgagctct tcgaccagat ggagaatccc gcaggctccg ccctccgccg 780
 catgcaggaa cgcctccaga ggataatggg cgtggctttg cctctctttc atgctagagg 840
 agttttccag tacagcttcg ggctgctgcc gtacaggaaa cccatacaca ccgtcgttgg 900
 acgtcccatt cctgtaagtc agactccgtg tcccagtaaa gaagacatcg acgctcttca 960
 cacactctac atgcagggcc tgacacaggt gtttgaggag aacaagaaac attacggcat 1020
 cgcggacgac aaacatctga agttcacctg aactgatctg tgtgtgtgtg tgcgcgtgcg 1080
 tatgtgtttg tttgtgtgtg tgtgtgtgtg tgtgtgtgtg cgttgcatat aatctagaat 1140
 gattatctac catctagtta aaggtcttgg tacgaggggt tgtgtagttg tgttgtcatg 1200
 tgatatgacg atattcagaa taaccacag cttgagctga agtacgggtta ctacacagtg 1260
 tataaaacat taaggagatc cttttactgt tctctcctt cactacagta tggtaaatat 1320
 tgttactgta ttttatgact tcattaaaac actatgtgac cgccattaag attattaaat 1380
 gttcagatta aacaggtcat gaacactctg ggttttacag atcttcagga ttccctcctg 1440
 tgcttaaatac agatttggtc ttaatcagca aatgtgtgtt gatgtgtatg acaacagaaa 1500
 cacaaataaa gcggtgctgt ttattatattt aaaaaaaaaa aaaaaaa 1547

<210> 15
 <211> 1279
 <212> DNA
 <213> Caenorhabditis elegans

<400> 15
 attcttcatt ggacctctac gaccaccaca cagtcagtca tgttagatcg aacagaaagt 60
 gaagataaga tgccacatct actaggagtt gagtgggctc cgctcaatat tccgttggct 120

pctau2012001598-seq1

cggcgtcttc aaactttggg agcatttcat ttcttcttca tcactctctt cacaccagta	180
ctcgtttctca ccgttccatt ctacatgtta tataccgtac tctggccttt gatctttctg	240
tatgggcttt ggatgattta cgattggaat tcaccaaaga agggagccta tatgagcaat	300
tggttccaga gacaaagaat tcattcgtgg tatgccaact attttccagt caaattgcac	360
acaacatctg acatgccaga agaacataac tatttgattg ggtaccatcc gcatggaata	420
atttcaatgg ccgcattcat caactttgca acaaatggaa ctggaattct cgatactctt	480
ccacgaattc gtttccattt gtgcacactt gttggtcaat tctggactcc gtggagacgt	540
gagtggggat tggtgcacgg aatgatagac tgcagtcgag aaagcatcaa gcacgttttg	600
gagcatgaaa agaaaggaaa agcagttgta ttggtggttg gtggagctga agaagcactt	660
gatgcacatc caggatgcc aattttgact ttgaaaaaaa ggaaaggatt cgtgaaaatt	720
gccctgcaaa ctggagctca actggttcca tgctattcat tcggtgaaaa tgatattttc	780
aatcaagccg aaaatccaaa gggatcaaca attcgacagt tccaaacgat aatgaaaaga	840
gtcttgggat tctcccctcc agcattctat gggagaggag tattcaacta tacatttggg	900
cttcttccat tcagaaaacc tatcaacact gttctcggcg ctccaatttc agtgacaaag	960
acagtgaatc caactcaaga acaaatcgac acatttcac agacatacat ggaccgtctt	1020
catgagcttt tcgaggagca caagacaaaa tacgatgtct ctccaactac acaacttggt	1080
atcaattaaa tattatctat ttaatcacgt ttgatgcact tttctcatat ttacatctct	1140
attatacctc attctgcaat ttagttcata ttttcagcca atttttcaat attacgagtt	1200
ctcatgtcaa atatcccat cagattcagt tactagttca gttattttta atcgatctct	1260
attataagtt tcgacacag	1279

<210> 16

<211> 1089

<212> DNA

<213> Caenorhabditis elegans

<400> 16

attcttcatt ggacctctac gaccaccaca cagtcagtca tgtagatcg aacagaaagt	60
gaagataaga tgccacatct actaggagtt gagtgggctc cgctcaatat tccgttggct	120
cggcgtcttc aaactttggg agcatttcat ttcttcttca tcactctctt cacaccagta	180
ctcgtttctca ccgttccatt ctacatgtta tataccgtac tctggccttt gatctttctg	240
tatgggcttt ggatgattta cgattggaat tcaccaaaga agggagccta tatgagcaat	300
tggttccaga gacaaagaat tcattcgtgg tatgccaact attttccagt caaattgcac	360
acaacatctg acatgccaga agaacataac tatttgattg ggtaccatcc gcatggaata	420
atttcaatgg ccgcattcat caactttgca acaaatggaa ctggaattct cgatactctt	480
ccacgaattc gtttccattt gtgcacactt gttggtcaat tctggactcc gtggagacgt	540
gagtggggat tggtgcacgg aatgatagac tgcagtcgag aaagcatcaa gcacgttttg	600

pctau2012001598-seq1

gagcatgaaa agaaaggaaa agcagttgta ttgggtggtg gtggagctga agaagcactt	660
gatgcacatc caggatgcc a tttttgact ttgaaaaaaa ggaaaggatt cgtgaaaatt	720
gccctgcaaa ctggagctca actggttcca tgctattcat tcggtgaaaa tgatattttc	780
aatcaagccg aaaatccaaa gggatcaaca attcgacagt tccaaacgat aatgaaaaga	840
gtcttgggat tctcccctcc agcattctat gggagaggag tattcaacta tacatttggg	900
cttcttccat tcagaaaacc tatcaacact gttctcggcg ctccaatttc agtgacaaag	960
acagtgaatc caactcaaga acaaatcgac acatttcac agacatacat ggaccgtctt	1020
catgagcttt tcgaggagca caagacaaaa tacgatgtct ctccaactac acaacttggt	1080
atcaattaa	1089

<210> 17
 <211> 1149
 <212> DNA
 <213> Caenorhabditis elegans

<400> 17	
tcagctttcc actgaaaatg ccgcaatttc tcggaataga atgggtggat ctattctcat	60
caattcaacg taaaaagaca tacttgggag ttgtctatca ctttatgctc acttaccac	120
ttgcactttt cgtcaccatt ttgccatttt tcttgctttt cactttccaa tggcacattt	180
tggctctcta cgcttgctgg tacttctacg atatggattc tccgaggaga ggtggatatt	240
ccagtgattg ggtcagaaaa tggcgtgtca acgactgggt cgcccaatac ttcccaatca	300
acttgcacaa aactgccgaa ctctccacgg acaagaacta tctggttgga atccatcctc	360
atggtatcat ctccatggct gcatggtcta attttgccac caatggaacg ggaatctatg	420
agaaattccc tggaattcgc tggaatttgt gcactttagc acttcagttc agaatggcca	480
tccgtcgtga gcttctcctt ctcaccggcc taatcgactg ctccagagaa tccattgaat	540
atgtgcttga caaatgtggc cagaaggac gtgcagtggg attggtgatt ggaggagccg	600
aagaagcttt agatgctcat ccaggctacc acacactcac tttggcatca agaaagggat	660
ttgttcggga ggctttgatt actggtgcct acttgggtccc ggtgtattct tttggagaga	720
atgacgtttt tgaacagatg gagaatccag ttggctcacg gctccgcaac ttccaagaat	780
ggtgcaaaag cattttcggc atctcatatc caattttcca tggctcgtggc ttcttccaac	840
taactttcgg atatcttcca ttccgtaaac caatcgatac cgtagtcgga gcccgaattc	900
ccgttgaaaa agtggaaaat ccgacgaaag agcaaatcga cgagcttcac accatttatt	960
gtcaaaagt gacggagctg ttcgatgagc acaaggaaaa atatggagtg gagaaggacg	1020
tgccgttggg gctccggtag agcggaaaat ttgaaatgtt ctgtgaaaag tgccagtttt	1080
tgatgagatt tctgtatttt atcgtgccat aatttttgtt tttcagacct tgttttaagt	1140
tgttacggt	1149

<210> 18
 <211> 1136

pctau2012001598-seq1

<212> DNA

<213> Caenorhabditis elegans

<400> 18

atgctaaact accaaattca caaaaagctc accgacatca agtgggtgaa catcttctcc	60
ccatgggatc gccagcgtgc ctacttcgcc ttggtcgtct ggttcgggct catctaccca	120
ttctgctgcc tgtgccaggt ggctccgttt gtgctctttt tcaccggcca gtggattatt	180
ttgggtctct acgcagtttg gtacctttac gatcgagaat ctccgagaag aggaggatat	240
cgggataatt ggttcagaaa tttgtcgtcg cacaagtggg tcgccgagta ttttcctggt	300
aaacttcaca aaactgcgga gttggatcca aacaaaaatt atttattcgg atatcatcct	360
catggaattc tcggtgtcgg agcgtggtct tgttttggat ttgatgcgtg caatgtgaag	420
caagtgttca aaggcatccg cttcaacatc tgcaccttgc ccggcaactt caccgcaatg	480
ttccgccgag agatcctcct cagcatcggg atgatcgaga gtcctcaaaga atccatcgag	540
cacgtgctca actccgagga aaagggccgt gccgttgtaa ttgtcgtggg tggagccgct	600
gaagctcttg aagctcacc aggaagcat actctaacac tggcaaactg caaagggttc	660
gtgagagaag ccgtgaagac cggagctcat ctggtgccag tttatgcgtt tggagagaat	720
gacatatata agcaaattga caaccggaa ggctcgaaat tacggaaaat tcaagaatgg	780
ggaaagaaga aaatgggaat ttcactgcca ctaatctacg gaagaggata ttttcaaag	840
gctcttgggc ttcttccaat gagccgagct gtgaatgtag ttgtcggagc gcctattcaa	900
gtggaaaaag agctcgatcc ttctaaggaa gtcattgatg aaattcatgg agtttatatg	960
gaaaagctcg ccgagttatt tgaagagcac aaggcaaagt tcggagtttc caaggacact	1020
cggctcgttt ttcagtgaat taacttttat tttagattat ttgtttgatg tttgatgtct	1080
gttatgttat ctttttttat caccaaactg tgtaattttg tttacggggt cattac	1136

<210> 19

<211> 1176

<212> DNA

<213> Caenorhabditis elegans

<400> 19

atgagactcc ggctgagctc gatatctgga aaggcgaagc ttcccgataa agaaatatgc	60
tcatcagttt cgagaatatt ggcaccattg cttgttccat ggaagcgacg actcgagact	120
cttgccgtga tgggtttcat tttcatgtgg gtaatcctac caatcatgga cctctgggtt	180
ccattccacg tcttgttcaa tactcgatgg tggttccttg ttccactcta cgctgtctgg	240
ttctactatg attttgatac accgaaaaaa gcttcaagaa gatggaattg ggccagaaga	300
cacgtagcct ggaagtactt tgccagctac ttcccattga gattgatcaa gactgctgac	360
cttccggcgg atcgtaatta catcattggc tctcatcctc atggaatgtt ctcggttggt	420
ggttttactg caatgagcac caacgcgacc ggatttgaag acaagttccc gggaataaaa	480
tctcacatca tgacgctaaa tgggcaattt tatttcccat ttcgtcgaga attcggaata	540
atgctcgggtg gaatcgaagt ttcgaaagaa tcacttgaat acactctaac taaatgtgga	600

pctau2012001598-seq1

aaaggacgag	catgcgcaat	tgtcattggc	ggagcctcgg	aggctcttga	agctcatccc	660
aataaaaaata	cattgacgtt	gatcaatcga	cgtgggtttct	gcaaatatgc	tctgaagttt	720
ggagcagacc	tcgtaccaat	gtacaatttc	ggagagaatg	atttatacga	gcagtatgaa	780
aacccgaagg	gatctagatt	gcgagaagtt	caggagaaaa	tcaaggacat	gttcggattg	840
tgtcccccct	tgctccgagg	tcgatcggtg	ttcaaccaat	accttatcgg	attgctgccg	900
ttccgaaaac	cagttacaac	agtcatggga	aggccaattc	gagtcaccca	aaccgacgag	960
ccaaccgttg	agcagattga	tgagctgcat	gcaaaatatt	gtgatgctct	ctacaatctg	1020
ttcgaggagt	acaagcatct	tcactccatt	cctcccgaca	ctcatctcat	cttcagtgga	1080
ctttttcaca	aattttgtat	attttgcaca	cagattgttt	tatcgccaag	ttaatgtata	1140
gtaatcttct	tgtaaataaa	tcagataaac	attaaa			1176

<210> 20

<211> 1239

<212> DNA

<213> Kluyveromyces lactis

<400> 20

atgggtaagg	aacaaatcag	aaagagacaa	actaaaaagg	tacgcaaggg	ttcacaaccc	60
gtggcaaatg	gtaaattctt	tgtagtaggg	gtctccaaag	atgggaaacc	agaattttgt	120
gctattgata	cgcccataga	aaggagactt	caaactctca	gcgtagcatg	gtatgtcgaa	180
tgtattccgt	taatgatcat	cataatgttg	tttgtttggg	ttaaccggtt	gatgtggtcg	240
tttgaatac	catacaccat	ttattatttc	atagaccgga	ctgcctccaa	tgggaacgct	300
gtgaagagac	actccaagtg	gttccgatct	ctaaagggtc	ggttttatct	ccgtgattac	360
ttccctataa	gtatgcacaa	aagcactgaa	ctcgagccaa	catttacaag	tattgattcc	420
acagagctgg	aaaatgatgc	gagtgaacca	ggctacctgg	atagctcaca	gcccgtttta	480
ccagataaat	ggtggaatcc	tttcagagaa	aaggatgaaa	ctgtagacc	aacaggaccg	540
aggtatattt	ttggctacca	tccacatgg	atagcagcat	tcggagcggt	tggagcattc	600
gctaccgagg	catgtaattg	gagtaaagtt	ttcccaggca	ttccagtgtg	tttgcttact	660
ttggtaaata	aatttcaaat	accagtgtac	cgtgattatt	tattggcctt	aggcataaca	720
tcggtcgcta	gaaaaaatgc	tatgaaagta	cttgagaaaa	attattcaat	agccatcgta	780
attggagggtg	ccagtgaatc	ccttcttact	aaccttggtt	catctgatat	cattttaagc	840
aagcggaag	gttttgtgaa	gttggctttg	caaactggga	atgtgagtct	tgtacctgtg	900
tatggatttg	gtgaaactga	tacttacaag	attctgaagc	ttaaaaatga	ctctataata	960
ggcaggatac	agatatggct	taaagagaac	tacagcttca	ctgtccctct	atttttcgct	1020
cgtgggtgat	tcaattatga	ttttggttta	ctaccgttta	gacatccagt	caatgttggt	1080
gttggaaacc	ctattcatat	taaagaaaag	atagatcatc	caacaattga	ggagattgat	1140
cattaccatt	ccttatacat	tgaagaattg	aaacgtcttt	acgatgataa	caaggcaaaa	1200

pctau2012001598-seq1
tttaactatt cagaaaaaac tttaaataata gtggaatga 1239

<210> 21
<211> 1386
<212> DNA
<213> Ashbya gossypii

<400> 21
atgcaggaca gcatggatga cagcttgagg gaggtgagg gcagacagga cgactcggag 60
gtcagcagcg ggacgacact cggaagcagc accccggagg acagcggggt gacggcgaag 120
ctgcgcaaga agtaccagat ggcgtcggca ctgctgcggc gggagctgga ggagctgagc 180
gtgtacgacg cgaagacggc gggggtaagc ggccgcagca gcggcagcgg cagcggcggg 240
ctggcgctgc tgggcggggc attccacgtg gcgccgctgc ggatccccgc gcggcggcgg 300
ctgcagacgc tgggtggtggc gtggcacacg agctcgttca tctacatgac ggtgctggtg 360
ctgtttcttg cggcgaacct gctgatgtgg tggttcatgg tgccgtatat ggtgtactac 420
gtgtggaacc ggtcgccggc gaacggcgga gtggtgcggc ggtactcgcc gcggctgcgg 480
tcgctggcgc tgtggcggtg ctactgcgag tactaccga tctactgca caagagcgag 540
gacctggcgc cgacgttcgt gccggacctg cggggggcgg agccgcgcga gtggaagctg 600
cggctgtggc tgtggccgac gcgcgtggag ctgctgaact tgacgctcca gtggacgcgc 660
gcgcggccgc aggttgcgac agggccgcgc tacatcttcg ggtaccacct gcacggcgtg 720
ggcgactgg gcgccttcg cgcgatcgcg accgagggct gcaactggag caaggtcttc 780
gcgggcatcc cggcctgcct gtgcacgtg gtgaaccagt tccagatccc catctaccgc 840
gactacctgc tgggcctcgg gtgcacctcc gtggcccga agaactgct caaggtgctc 900
gagcagaact actcggctctg catcgtcgtg ggcggcgccc aggaagccct gctcagccgt 960
gtgggctcca ccgagctcgt gctgaacaag cgcaagggct tcatcaagct cgcgctggag 1020
acgggcaacg tcaacctggt gcctatctac gccttcggcg agacagactg cttcaacgtg 1080
ctcgacacgg gcaacgagag ctacctacgc aagttccagc tgtggatcaa gaagacctac 1140
ggcttcacca tcccgttctt ttttgccgc ggtgtcttca actatgactt cggcttcctg 1200
cccttcagga atccgatcaa cgtcgtcgtg ggcaagcctg tgtacgtaga caagcgcaga 1260
acgaacccca ccatggagga aatcgaccac taccacgacc tgtacgtgca ggagctgcgc 1320
aatgtgttcg acaagaacaa acacaagttc ggctacgccg ggaaggaact taaaatagta 1380
gagtag 1386

<210> 22
<211> 1653
<212> DNA
<213> Magnaporthe oryzae

<400> 22
atggccacgg tagatgatac ccagaacct gtccaactga atgcagagtc cagcctaaac 60
gaatttccgt accaagcaag cgcggacaag gaggagtatc caagccttcc ttctgccccg 120

pctau2012001598-seq1

ctccctgacg	tggacggcag	ctcccaaat	gggcatacag	ccgccaacgc	tgaccttccg	180
ccctcggagt	cgaccgaaca	gccatggaac	aaggtagcca	gtggcgactt	catccatgaa	240
accaagatgc	ccagtgaaga	ccgggacaac	cttgccaggc	ccggcagtct	accaactccc	300
gagccgagcc	gtagaggcag	caacacgctg	gtcggcagcc	cgcgtcaggg	cgacggattg	360
caaatgctca	tggacgaatt	cgcagacgat	gaatcctcag	atacctcgtg	ggacattcac	420
gagaccatca	ggagtcgcac	ccaggccttg	gctgagaagg	accaccaggg	gatatttctca	480
aatcagcggc	gaggcaacca	ggccgacgcg	ccagacacgg	taaccagtgc	caacgcttgg	540
ttgccggctg	gcttcgaaga	cgacctggac	gaagtgaccg	gtgctgacga	cgagttcgag	600
ccaacccgtg	ggagtcgctg	gagagtgagg	ctggcgcccc	tggatacccc	ttttcatcgc	660
cggatgcaga	ccctggtggt	tctgctacac	gtcctcggca	tgggaatcac	tttcagcttt	720
ttctgcttcc	tatgcacgct	ccccctgttt	tggcctatca	tcatcgcccta	cctcgtattc	780
atcagactgt	cgagagctgg	atccgatggg	aagacaaacc	gtcgcataga	atggctccgc	840
agggctcgca	tctggaagta	ctttgccgat	tacttccctg	tcaagctaca	caagacagcc	900
gacctgcccc	ccacccgcaa	gtacattttt	gcgattcatc	cacacggaat	catctcgcat	960
ggtgcctttg	cctccttcgc	cacggaggcc	ctcggcttct	ctgagaagtt	tcccggcatc	1020
acaaacagtg	tgtgactctt	tcacggcaac	ttcaagacc	cgttctaccg	cgagtatctg	1080
atggccatgg	gcctgatctc	ggtttcaaag	gagtccatta	ttaacacttt	gatgaccggc	1140
gggatcaacg	gtgaaggcat	gggcagggcc	gtctccataa	ttgtcgggtg	ggcgcgagag	1200
gcactcgagg	cctgccccaa	gacgatgcgc	ttgatactga	aacggcgggg	tttttgtcgc	1260
atggcactac	gcaccggtgc	agaccttgtg	ccggtgctct	gctttggtga	gaacgacctc	1320
taccaacagt	gggggcccga	ggatcatccg	cgcttccgtc	gcctgcaa	gaaggctctc	1380
gcgtacctag	gctttgcggt	ccccgttctg	aggggaaggg	gcgtcttcaa	ctatgacttt	1440
ggagtccctc	ctcagcggcg	gccgatcaac	gttggtgttg	gtgagcctat	caaggttgag	1500
cagtttcggg	gtagcggtaa	cattgagcca	cgcgtcgagg	agctgtggaa	gctatatattgc	1560
cagaagcttc	aggaactgta	cgaccagaac	aagcacgtct	atttcaagga	tcgtcaggag	1620
gatatgaggt	ttgttgactc	ggaggagttg	taa			1653

<210> 23
 <211> 1410
 <212> DNA
 <213> *Ciona intestinalis*

<400> 23	
gccataaaac	gttgagaaaa tgttttgtgt tacgtcaagt tactcattaa ctcaatgtaa 60
agattgatat	agccagagca aattgtagtt aagtagcgca tatacccaac ttggaaactg 120
ctatatagtt	ggtaaagcgg attaggaacc atgtgggaat tcgcaccgct tcgtattcct 180
atgataagaa	gacttcaaac gttggctgtt tctcattttg tgttctgctt tcttgcccta 240
gctgccacat	gcactttatt gactatctac attactttgt tcacttcata ctggccgatac 300

pctau2012001598-seq1

accaccgctt	accttgtgtt	cttattcttg	gaccgaaaca	ccccggaatc	tggtggaaga	360
agaagcgagt	gggtgagaaa	ttggagtctc	tggaatgga	tgagcgatta	ctttccttgt	420
acgttgca	aaaccgttga	cctcgatcca	aaaagaaact	acatatttgg	catccatcca	480
catggggttt	tgtgcatcgg	atctttcact	cacttctcca	cgaatggctc	cggattttcg	540
cacgtttttc	cgggattttac	atcttatctt	acgatgttgc	ctttctgggt	caaaatgccc	600
ttttttcgtg	actatgttat	gagcggcggt	ttgactccag	ccacaaggaa	agccatcaaa	660
cacacaataa	caaggccagg	tggtgggcat	atatgctgca	ttatacctgg	tggtgctcca	720
gaatctctaa	acgcacggcc	tggagatgtt	gtattgttgc	tgaaacaaag	acttggattt	780
ctcaagctgg	caataacgaa	cggagttcct	ttggttcccg	tgttttcatt	tggcgatcac	840
gcgttgtggg	agcagaaacc	gaatcccccc	ggttcctca	tccgtaggtt	ccaagatagc	900
tctcagaagt	ggatgcaagt	cgctcttccc	gttttccacg	cacgggggtat	ctttcagtag	960
aattttggac	taattccgta	caggcgtagc	gtccataccg	tagtgggcca	gcccatagaa	1020
gtgccacaga	acagtaacc	gacatcgag	gaccttatga	gtttgcaaga	ggattatatt	1080
aatcgattga	gagcaatttt	tgacgagcac	aagagcaaat	atctcccaga	agattgcaaa	1140
ctgatcatca	actaaactaa	actaaactat	gcaaccaaag	ctttttaaat	tgagttagtt	1200
atagtagtgt	gaagtaagat	agagcacctt	tagcaataat	gctgtccaaa	tttcttggtt	1260
caaacaatta	ataacggacg	ggagtcgtga	tgaaacgggt	ttagaatact	ttaagtgttt	1320
ttgtttacta	ccaaatgaaa	agataaggta	gaataaaaag	gtgtctcatc	ttccccacc	1380
gtactatata	tatacacgat	cgaatcattg				1410

<210> 24
 <211> 1005
 <212> DNA
 <213> Homo sapiens

<400> 24	
atggtagagt	tcgcgccctt gtttatgccg tgggagcgca ggctgcagac acttgctgtc 60
ctacagtttg	tcttctcctt cttggcactg gccgagatct gcactgtggg cttcatagcc 120
ctcctgttta	caagattctg gctcctcact gtctgtatg cggcctgggtg gtatctggac 180
cgagacaagc	cacggcaggg gggccggcac atccaggcca tcagggtgctg gactatatgg 240
aagtacatga	aggactatth ccccatctcg ctggtcaaga ctgctgagct ggacccctct 300
cggaaactaca	ttgcgggctt ccacccccat ggagtccttg cagtcggagc ctttgccaac 360
ctgtgcactg	agagcacagg cttctcttcg atcttccccg gtatccgccc ccatctgatg 420
atgctgacct	tgtggttccg ggcccccttc ttcagagatt acatcatgtc tgcaggggtg 480
gtcacatcag	aaaaggagag tgctgtcac attctgaaca ggaaggggtg cggaaacttg 540
ctgggcatca	ttgtaggggg tgcccaggag gccctggatg ccaggcctgg atccttcacg 600
ctgttactgc	ggaaccgaaa gggcttcgtc aggctcgccc tgacacacgg ggcaccctg 660

pctau2012001598-seq1

gtgccaatct	tctccttcgg	ggagaatgac	ctatttgacc	agattcccaa	ctcttctggc	720
tcctgggttac	gctatatcca	gaatcggttg	cagaagatca	tgggcatctc	cctcccactc	780
tttcatggcc	gtggtgtctt	ccagtacagc	tttggtttaa	taccctaccg	ccggccccatc	840
accactgtgg	tggggaagcc	catcgaggta	cagaagacgc	tgcattccctc	ggaggaggag	900
gtgaaccagc	tgcaccagcg	ttatatcaaa	gagctgtgca	acctcttcga	ggcccacaaa	960
cttaagttca	acatccctgc	tgaccagcac	ttggagttct	gctga		1005

<210> 25
 <211> 1728
 <212> DNA
 <213> Mus musculus

<400> 25		
agcatggtgg	agttcgcccc	cctgttggtgta ccatgggagc gcaggctaca gaccttcgcg 60
gtccttcagt	gggtcttctc	cttcctggcc ttggcccagc tctgcatcgt catcttcgta 120
ggcctcctat	tcacaagggtt	ctggctcttc tctgtcctgt atgccacctg gtggtacctg 180
gactgggaca	agccgcggca	gggaggccgg cccatccagt tcttcagacg cttggccata 240
tggaaagtaca	tgaaggatta	tttcctgtc tctttggtca agacagctga gctggaccct 300
tcccggaaact	acatcgcggg	cttcacccc catggagtcc tagcggctgg agcctttctt 360
aacctgtgca	ctgaaagcac	gggctttacc tcgcttttcc cgggcatccg ctcctatctg 420
atgatgctga	ctgtgtggtt	ccgggcccc ttcttccgag attacatcat gtctgggggg 480
ctggtctcat	cagaaaaggt	gagtgccgat cacattctgt ccaggaaggg cggcgggaac 540
ttgcttgcca	tcatcgttgg	gggcgcgcag gaggcactgg acgccaggcc tggagcctac 600
aggctgctgc	tgaagaatcg	caagggttc atcaggctcg ccctgatgca tggggcagct 660
cttgtgccaa	tcttctcctt	tggagaaaac aacctgttca accagggttg gaacaccctt 720
ggtacctggc	tgcgctggat	ccagaaccgg ctacagaaga tcatgggcat ctccctccct 780
ctcttccacg	gcagaggtgt	cttcagtac agctttggcc tcatgccctt ccgccagccc 840
atcaccacca	tagtggggaa	gccatcgag gtgcagatga caccacagcc ctcaaggag 900
gaggtggacc	ggcttcacca	gcgctatatc aaggagctct gcaagctctt tgaggagcac 960
aaactcaagt	tcaacgtccc	tgaggaccag catctggagt tctgctaagt gtctccagcc 1020
ggaagacagc	tgcattctgag	cgcctgcagg agtgtgggat tagggggact tccacagcca 1080
ccagacactc	ctacaaacct	agccacaact gccaagatgg aagagggggc agctcctaata 1140
cctgggattt	gaacctgcag	ccaaagctct gaggtctccc tgtccttggc ctgtctgcac 1200
atctgtagaa	tgggggaaaa	gcaggcagag agaaattcct gaggtctctt cccacagttg 1260
taatgtcatt	caaacatgac	caaaggacaa acaggagaaa agagaacaaa actgttcttc 1320
atctaccctt	gaggacagct	gcaagagaag ccagcacccc aggcctccct gtgcatgctc 1380
cctgatgctg	cttcttcctt	ctgaggcaga gacggggagc caagtctgcc ctggcaccta 1440
ctctatgttt	cttcagattc	tgggtcctct gagctatgat accaaaggag cccagaaggc 1500

pctau2012001598-seq1

agataaggag	ggcaggggtc	actgactatg	accgagggta	ggctctcttc	ccatatacctg	1560
agcctcagtt	tccccagcct	taatgacctg	ggagcgccac	actgctcacc	acagaggctc	1620
caccagagag	cctcttactc	atgctttcta	gtgaactcca	gcctctgtct	tggcactgaa	1680
gggcagcact	gtacatgtta	cctcaataaa	tgaaggaggt	ctgtctta		1728

<210> 26
 <211> 1128
 <212> DNA
 <213> Pan troglodytes

<400> 26	
aaaaacctgt	60
gggtgcctca	
gaccacagca	
gagctcacag	
aacctgcggg	
agccaggctg	
acccgccagc	120
atggttagagt	
tcgcaccctt	
gtttgtgccg	
tgggagcgca	
ggctgcagac	
acttgctgtc	180
ctacagtttg	
tcttctcctt	
cttggcactg	
gctgagatct	
gcactgtggg	
cttcatagcc	240
ctcctgttta	
caagattctg	
gctcctcact	
gtcctgtatg	
cggcctgggtg	
gtatctggac	300
cgagacaagc	
cacggcaggg	
gggccggcac	
atccaggcca	
tcagggtgctg	
gactatatgg	360
aagtacatga	
aggactatct	
ccccatctcg	
ctgggtcaaga	
ctgctgagct	
ggacccctct	420
cggaactaca	
ttgcgggctt	
ccaccccat	
ggagtcctgg	
cagtcggagc	
ctttgccaac	480
ctgtgcactg	
agagcacagg	
tttctcttcg	
atcttccccg	
gtatccgtcc	
ccatctgatg	540
atgctgacct	
tgtgggttcg	
ggcccccttc	
ttcagagatt	
acatcatgtc	
tgcaggggtg	600
gtcacatcag	
aaaaggagag	
tgctgctcac	
attctgaaca	
ggaaggggtg	
cggaaacttg	660
ctgggcatca	
ttgtaggggg	
tgcccaggag	
gccctggatg	
ccaggcctgg	
atccttcacg	720
ctgttactgc	
ggaaccgaaa	
gggcttcgtc	
aggctcgccc	
tgacacacgg	
ggcacccttg	780
gtgccaatct	
tctccttcgg	
ggagaatgac	
ctatttgacc	
agattcccaa	
ctcttctggc	840
tcctgggttac	
gctatatcca	
gaatcggttg	
cagaagatca	
tgggcatctc	
cctcccactc	900
tttcatggcc	
gtgggtgtctt	
ccagtacagc	
tttgggtttaa	
taccctaccg	
ccggcccatc	960
accactgtgg	
tggggaagcc	
catcgaggta	
cagaagacgc	
tgcatccctc	
ggaggaggag	1020
gtgaaccagc	
tgcaccagcg	
ttatatcaaa	
gagctgtgca	
acctcttcga	
ggcccacaaa	1080
cttaagttca	
acatccctgc	
tgaccagcac	
ttggagttct	
gctgagccca	
aagggcaggg	1128
ccaacattag	
ggagcccagc	
aggaggtgct	
gtgctgag	

<210> 27
 <211> 1005
 <212> DNA
 <213> Canis familiaris

<400> 27	
atggtgaagt	60
ttgcaccact	
gttcgtgcca	
tgggagcgca	
ggctgcagac	
cttcattggtc	
ctgcagtggg	120
tcttctcctt	
cctggccctg	
gccagatct	
gcactgtggg	
cttcgtaggc	
ctcctgttca	180
caaggttctg	
ggtggtcagc	
gtcctgtatg	
ccgcctgggtg	
gtaccgggac	
cgggacacgc	240
cgaggcaggg	
gggcaggccc	
gtccaagccc	
tgaggcgctg	
cttcttgtgg	

pctau2012001598-seq1

aagtacatga	gggactactt	ccccgtcacg	ctgggtcaaga	ctgctgagct	ggacccctcc	300
cggaactacc	ttgctggcct	ccacccccac	ggggctcctg	ccgcaggagc	ctttgtcaac	360
ctgtgcacgg	agagcacggg	cttccctttg	ctcttccttg	gcatccgctc	ccatctgatg	420
atgctgacct	tgtggttccg	ggtccctttc	ttcagggatt	acatcatgtc	cgggggcctg	480
gtcacttcgg	acaaggagag	cgctgctcac	atcctgagca	ggaaggaggg	tggcaatctg	540
ctggccatca	ttgtaggggg	tgcccaggag	gcgctgaacg	ccaggcctgg	agactccacg	600
cttctgctcc	ggaaccgcaa	gggcttcac	agacttgccc	tgatgcacgg	ggcagctctg	660
gtgccgatct	tttccctttg	ggagaatgag	ctatttgacc	aggttgagaa	ctctcctggc	720
tcctggttgc	gtaggattca	gaaccggctg	cagaagatca	tgggtatctc	cctcccgctc	780
ttccatggcc	gtggtgtctt	ccagtacagc	tttggcttta	taccctaccg	ccagcccatc	840
accaccgtgg	tgggaaagcc	catcgaggtg	cagaagacac	tgtgtccctc	aaaggaggag	900
gtggacaagc	tgcatcagcg	ctatgtcaag	gaactgtgtg	aactcttcga	aaccacaaag	960
ctcaagtaca	atgtccctgt	ggaccaacat	ttggagttct	gctga		1005

<210> 28
 <211> 1820
 <212> DNA
 <213> Bos taurus

<400> 28						
gaggagccag	catggtggag	tttgcacctt	tgtttgtgcc	actggagcgc	agactgcaga	60
cctttgcagt	cctgtactgg	atcttctgct	tcattgtctt	accccttta	tgcttctgtg	120
tcttcatagg	cctcctgttc	acaagattct	ggctcttcaa	tatcctgtat	gtgatctggt	180
tgtacttggg	ccaaaatagg	ccacggcagg	ggggcagtc	caataagttc	ctaaagcgct	240
gggtcctgtg	gaagtacatg	aaggactatt	tcccatcac	gctagtcaag	accactgagt	300
tggacccctc	ccggaactac	cttgtctgct	tccaccctca	tggactcctg	tcagtcggag	360
cttttaccaa	catgtgcaca	gacagcacgg	gcttctcctc	acttttcctt	ggcatccgcc	420
cacatctgac	aacgataaac	atatatttca	ggatcccact	cttcagagat	tacatcatgc	480
aaggggggct	ggtctcatca	gacaaggaga	gtattgttta	cattctgagc	aggaaaggag	540
gcggcaacct	agtggccatc	actgttgggg	gcatcagggg	ggcactgaaa	accaggcctg	600
gagccaacaa	gctggtgctg	cggaaccgca	agggttcat	caggctcgcc	ctgatgcatg	660
gagcagttct	ggtgccgatc	ttctcctttg	gggacaatga	attatatgct	aagacttctc	720
ctggcttttg	gtggaaatgg	ttcagagatc	aactgtacaa	gaagacacgt	cttgccatcc	780
cgttcttcta	tggccgtgga	gtcttccagt	acagttttgg	ttttatgccc	taccgccgac	840
ccatcaccac	cgtggtgggg	aagcccatcg	aggtgccgaa	gataccacat	ccctcccaag	900
aggaggtgga	caggctgcac	cagcattacc	tgaaggagct	gagcaacctc	ttcgaacccc	960
acaagctcaa	gtacaacatc	cccaaagacc	aacacttgga	gttctgctga	gtacctcccc	1020

pctau2012001598-seq1

agaggagggg	cttgcctcaa	agggcagggc	tggcattagg	aaggagggg	ctgtgatccc	1080
agaaggtttc	ctggaggtat	ctgttgaaca	catctccaga	gccttactg	actcctgcat	1140
ctccaccca	tgccaggccc	caggtcacat	ctggctcata	gaagaggaga	ccagactctg	1200
gcatcaatga	tttgctctcc	tccactaact	cctctgtaac	gggccctctc	tttggttctg	1260
cccacctgag	aaatgggaga	gatactggga	gaaaatgctg	tctgtgtgtc	cgtcctctgt	1320
ggtcctgtca	ctcaataagc	aggacaaaca	gcagctccag	tgagcagggt	tcccttcctc	1380
tgccctgagt	gacagctgga	agcagcccc	agctgcagcc	ctttctccta	caagctccct	1440
tatcctacat	cctgcctgtc	ccatcttctc	cctccctgag	ccggagagga	gggggagggc	1500
acaacctcct	ggctcatcct	ggcacctctg	agtttctgag	ccccatatgg	cagccagaag	1560
ggcctggagg	tggggaagac	ctgagctctg	tactggcta	aatgactgta	caccaacctc	1620
ttcaacctcc	ctctgagtct	cagttttccc	aaggactaaa	atgacctctg	acctcggtct	1680
taatcacaca	taggctgaaa	gctgcttatg	gtaactccta	ctctatcctt	tgtaattcag	1740
ccctgacctt	gggtccata	tatgtcaatt	catcacatga	aagattcgat	tttaaaaaaa	1800
aaaaaaaaaa	aaaaaaaaaa					1820

<210> 29
 <211> 1774
 <212> DNA
 <213> Rattus norvegicus

<400> 29						
gtgtgtggag	cctcctgagc	cagctcagca	tgggtggaatt	cgccccctg	ttggtaccat	60
gggagcgag	gctacagacc	ttcgcggtcc	ttcagtgggt	cttctccttc	ctggccttgg	120
cccagctctg	tatcttcac	ttcatcggtc	tcctattcac	gaggttcttg	ctcttctcag	180
tcctatatgc	cacctggtgg	tacctggact	gggacaggcc	gaggcaggga	ggccggccca	240
tccaattctt	cagacgcatg	gccatatgga	agtacatgaa	ggatttcttc	cctgtctctt	300
tggtaagac	agctgagctg	gacccttccc	ggaactacat	cgcaggcttt	cacccccatg	360
gagtcctggc	ggccggagcc	tttcttaacc	tgtgcactga	aagcacaggc	tttacctcgc	420
tcttcccagg	cattcgctcc	tatctgatga	tgctgactgt	gtgggtccgg	gcccccatct	480
tccgagatta	catcatgtct	ggggggctgg	tctcgtcaga	aaaggtgagt	gctgatcaca	540
ttctgagcag	aaagggcggc	gggaacttgc	tagccatcat	cgttgggggc	gcacaggaag	600
cactggatgc	caggcccgga	gcctacaggc	tgctgctgaa	gaatcgcaag	ggcttcatca	660
ggctcgcct	gacacatggg	gcagctcttg	tgccgatctt	ctcctttggg	gagaacaacc	720
tgtttaacca	agttgagaac	accctggca	cctggctgcg	ctggatccag	aactggctgc	780
agaagatcat	gggcatctcc	cttccgctct	tccatggcag	aggtgtcttc	cagtacagct	840
ttggcctcgt	gcccttccgc	cagcccatca	ccaccgtagt	ggggaagccc	atcgaggtgc	900
agatgatacc	acatccctcg	gaggaggagg	tgaaccggct	tcaccagctc	tatatcaagg	960
agctatgcaa	gctctttgag	gagcacaac	tcaagttcaa	cgtccctgag	gaccagcatc	1020

pctau2012001598-seq1

tggagttctg	ctaagtgtct	ccagccggaa	gacagctgca	tctgagcgcc	tgcaggagtg	1080
ctgggatgag	agggatttcc	acagccacca	ggaactccta	ggagcccagc	cacagctggc	1140
aagatggagg	cagttcctaa	ccctgggatt	cgtccctgca	gccaaagctc	tgagctctcc	1200
ctgtccttgg	tctgcacatc	tgtagaatgg	gggggaaagc	aggcagagag	caattcctga	1260
ggtctctctc	cacagttgta	atgtcattca	aacatgacca	aaggacagcc	atggagaaag	1320
agaacaaaat	tgttctttac	ctacccttga	gggacagtgc	aagaggagcc	agcaccccag	1380
gcctctctgt	acatgcccc	tggatgcctc	ttcttaccct	ctgagctgga	gaggggaagc	1440
catatatgcc	ctggcaccca	ctctatgttt	cttcagatt	ctgagtcctc	tgagctatga	1500
catgaaggag	cccagagaca	gttaaggagg	acagagctca	ctgactatga	ccgaggttag	1560
atctccttcc	catgtcctga	gcctcggttt	cctcagcctt	aatgacctgg	gagccacacc	1620
tctcacatca	cagaggctcc	cccaggagc	ctcttactca	tgctttctag	tgaactctag	1680
cctctgtctt	ggcacaaaag	ggcagtgtcg	tacatgttgc	ttcaataaat	aaaaggagtc	1740
tgtcgtaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaa			1774

<210> 30
 <211> 1056
 <212> DNA
 <213> Gallus gallus

<400> 30						
atggagcgtg	cctataaaat	ctgtgatgac	accacgtcca	aaagactgaa	aaacctcttc	60
tgggctgtgc	agaacgggtc	ccccagggtg	agccgttcct	ctcccaccac	cctcaccgct	120
caccgcctgg	cttttctccc	agcacagtgc	tgcaccgctg	ccttcacggg	cctcctcttc	180
acccgcttct	ggctgctcag	tgtcctctat	gccgcctggg	ggttcgtgga	caggggaagca	240
ccgctcaggg	gtgggcgacg	gatccacatg	gtgcggaaca	gcgcagtctg	gagacacatg	300
agggacttct	tccccgtcac	gctgggtgaa	acagcagagt	tggaccccag	gcagaactac	360
ttggtgggat	ttcaccccca	cggggctcctg	gcagtgggag	ccttcaccaa	ttttggaaca	420
gaggcgacag	gcttttccac	gatcttccca	ggcatcaccc	cacacctgat	gatgctctcc	480
ttatggtttc	gtgtcccctt	cttaagggtg	tacctgatga	gcggaggcct	tgtgtcctct	540
gataaggaga	gtgcgtacca	tgtgctgcag	aggccagagg	gggggaacct	gttggccatc	600
atcgttggtg	gggctcagga	ggctctggat	gcccgccag	gatcctgcac	cttgctgctg	660
aagaatagga	agggatttgt	ccgcgtggcc	atcgagcagg	gcacccact	ggtccttgcc	720
ttttcctttg	gggagaatga	gctcttcgac	caggtgagca	acccaaggg	ctcgtggctg	780
cgggtgatcc	agcaccggct	ccagcagatc	atgggcatct	ccctaccct	cttcacgca	840
agaggtatct	tccagtacag	cttcgggctg	gtgccctacc	gccggcctat	caacaccgtg	900
attgggaagc	caattcccg	gctgaagaag	cacaagccaa	cggaggaaga	ggttgatcga	960
gtccataaga	aatacttaga	agagctgagc	aagctctttg	aggagcacia	agctaaatac	1020

pctau2012001598-seq1
aacgtcccag aagacagtca cctcgaattt atatag 1056

<210> 31
<211> 1287
<212> DNA
<213> Danio rerio

<400> 31
gtgttttggtt cagaatgacc cagtaaactt gttcatttgt ggagcctgtg ggaataagta 60
aacaaatcat acctggatct ttgagttctc cagagctgga ttttactcgc cctgctgtat 120
tgtcctgttc ttcacctcgc tcttcacccg tttctggctc atcagtgtgc tttatgccac 180
ctggtggtac atcgactggg acacaccttc acgaggaggg agaatagtgc catatctctg 240
caaaactgcgt gtgtgggagt acatgagaga ttatttcccc atcaggctga taaagacggc 300
tgacttggac actaggaaga actatgtgat gggttttcat cctcatggta ttctcgtggc 360
tggtgctttt acaaatttct gcacggaggc gactggattt agcaaactct tccctgggat 420
aaagagcaac cttcttatgc tgcctctctg gttcagagcc cccttcttca gagattacat 480
catgtcagca gggctgggtc catcagacaa ggagagtgcc agttatctgc tccggaggaa 540
gggaggaggg aatgcggtgg tcattgcagt cggaggagct cctgaagctt tggatgcaca 600
tcctggtgac tacaccgttc acctggccaa taagaaaggc tttattaagc ttgcaataga 660
gcacggtgca gatttgggtc ctatatattc gtttggagag aacgagggtg ttgaccaggt 720
gcagaatcct cgtggcacat ggctgcgata cattcaagag cgtctgcagc gcatcatggg 780
tgtttctctt cctctctttc atgcgcgagg agttttccag tacaccttcg gcctgatgcc 840
ttacagaaaag cccatcaaca ctgttgtggg aaggccaatt ccagtggaga aaaatgagaa 900
accaacagca gaagaattgg acgtctatca tcagcgctat atggatgaac tggcacggct 960
gttcgaggac cacaaggga attatggagt acctgaagac actcatctgg tgttccagtg 1020
agatgatcac ttatgttgac gatgagtttg tgtgtttgta ccttgtatctt cttacattgt 1080
aggaaccaa tgtcttcaca tatatagaca taccactaca tttgaccttg tggggacatt 1140
tgtttggttc taatgttcaa aatggctcat gaatcatgct gaatttaata tttagaaaat 1200
gtaaaaatgc cgattgtttt ctgtgagggt tgggttttagg ggtagaaaaa aaaccaaaaa 1260
aaaaaaaaa aaaaaaaaaa aaaaaaa 1287

<210> 32
<211> 1268
<212> DNA
<213> Drosophila melanogaster

<400> 32
ccgagagccg ataaggataa gtataagtct agtacagggc aatttgacca tcgctggctt 60
ccatatcatt cagtcgtggg agaaccgccg aaagagcagc tgggtggagt tttttgggca 120
tttttgtaac acagcagaca tgaaaatcga gtgggcacca ctgcgggttc ctctggaacg 180
ccgactgcag atactggtca cggccttttt cacctccatg ctgctgatac tattgtcagt 240

pctau2012001598-seq1

ttccttcctt ttggtagctg gatcactgat ctacggaggt cttttggtgc gtagtctgat	300
ggtaacttac ttggcctacg tctttgtgca ccacaagaaa acccaatccg ttgtggatgg	360
caatggctgg atgataacac gcaccaacct tttgcatcgc cactatcgtg attactttcc	420
cgtggagctg gtgaaaacag ccgaactgcc agctactaag aactacatct tggccagctt	480
tccccacgga attctgggca caggcattgg cattaacatg ggcttggaat tctccaagtg	540
gctggagcta ttcccccaag tgcgtcccaa actgggcact ctggatcagc atttccatgt	600
tccgttcatt cgtgaggtcc tccgctgctg gggctctggtg tcagtgtcca aagaggcgct	660
gatccgtatg ctacgaagt caaatgatcc caagcacaag gataatcggg atggtttcac	720
ctccaatgcg gtggccattc tggttggcgg tgcccaggaa gccatggact ctcatcctgg	780
gcagtacatt ttaaccttga agaataaggaa aggcttcgtg cgaatggcca ttagaacggg	840
ctcatcgtat gttccttcatt tttcctttgg agagggtggac attttcgatc aggtggcaaa	900
tcccccaac tcgctgtctc gacggtttca ggactttgtc aagaagctca ccggagtctc	960
tccgctgatt cctgtgggcc gcggattctt caactacacc tttggcttcc tcccattccg	1020
acgacgcatt gtccaagttg ttggtgctcc catcgatgtt gttaagaacg agcaccacga	1080
ctcggagtat gtggataaag tgcatggaca ggtcattgag tcgctggaga agttattcga	1140
tcagtacaaa gacaagtact tggagaattc gaagagtgcc actctagttg tacactagtt	1200
ttaaatacatg ttgatacaaa atattcctat gtattgaaat agaaacgagc tttttaattg	1260
aatttaaa	1268

<210> 33
 <211> 1269
 <212> DNA
 <213> *Drosophila melanogaster*

<400> 33	
agttaaatgt tgtacgctcg catcgctcgg tctgctgctc aaacgcgata acgatccatt	60
actcgccgga ggaccgtgct accaagtagt tcaactttcc gagttcgaga gctctgacat	120
catgaaaatc gagtgggctc ccaaaggagt tcccatggaa cggcgctcgcc agacgtttgc	180
catggccttc ctaatcctat cttttatgat actctccttt ggatcctatt tcttcgttgc	240
tgccgtgctg ttctatggaa gcctttttgtg gcgcaccatt atggtcatct atttgggtcta	300
cgtctacgcg aatcacaaga gaaccactc cattatggat ggcaatggct ggaaaattaa	360
ccgcaacaat tggctgtttc ggcatatcgc ggactacttt cccgtacagc tggtaagac	420
cgcagagctg ccgccgaata agaactacat tttggccagt tttccgcacg gaattttggg	480
gactggcatt tccattaata tgggtctaga catctccaag tggctgcaac tgttcccaca	540
ggtgaggccc aagggtggcca ctttggatca aaacttcctg acgcccacgc tgcgtggtct	600
cctgagatct tggggcctgg tgtcggtttc caaggaggcc ttagtatatc tgctaaccaa	660
atcaaacgat cccaagcaca aggacaatcg tgatgggttc acctccaacg cgggtggctat	720
tctggtgggt ggtgcccagg aggcactaga ctgcgacctt ggcaaatata tattgacact	780

pctau2012001598-seq1

gaagaaccgc aagggattcg ttaaaatggc cattaggaca gggctcttcga ttgttccac	840
gttttccttt ggcgagggtg acattctgga tcagggttgcc aatccaccga actctcgggt	900
tcgtcgcttt caagactttg taaagaggat aacggggata tctccgctga ttcccgttgg	960
ccggggcatc ttcaactact cctttggctt tctgccaac cgtcgacgca ttgtccaagt	1020
tggtggcgct cccatcgacg tggttcagag cgatcaacca gacgcagcct atgtggataa	1080
gatacacaaa caggttattg acgacctgga gaagatgttc gccaaagtaca aggatcagta	1140
cataccgaac tccaagcagg acaagctgat tatacactag tcgagaagtt attgtctaaa	1200
caatgtgatc aaatgagtgc ttataagca tttaatataa aaattacgaa ttttatcagg	1260
caaataaaa	1269

<210> 34
 <211> 1139
 <212> DNA
 <213> *Drosophila melanogaster*

<400> 34	
atgacaatcg aatgggctcc tctccgggtt ccgctggaac ggcggttca gaccctggtt	60
acgtcattct tcacctatac cttctttacg ctgcccattc cgtcctgcct cgcagtagct	120
atcctgctgt actacggcga aatgtttgtg cgtagtctgc ttcttatcta ttttgtgaaa	180
atttatttgg attataaaag aaattacggc attatggaag gtaatgggtg gttattctac	240
cgcagcaatt ggagatatcg aaactacttc cccgtcgaac ttgtaaaaac cgccgagctg	300
cctccaaca ggaactatat cgtagcaagc tttccacacg gaatacttgg tactggaact	360
tgcattaaca tgagcctgga catcgacaat tggctatccc ttaccgcga cgttcgacct	420
aagatcgcca ccctggatca tcacttcaag acgcctttcc tgcgtgacat attgcgctgg	480
tggggcatgg tatcggtttc caaggaatcc ctgtcctatt tactcagcaa gtcaaagat	540
ccaatgcaca aggataatcg ggatggtttc acatccaatg cgggtggcgt acttggttgg	600
ggcgccaagg aagccatgga ttcgcatccg ggacagtaca ttttgaccct caaggataga	660
aagggtttcg tcaaaatggc tggtcgaacc ggatcgtcga ttgtgccctc gttgtccttt	720
ggcgagggtg acatatttga tcagggtggt aatcctccgg actcctcgct ccgtcgcttc	780
cagaatgtgg tcaagaaatt cacaggcatt tcaccgctcc ttccaaggg tcgcggtatc	840
ttcaactaca actatggcat tctcccacat cgacgacgta ttgtccaagt tggtggttcc	900
cccatcgatg tggaaagatg cgagactccc gatcccgaat atgtggacaa gatccatggg	960
caggtgatcg acgcgctggc gaggatgttc gatgaataca aagagaagta cactccaaac	1020
tcaaagcata tcaaacttat aatacagtga ctcagtagag tcccttagat tgtgattttt	1080
tatagaatta aggaaacttt tgaaacgcaa cacagtttta gactaattta aaaaaaaaa	1139

<210> 35
 <211> 1971
 <212> DNA

<213> Anopheles gambiae

<400> 35

cgcatgagga gaaaagttgg ccactcagtt attatacgcg acaagctccc agtataagcg	60
tcattgttag atcgcattct acaatgacaa acggacccga catcattagt catgtttggc	120
cgggtggtcag aagaaagcac agagggtgga agaaaagctg attcattggg tgaggacggt	180
agtctgcggg cgtgttttta tcgtctcagt ttgcatcatg ccccagtga agtgataccg	240
acacgacatt tgggtgaagcc attcgggttca agattaaggg acaggctaaa ccagtacaag	300
tgagtattcg ttgattcatt cgtaaaccga acgttgtttg ctcagcaaga agcaatgact	360
gataagtgat aacgtaattg catcatcggc atcaatgctg ataagatctc gagctgcggg	420
aagggaaaacg catcaagtgg tacatcgttt gtgggtgtgat agattttggg agggcgagtc	480
gtagtggtgg aacatcccc ggtaccggtt ctagacgcgc tcaccaaacc ctaagatggg	540
taaaatcgaa tgggccccac tgaatgtgcc gatgaggcgg cggatcgaaa ccctgtccac	600
tgccctgtgg atgtggctga tactgttcgg cgagctgggc atgctgatat cgtacttcct	660
gctgctgatc tatggtaact tgtttattaa atcgctgtgc gtcactctac gatacttcat	720
ctacactgat cgaaaagtga ctttgaacgg gggccgtgga caggggtgtaa catggtggcg	780
agagctgttc tgggtggaagc tctaccagag ctacttcccg gcgaagctgc acaaaacggt	840
tgatctcgac cctaaccgga actacctgtt tgccgccttc ccgcacggcg tgcttggcct	900
cggtgccctc atcaactttg ccaccaatgc gaccggcttt caccgacaagt ttcccaaat	960
ccgttccccg ccggtgacgc taaacttcca ctttgtgata ccgttcttcc gcgagctgct	1020
gctcagctgg ggctcgtgt cggccaacc gaacagcatc ctgagcctgc tgaaggcgcc	1080
gaacaaaccg gaccaccgc tgaacgacga cggctacac gccaacgcgg tcgtgatagt	1140
ggtgggcggg gcggccgaat cgctgcactg ccggccgaac aactacacgc tgggtgctgcg	1200
caagcgggaag ggcttctgca agctggcgat caaggcgggc acccgctgg tgccggtgat	1260
gacgttcggc gaggtcgacc tgttcgatca gccgccgaat ccgcccggct cgcggttgcg	1320
ccggttccag gagtttgtga aaaacacgac cggcatcgcg ccggccgcct tcgtcgggcg	1380
gggcttcttc cagtacagct acgggttgat cccgcgccgg aagccactga acacagtcgt	1440
cggtgctccg gtggaggatga cgaaataga caaccgacg caggaaacagg tggatgaggt	1500
gcacgagcgg ttctgccgcg cgctggacaa tctgttcgaa acgaacaaat cgcgcttcat	1560
tgccgactac aagaatgtga agcttgtcat ggagtaatat gtaaaatgca gatacttccc	1620
gtacttcccc ccatctttac acacaaacga tcggtgtata gcggtgggaa aacggggcaa	1680
aatcacgcac aagagacaca gaaacacact atcatcacga taccctacgc aaccggtgga	1740
aattgtgtta taattcgcat cgtaaggatt gagactgtac caaaaaaaaa aaaaaaacga	1800
acgatcaaaa ccaactcttt tcatacctaa gatccaacca ttgtgcggtg ccaacaggtg	1860
atattgttat acagctctgc gagaggagat caattactag ctgcgtaaga gaggacacac	1920
tacattggga gtgttttagca tttaaattct cgtgtaagtg taattgtgta a	1971

pctau2012001598-seq1

<210> 36
 <211> 1612
 <212> DNA
 <213> *Anopheles gambiae*

<400> 36
 gtcacattct aaccttacgc ccgtgctgaa cccagggga gaccagtgcg cgctccactg 60
 tgaactccct agtgtgtagt gtattctttt cctgttggtt tagatgtatt aagggcgagt 120
 cgtagtgggtg gaacatcccc cgggtaccggt tctagacgcg ctcacaaaaa cctaagatgg 180
 gtaaaatcga atggggcccca ctgaatgtgc cgatgaggcg gcggatcgaa accctgtcca 240
 ctgccctgtg gatgtggctg atactgttcg gcgagctggg catgtctgata tcgtacttcc 300
 tgctgctgat ctatggtaac ttgtttatta aatcgctgtg cgtcatctac ggatacttca 360
 tctacactga tcgaaaagtg actttgaacg ggggccgtgg acaggggtgta acatgggtggc 420
 gagagctgtt ctgggtggaag ctctaccaga gctacttccc ggccaagctg caaaaaacgg 480
 ttgatctcga ccctaaccgg aactacctgt ttgccgcctt cccgcacggc gtgcttggcc 540
 tcggtgcctt catcaacttt gccaccaatg cgaccggctt tcacgacaag tttcccaaaa 600
 tccgttcccc cccggtgacg ctaaacttcc actttgtgat accgttcttc cgcgagctgc 660
 tgctcagctg gggcctcgtg tcggccaacc cgaacagcat cctgagcctg ctgaaggcgc 720
 cgaacaaacc ggaccaccg ctgaacgacg acggctacac ggccaacgcg gtcgtgatag 780
 tgggtggcg tgcggccgaa tcgctgcact gccggccgaa caactacacg ctggtgctgc 840
 gcaagcgga gggcttctgc aagctggcga tcaaggcggg cccccgctg gtgccggtga 900
 tgacgttcgg cgaggtcgac ctgttcgatc agccgccgaa tccgcccggc tcgcggttgc 960
 gccggttcca ggagtttgtg aaaaacacga ccggcatcgc gccggccgcc ttcgtcgggc 1020
 ggggcttctt ccagtacagc tacgggttga tccgcgccg gaagccactg aacacagtcg 1080
 tcggtgctcc ggtggaggtg acgcaaatac acaaccgcg gcaggaacag gtggatgagg 1140
 tgcacgagcg gttctgccgc gcgctggaca atctgttcga aacgaacaaa tcgcgcttca 1200
 ttgccgacta caagaatgtg aagcttgtca tggagtaata ggtaaaatgc agatacttcc 1260
 cgtacttccc gccatcttta cacacaaacg atcgggtgat agcgggtggga aaacggggca 1320
 aaatcacgca caagagacac agaaacacac tatcatcacg atatcctacg caaccctggg 1380
 aaattgtgtt ataattcgca tcgtaaggat tgagactgta ccaaaaaaaaa aaaaaaacg 1440
 aacgatcaaa accaactctt ttcataccta agatccaacc attgtgcggt gccaacaggt 1500
 gatattgtta tacagctctg cgagagggag tcaattacta gctgcgtaag agaggacaca 1560
 ctacattggg agtgtttagc atttaaattc tcgtgtaagt gtaattgtgt aa 1612

<210> 37
 <211> 1017
 <212> DNA
 <213> *Tribolium castaneum*

pctau2012001598-seql

<400> 37
atgaagattt tggggatcaa gttcgcccct ttgcacatcc cctcgcgagcg ccgcttgcaa 60
acccttgccg ccggatgctg gttcacaacc ctggcctttg gtaccttcat cggtagcttc 120
atctgggtgt acgtcttctt cacccgattc tggtagctgt ccgtccttta cgccaccata 180
atatacctcg agaaaagcaa gtgtgagaag ggcggccgcc ccatcgaatg gatcagacat 240
tggggatggt ggtactactt gaaaaactac ttcccctgta agcttgactt tgtcccgggc 300
ctaacctttg atccgaagag gaattatcta tttgcctgtt acccccatgg gatcctcccg 360
gcggggccctt tcaacaccat tggtagccct tacagcgaat ttagcaagtt gtttccaaag 420
tttcgggtga ggttggtgat tctgcaccaa catttcttca ttccgtttct gcgggagatt 480
gcttatggga caggggggat ttcggcctcg gccaaagtcgt tgaatcatgt tttggggagt 540
ccggaagggg ggtatattgc ggtgttgatg cccgggggag cagttgaggc ctataacagc 600
cggcctggac agtacagaat catcctgaag aacaggaaag gcttcgtaaa actggccttg 660
aggaatgggt caccactagt accagttatc tcgttcgggg agccggaatt gtttgatcaa 720
gttgaaggga ggactttgag gaagattcag gagtccatca ggaagtacct cgggctggca 780
cctgtgattt tctcaggag agggttcttc cagtactcgt ttggcgttat tccacaaagg 840
aggcctatta cgaccgttgt cgggcacccg attgaagtta ccaaaatcga aaaaccaaca 900
aacgaggaag tggacgaact acataaaaag tacatgcaag agctggagaa cttgttcgaa 960
gagtacaaat tcaaataattt ggagaatccg aaggacatac acttggagtt tgaataa 1017

<210> 38
<211> 1420
<212> DNA
<213> Homo sapiens

<400> 38
cactcacaca cctacggaca cacgctactc tgggaggtga tttgcgactt agccaggccc 60
ccaaagctgg gtcctgtag ggagaaagtc tgcccaggtc cacatccaag ctttcatcgt 120
ttgtcctccg ggttctggga tcctgctgga agaggggagc ttctgcaatg ggagttgcca 180
caaccctgca gcccacaacc acttccaaaa cttgcagaa gcagcatcta gaagcagtgg 240
gcgcctacca atatgtgctc actttcctct tcatggggcc tttcttctcc cttcttgtct 300
ttgtcctcct cttcacgtca ctctggccct tctctgtttt ttacttgggtg tggctctatg 360
tggactggga cacaccaac caaggtggaa ggcgttcgga gtggataagg aaccgggcaa 420
tttgagaca actaagggat tattatcctg tcaagctggt gaaaacagca gagctgcccc 480
cggatcggaa ctacgtgctg ggcgccacc ctcattggat catgtgtaca ggcttctctt 540
gtaatttctc caccgagagc aatggcttct cccagctctt cccggggctc cggccctggt 600
tagccgtgct ggctggcctc ttctacctcc cggcttatcg cgactacatc atgtcctttg 660
gactctgtcc ggtgagccgc cagagcctgg acttcatcct gtcccagccc cagctcgggc 720
aggccgtggt catcatggtg ggggggtgag acgaggccct gtattcagtc cccggggagc 780

pctau2012001598-seq1

actgccttac gctccagaag cgcaaaggct tcgtgcgccct ggcgctgagg cacggggcgt	840
ccctggtgcc cgtgtactcc tttggggaga atgacatctt tagacttaag gcttttgcca	900
caggctcctg gcagcattgg tgccagctca ctttcaagaa gtcctatgggc ttctctcctt	960
gcctcttctg gggtcgcggg ctcttctcag ccacctcctg gggcctgctg ccctttgctg	1020
tgcccatcac cactgtggtg ggccgccccca tccccgtccc ccagcgcctc caccaccacg	1080
aggaggaagt caatcactat cacgccctct acatgacggc cctggagcag ctcttcgagg	1140
agcacaagga aagctgtggg gtccccgctt ccacctgcct caccttcctc taggcctggc	1200
cgcggccttt cgctgagccc ctgagcccaa ggcactgaga cctccacca ctgtggactc	1260
catgcctcca ataaaaggta gttctggggc cagcgcagtg cctcgtgcct gtgatccag	1320
cactttggga ggccaggggtg ggaggatcgt ttgagcccag gagttgaaga ccagcctggg	1380
caacacagtg agacttcatt tctacaaaaa aaaaaaaaaa	1420

<210> 39
 <211> 1934
 <212> DNA
 <213> Pan troglodytes

<400> 39

ccactcacac acctacggac acacgtact ctgggaggtg atttgcgact tagccaggcc	60
cccaaagctg ggctcctgta gggagaaagt ctgcccaggt ccacatccaa gccttcctcg	120
tttgtcctct gggttctggg atcttgctgg aagaggggag cttctgcaat gggagttgcc	180
acaacctgc tgccccaac cacttccaaa accctgcaga agcagcatct agaagcagtg	240
ggcgcctacc aatatgtgct cactttcctc ttcattgggc ctttcttctc ctttcttgct	300
tttgtcctcc tcttcacgtc actctggccc ttctctgttt ttacttggt gtggctctat	360
gtggactggg acacacccaa ccaaggtgga aggcgttcgg agtggataag gaactgggca	420
atttgagac aactaaggga ttattatcct gtcaagctgg tgaacacggc agagctgccc	480
ccggatcgga actacgtgct gggcgccac cctcatggga tcatgtgtac aggccttctc	540
tgtaatctt ccaccgagag caatggcttc tcccagctct tcccggggct ccggccctgg	600
ttagccgtgc tggctggcct cttctacctc ccggtctatc gcgactacat catgtccttt	660
ggactctgtc ccgtgagccg ccagagcctg gacttcctcc tgtcccagcc ccagctcggg	720
caggccgtgg tcatcatggt tgggggtgcg catgaggccc tgtattcagt ccccgaggag	780
cactgcctaa cgctccagaa gcgcaaaggc ttcgtgcgcc tggcgctgag gcacgggtggg	840
ccgccccatc cctgtcccc agcgcctcca cccaccgag gaggaagtca atcactatca	900
cgccctctac atgatggccc tgcagcagct cttcgaggag cacaaggaaa gctgtggggg	960
ccccgcttcc acctgcctca cttctatcta ggcttggccg cggcctttcg ctgagccct	1020
gagcccaagg cactgagacc tccaccact gtggactcca tgcctccaat aaaaggtagt	1080
tctgggcca gcgcagtgc tcacgcctgt aatcccagca ctttgggagg ccaaggtggg	1140
aggatcgttt gagcccagga gttgaagacc agcctgggca acatagttag acttcatttc	1200

pctau2012001598-seq1

tacaaaaatt	aaaaaataat	gttggttaatt	agccgggcat	ggtggcatgt	gcctgtaatc	1260
ccagctat	gggaggctga	ggcaggagac	ttgcttgaac	ccaggaggcg	gagggttcag	1320
tgagctgaga	tcacaccaca	gtactccggc	ctggggcaaca	cagcgagact	caatctcaaa	1380
aaaaaaaaa	aaaaaaaaa	aagatggtta	aagctatagt	aggctgggcg	tggtggctca	1440
tgcccataat	cccaacactt	tgggaagcca	aggcagggtg	atcacctgag	gtcaggagtt	1500
caagaccagc	ctggccaaaa	tggcaaaacc	ccctctctac	taaaactaca	aaaattagcc	1560
gggcatggtg	gctcatgcct	ataatcccag	gtactccgga	ggctgaggca	ggagaatcgc	1620
ttgaaccag	gaggcagagg	ttgcagttag	ccgagatcac	accactgcac	tccagtctgg	1680
gcaacagagc	gagattccat	ctcaaaaaaa	aaaaaaaaa	gttggttaaag	ctataatcat	1740
tctggaagt	aacatggcca	ggcatgacct	gcccccatcc	tgatgtgcta	caaagctgct	1800
gccttagatt	gggtatcagg	accctgaact	gtcctcccag	tgagagttca	gtctgtggca	1860
tgaccttga	caagcctctt	ccttttgcct	gtttacttct	ttgtctgtaa	ataaaggcat	1920
ctgcttagca	aaga					1934

<210> 40
 <211> 2223
 <212> DNA
 <213> Pan troglodytes

<400> 40	
ccactcacac	acctacggac acacgtact ctgggaggtg atttgcgact tagccaggcc 60
cccaaagctg	ggctcctgta gggagaaagt ctgcccaggt ccacatccaa gccttcacgc 120
tttgtcctct	gggttcttggg atcttgctgg aagaggggag cttctgcaat gggagttgcc 180
acaaccctgc	tgcccccaac cacttccaaa accctgcaga agcagcatct agaagcagtg 240
ggcgctacc	aatatgtgct cactttcttc ttcatgggcc ctttcttctc ctttcttgct 300
tttgtcctcc	tcttcacgtc actctggccc ttctctgttt ttactttggg gtggctctat 360
gtggactggg	acacacccaa ccaaggtgga aggcgttcgg agtggataag gaactgggca 420
atttgagagc	aactaaggga ttattatcct gtcaagctgg tgaacacggc agagctgccc 480
ccggatcggg	actacgtgct gggcgccac cctcatggga tcatgtgtac aggccttctc 540
tgtaatttct	ccaccgagag caatggcttc tcccagctct tcccggggct ccggccctgg 600
ttagccgtgc	tggctggcct cttctacctc ccggtctatc gcgactacat catgtccttt 660
ggactctgtc	ccgtgagccg ccagagcctg gacttcaccc tgtcccagcc ccagctcggg 720
caggccgtgg	tcatcatggg tgggggtgcg catgaggccc tgtattcagt ccccggggag 780
cactgcctaa	cgctccagaa gcgcaaaggc ttcgtgcgcc tggcgctgag gcacggggcg 840
tccctgggtg	ccgtgtactc ctttggggag aatgacatct ttagacttaa ggcttttgcc 900
acaggctcct	ggcagcattg gtgccagctc accttcaaga agttcatggg cttctctcct 960
tgcatcttct	ggggctgctg tctcttctca gccacctcct ggggcctgct gccctttgct 1020

pctau2012001598-seq1

gtgcccata	ccactgtggg	tgagtgccca	cctccggggg	gacggccacc	agcagctgcg	1080
tgggcatcag	ggatccctcg	cccacctgtc	tccctctccc	tgcaagtggc	cgccccatcc	1140
ctgtcccca	gcgcctccac	cccaccgagg	aggaagtcaa	tcactatcac	gccctctaca	1200
tgatggccct	gcagcagctc	ttcgaggagc	acaaggaaag	ctgtgggggc	cccgtttcca	1260
cctgcctcac	cttcatctag	gcttggccgc	ggcctttcgc	tgagcccctg	agcccaaggc	1320
actgagacct	ccaccactg	tggactccat	gcctccaata	aaaggtagtt	ctggggccag	1380
cgcagtgcct	cacgcctgta	atcccagcac	tttgggaggc	caagggtgga	ggatcgtttg	1440
agcccaggag	ttgaagacca	gcctgggcaa	catagtgaga	cttcatttct	acaaaaatta	1500
aaaaataatg	ttgttaatta	gccgggcatg	gtggcatgtg	cctgtaatcc	cagctatttg	1560
ggaggctgag	gcaggagact	tgcttgaacc	caggaggcgg	aggttgcagt	gagctgagat	1620
cacaccacag	tactccggcc	tgggcaacac	agcgagactc	aatctcaaaa	aaaaaaaaaa	1680
aaaaaaaaaa	agatggttaa	agctatagta	ggctgggcgt	ggtggctcat	gcccataatc	1740
ccaacacttt	gggaagccaa	ggcaggtgga	tcacctgagg	tcaggagttc	aagaccagcc	1800
tggccaaaat	ggcaaaaccc	cctctctact	aaaactacaa	aaattagccg	ggcatggttg	1860
ctcatgccta	taatcccagg	tactccggag	gctgaggcag	gagaatcgct	tgaaccagg	1920
aggcagaggt	tgcagtgagc	cgagatcaca	ccactgcact	ccagtctggg	caacagagcg	1980
agattccatc	tcaaaaaaaaa	aaaaaaaaatg	ttgttaaagc	tataatcatt	ctggaagtga	2040
acatggccag	gcatgacctg	cccccatcct	gatgtgctac	aaagctgctg	ccttagattg	2100
ggtatcagga	ccctgaactg	tcctcccagt	gagagttcag	tctgtggcat	gaccttgac	2160
aagcctcttc	cttttgctg	tttacttctt	tgtctgtaaa	taaaggcatc	tgcttagcaa	2220
aga						2223

<210> 41
 <211> 2137
 <212> DNA
 <213> Pan troglodytes

<400> 41	
ccactcacac	acctacggac acacgtact ctgggaggtg atttgcgact tagccaggcc 60
cccaaagctg	ggctcctgta gggagaaagt ctgccaggt ccacatccaa gccttcatcg 120
tttgtcctct	gggttctggg atcttgctgg aagaggggag cttctgcaat gggagttgcc 180
acaaccctgc	tgccccaac cacttccaaa accctgcaga agcagcatct agaagcagtg 240
ggcgcctacc	aatatgtgct cactttcctc ttcattgggc ctttcttctc ctttcttgct 300
tttgtcctcc	tcttcacgtc actctggccc ttctctgttt ttacttggt gtggctctat 360
gtggactggg	acacacccaa ccaagggtgga aggcgttcgg agtggataag gaactgggca 420
atttggagac	aactaaggga ttattatcct gtcaagctgg tgaaaacggc agagctgccc 480
ccggatcgga	actacgtgct gggcgccac cctcatggga tcatgtgtac aggcttcctc 540
tgtaatttct	ccaccgagag caatggcttc tcccagctct tcccggggct ccggccctgg 600

pctau2012001598-seq1

ttagccgtgc	tggctggcct	cttctacctc	ccggtctatc	gcgactacat	catgtccttt	660
ggactctgtc	ccgtgagccg	ccagagcctg	gacttcatcc	tgtcccagcc	ccagctcggg	720
caggccgtgg	tcatcatggt	tgggggtgcg	catgaggccc	tgtattcagt	ccccggggag	780
cactgcctaa	cgctccagaa	gcgcaaaggc	ttcgtgcgcc	tggcgctgag	gcacggggcg	840
tccctggtgc	ccgtgtactc	ctttggggag	aatgacatct	ttagacttaa	ggcttttgcc	900
acaggctcct	ggcagcattg	gtgccagctc	accttcaaga	agttcatggg	cttctctcct	960
tgcactttct	ggggtcgtgg	tctctttctc	gccacctcct	ggggcctgct	gccctttgct	1020
gtgcccata	ccactgtggt	gggccgcccc	atccctgtcc	cccagcgcct	ccaccccacc	1080
gaggaggaag	tcaatcacta	tcacgccctc	tacatgatgg	ccctgcagca	gctcttcgag	1140
gagcacaagg	aaagctgtgg	ggtccccgct	tccacctgcc	tcaccttcat	ctaggcttgg	1200
ccgcggcctt	tcgctgagcc	cctgagccca	aggcactgag	acctccaccc	actgtggact	1260
ccatgcctcc	aataaaaggt	agttctgggc	ccagcgcagt	gcctcacgcc	tgtaatccca	1320
gcactttggg	aggccaaggt	gggaggatcg	tttgagccca	ggagttgaag	accagcctgg	1380
gcaacatagt	gagacttcat	ttctacaaaa	attaaaaaat	aatgttggtta	attagccggg	1440
catggtggca	tgtgcctgta	atcccagcta	tttgggaggc	tgaggcagga	gacttgcttg	1500
aaccaggag	gcggagggtg	cagtgcgctg	agatcacacc	acagtactcc	ggcctgggca	1560
acacagcgag	actcaatctc	aaaaaaaaaa	aaaaaaaaaa	aaaaagatgg	ttaaagctat	1620
agtaggctgg	gcgtggtggc	tcatgcccac	aatcccaaca	ctttgggaag	ccaaggcagg	1680
tggatcacct	gaggtcagga	gttcaagacc	agcctggcca	aaatggcaaa	acccctcttc	1740
tactaaaact	acaaaaatta	gccgggcatg	gtggctcatg	cctataatcc	caggtactcc	1800
ggaggctgag	gcaggagaat	cgcttgaacc	caggaggcag	aggttgagct	gagccgagat	1860
cacaccactg	cactccagtc	tgggcaacag	agcgagattc	catctcaaaa	aaaaaaaaaa	1920
aatgttggtta	aagctataat	cattctggaa	gtgaacatgg	ccaggcatga	cctgccccca	1980
tcctgatgtg	ctacaaagct	gctgccttag	attgggtatc	aggaccctga	actgtcctcc	2040
cagtgcagag	tcagtctgtg	gcatgacctt	ggacaagcct	cttccttttg	cctgtttact	2100
tctttgtctg	taaataaagg	catctgctta	gcaaaga			2137

<210> 42
 <211> 1749
 <212> DNA
 <213> Canis familiaris

<400> 42	
atgacgttcc	ctgaccccaa gacgggacag tgctgcctct ctccggtcct gtccgggcct 60
ggggggcgcc	gggtggccct gctggccctg ggagggcccc tggaagcgct ggaagcaaaa 120
ccgggcgcgc	tgttttgcgg atccggaatc agaagagatt tgtccagctg gcaccgggac 180
acggagcgag	ctcttcagc cgttcgcgaa ccctccggac ccttgggcct ccttctgccc 240

pctau2012001598-seq1

ttccggggcgc	gcccattccac	ccgggtccggg	agtcccagcc	tctgccccct	cggtctcccct	300
gcgcgcacca	gctccgcagc	gctcggggcca	gcgcgctccc	cgcgccgggg	cggaaccccc	360
gcgctgctgc	ccacgcctgg	gcacaccggc	tcacctcttt	tgacactccc	agtccttgac	420
tctgtcgaga	cggtcgcgcg	tggcaatgaa	gaggccgggg	ggtgccgggg	atcccgttct	480
gcgcgcatgg	cgcggtctac	gcggctgttc	cgggagcgca	aggcgcgcg	cggcctccc	540
ggccaccgc	acctccccgc	cgtgtccctg	aggtgccgac	actcccgggg	acctgcacac	600
ggagccgggg	cgcgaccag	gggcctgccc	ggccacgctg	ggccccggag	gcgtggagcg	660
tttggtcgct	ccagcaagaa	agccttctact	ccgtgcccaa	accttcgctg	tttgcgctcc	720
aggccctgga	acccactgg	cggaggggag	cgccccacca	tgagaaccac	gcagaaacag	780
tggctagaag	ccttaagcgt	ctcctattac	gtgttctactt	tcctcttcat	gggccttttc	840
ttctcccttc	ttgtcctctt	cctcctcttc	acttcattct	ggtccatctc	tgttctctac	900
ttggtgtggc	tcttcctgga	ctgggacaca	cccaaccaag	gtgggaggtg	ctttgagtgc	960
aacaggagct	gtaccatttg	gaaacacttg	aaggattatt	ttcctattaa	gctggtgaag	1020
acagcagagc	ttccccctga	ccggaactat	gttgtgggtt	gtcaccacac	tgggatcatg	1080
tgcatgggaa	ccttctgcaa	tttcttcacg	gaggccaaca	acttctccaa	gcagttccc	1140
gggattcaga	catcaccctg	aacactggcc	tttctcttgc	acctcccagt	ctatcgtgac	1200
tacctcatgt	accttggact	gtgttctgtg	aaccgccgga	gcctggactt	tattttgtcc	1260
aagccccagc	ccgggcaggc	ggtagccatc	gtagttgggg	gtgcccacga	gtccctgttt	1320
gcgatcccag	ggatgactg	cctcgttctc	cggaaccgca	aaggcttcgt	gcgcctggca	1380
ctgaggcacg	gagcttccct	ggtacctgtg	tactcctttg	gggagaatga	tatcttcaat	1440
tttaaggctt	ttcccacaaa	ctcctggcag	tatctgtgcc	agatcaccat	caagaagatc	1500
atgaaatttt	ctccttgcat	cttctggggc	cgcggactct	tctcagctga	ctcctgggg	1560
ttactgccct	tcgccaaagg	tatcaccact	gtggtggggc	gccccatccc	tgtgccacag	1620
cgcctcaacc	ccactgagga	ggaagttgac	cactatcaca	tgctctacat	ggaggctcta	1680
gagcaactgt	ttgaggaaca	caaggaaagc	tgtggtgtcc	cagcctctac	tcacctcatc	1740
tttaaataag						1749

<210> 43
 <211> 1221
 <212> DNA
 <213> Bos taurus

<400> 43						
atgataagtg	agagcagact	ggagacctct	ggggctctga	tgagagagtc	tccccacgca	60
cacagccccc	tacaagcgat	cctggggaggc	tccaagcgc	ttggccactc	agaccagaaa	120
gatttcctgc	gctccaacca	aagccttcat	ccatttttct	ctgggttcgga	agccctgctg	180
gctgaagggg	agcacctggg	agtctccacc	acccaccac	ccacccccag	catgaaaacc	240
ttaaagaaac	agtgggttaga	agtactgagc	acctaccaat	atgtgctctg	tttctcttct	300

pctau2012001598-seq1

ctgggccctt	tcttttcct	tgttggttc	ttctctct	tcacctcgct	ctggtatctc	360
tctgttctct	acttggtatg	gttattcctg	gactgggaca	caccccagca	aggtggaagg	420
cgtaatcagt	ggttgaagaa	ctgcactgtg	tggaacacc	tgagcgatta	tttccccatt	480
aagctggtga	aaacagtaga	gctgccccca	gacaggaact	acgtgctatt	gtcccaccca	540
catgggatca	tggtgtttgg	aaccgtctgt	aacttcagca	cggagggcac	tggtgtctcg	600
cagctgttcc	cagggcttcg	gttctcactg	gccgtgttga	attgtctctt	gtacgtgcca	660
ggctgtcgag	agtacattat	gtcctgtgga	acatgttctg	tgaaccgtca	gagcctggat	720
tatgttctat	ctcaaccaca	gctcggcagg	gctgtggtca	tcattggtcgg	aggggccaat	780
gaggccctgc	ataccgtccc	aggggagcac	tgctcactc	tccggaatcg	taaaggcttc	840
gtccgcctgg	cactgaggca	cgggtgcctcc	ctggtgcccg	tgtactcctt	tggggagaat	900
gacgtcttca	gagttaaggc	ttttgccccca	gactcctggc	agcgtctgtt	ccaggtcacc	960
atcaagaggc	tcctgagctt	ctctccttgc	atcttctggg	gccgtggtct	cttctcagcc	1020
aagtcctggg	gcctgatgcc	cctcgccaga	cccatcacca	ctgtggtggg	ccgccccatc	1080
ccggtgcccc	agtgtccaca	gccaccgag	gagcaggtgg	accactatca	caggctctac	1140
atgaaggctc	tggagcaact	gtttgaggag	cacaagaaga	gctgcggcct	cccggcttct	1200
actcacctca	ccttcatcta	g				1221

<210> 44
 <211> 1872
 <212> DNA
 <213> Danio rerio

<400> 44						
acacacagtc	gctccaccct	catacagacg	tcgcacgagc	gaggagcata	cattcactca	60
aactcaaggg	atatttaaaa	gatatttccc	gacttaatga	cattttctag	ctcaagcagg	120
tattcgcggc	gtcttctagt	caccagtcg	tgacagcaaa	gactcgaatt	ctttctctgt	180
ttacaaatga	ttcacgttaa	agcaaagcgt	cgtttacagg	gggttcagtc	atgcggacgc	240
ggattcaaaa	gtggctggag	ataaacctca	gatagccagt	ccagaaacta	cccacataca	300
tttgcaatga	cgcgataaga	caacgaatcg	ttatctgttt	gaccaattga	accgatcaaa	360
gtgcgagcca	gacaaataca	agacttcggt	cagtgaaggt	gaggaagagg	attcgtttgc	420
cacaggatca	ctaaacttgg	ccactttaat	gggaacagat	aacgtttcag	aggtaaagga	480
gaagggggac	aggtcaccat	ggaaagatat	aatcgaggat	ataagtgttt	tgcaattggt	540
gttgtcattt	cttttttttag	gagtggcttg	tctactattg	atgatctacc	tgatgttcac	600
ctctctgtgg	attttcccca	ctctgtactt	cacctggcag	atctatgatt	ggcacacacc	660
tgagagaggt	ggcagaagaa	caaagtttgt	gagaggctgg	gaagtatgga	aacacttgcg	720
agactatttt	cctgtgaagc	tggtaaagac	tgctgaatta	aatcccaata	agaactatat	780
tatgggctgt	cacccgcacg	gcatcatgtg	ttttggagcc	ttctcctgct	tcagcacaga	840

pctau2012001598-seq1

```

ccgcaatgga tttgcggaaa cctttcctgg aatacgggcc actcttgcaa tcttagccgg      900
acttttccgc cttcctctct tcaggaata catttttagct gcaggattgc tgcctgtcag      960
taaggcaagt ttggattatt tgctgagcca aactggtgtg ggtaacgctg tggatcatcat    1020
cattggtgga gcagaggaat cactgacctc ctctacagga gtaaactctg tggtcataaa    1080
acacaggaaa ggctttgtac gactggctct tgaacacggg gcagacctgg ttccagtgtg    1140
ctcttttgga gagaacgagc tgttcccgca ggtggttctg tctgagggca gtgtaggctg    1200
acggctgcag gttcttttta agcagattat gggttttgct ccatgcatct tcacaggagg    1260
ccggtggctt cttctgccgt acaaacttcc agtcaccact gtggtgggct gcccataaa    1320
cgtccctctg gtgaaaattc caactcagga acagggtgat cattaccacg gtttatacat    1380
ggcatctctg gctgatcttt tccacaagca taagaccagc tatggactgg cagagacgca    1440
cgagctgcac ttcatttagc agtgctggaa agcacaaagc catttaaaaa gctgactggg    1500
atataatgtg gggctgatat aggaaaagga caatgcactt atttctgtac tgttccatgt    1560
gaatgtttta acttcttggg tctctgaaat gattcacaaa gcacaaaatg aaccaaata    1620
atacggagat gggcatttgt ctgtaacaaa acactcatcc ttaagaaagg caaaagggtt    1680
gcttacgggt atctcagctt tgttctgtaa atttgactg aaatttttac aacatgttag    1740
tgtttattat aaagggtata caaagacttt tgaagaaaat catttttaaa aattggttta    1800
atataaatga tttacatatg tattttgtta tgtctattca tgtgaaagat tataaaacca    1860
gcacattttt ac                                                              1872

```

<210> 45
 <211> 334
 <212> PRT
 <213> Homo sapiens

<400> 45

Met Lys Val Glu Phe Ala Pro Leu Asn Ile Gln Leu Ala Arg Arg Leu
 1 5 10 15

Gln Thr Val Ala Val Leu Gln Trp Val Leu Ser Phe Leu Thr Gly Pro
 20 25 30

Met Ser Ile Gly Ile Thr Val Met Leu Ile Ile His Asn Tyr Leu Phe
 35 40 45

Leu Tyr Ile Pro Tyr Leu Met Trp Leu Tyr Phe Asp Trp His Thr Pro
 50 55 60

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

Lys His Phe Lys Asp Tyr Phe Pro Ile His Leu Ile Lys Thr Gln Asp
 85 90 95

pctau2012001598-seql

Leu Asp Pro Ser His Asn Tyr Ile Phe Gly Phe His Pro His Gly Ile
100 105 110

Met Ala Val Gly Ala Phe Gly Asn Phe Ser Val Asn Tyr Ser Asp Phe
115 120 125

Lys Asp Leu Phe Pro Gly Phe Thr Ser Tyr Leu His Val Leu Pro Leu
130 135 140

Trp Phe Trp Cys Pro Val Phe Arg Glu Tyr Val Met Ser Val Gly Leu
145 150 155 160

Val Ser Val Ser Lys Lys Ser Val Ser Tyr Met Val Ser Lys Glu Gly
165 170 175

Gly Gly Asn Ile Ser Val Ile Val Leu Gly Gly Ala Lys Glu Ser Leu
180 185 190

Asp Ala His Pro Gly Lys Phe Thr Leu Phe Ile Arg Gln Arg Lys Gly
195 200 205

Phe Val Lys Ile Ala Leu Thr His Gly Ala Ser Leu Val Pro Val Val
210 215 220

Ser Phe Gly Glu Asn Glu Leu Phe Lys Gln Thr Asp Asn Pro Glu Gly
225 230 235 240

Ser Trp Ile Arg Thr Val Gln Asn Lys Leu Gln Lys Ile Met Gly Phe
245 250 255

Ala Leu Pro Leu Phe His Ala Arg Gly Val Phe Gln Tyr Asn Phe Gly
260 265 270

Leu Met Thr Tyr Arg Lys Ala Ile His Thr Val Val Gly Arg Pro Ile
275 280 285

Pro Val Arg Gln Thr Leu Asn Pro Thr Gln Glu Gln Ile Glu Glu Leu
290 295 300

His Gln Thr Tyr Met Glu Glu Leu Arg Lys Leu Phe Glu Glu His Lys
305 310 315 320

Gly Lys Tyr Gly Ile Pro Glu His Glu Thr Leu Val Leu Lys
325 330

<210> 46
<211> 335
<212> PRT
<213> Mus musculus

<400> 46

Met Met Val Glu Phe Ala Pro Leu Asn Thr Pro Leu Ala Arg Cys Leu
Page 35

pctau2012001598-seql

```

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

```

275

Ile Pro Val Gln Gln Ile Leu Asn Pro Thr Ser Glu Gln Ile Glu Glu
290 295 300

Leu His Gln Thr Tyr Leu Glu Glu Leu Lys Lys Leu Phe Asn Glu His
305 310 315 320

Lys Gly Lys Tyr Gly Ile Pro Glu His Glu Thr Leu Val Phe Lys
325 330 335

<210> 47

<211> 334

<212> PRT

<213> Pan troglodytes

<400> 47

Met Lys Val Glu Phe Ala Pro Leu Asn Ile Gln Pro Ala Arg Arg Leu
1 5 10 15

Gln Thr Val Ala Val Leu Gln Trp Val Leu Ser Phe Leu Thr Gly Pro
20 25 30

Met Ser Ile Gly Ile Thr Val Met Leu Ile Ile His Ser Tyr Ser Phe
35 40 45

Leu Tyr Ile Pro Tyr Leu Met Trp Leu Tyr Phe Asp Trp His Thr Pro
50 55 60

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

Lys His Phe Lys Asp Tyr Phe Pro Ile His Leu Ile Lys Thr Gln Asp
85 90 95

Leu Asp Pro Ser His Asn Tyr Ile Phe Gly Phe His Pro His Gly Ile
100 105 110

Met Ala Val Gly Ala Phe Gly Asn Phe Ser Val Asn Tyr Ser Asp Phe
115 120 125

Lys Asp Leu Phe Pro Gly Phe Thr Ser Tyr Leu His Val Leu Pro Leu
130 135 140

Trp Phe Trp Cys Pro Val Phe Arg Glu Tyr Val Leu Ser Val Gly Leu
145 150 155 160

Val Ser Val Ser Lys Lys Ser Val Ser Tyr Met Val Ser Lys Glu Gly
165 170 175

Gly Gly Asn Ile Ser Val Ile Val Leu Gly Gly Ala Lys Glu Ser Leu
180 185 190

pctau2012001598-seql

Asp Ala His Pro Gly Lys Phe Thr Leu Phe Ile His Gln Arg Lys Gly
195 200 205

Phe Val Lys Thr Ala Leu Thr His Gly Ala Ser Leu Val Pro Val Val
210 215 220

Ser Phe Gly Glu Asn Glu Leu Phe Lys Gln Thr Asp Asn Pro Glu Gly
225 230 235 240

Ser Trp Ile Arg Thr Val Gln Asn Lys Leu Gln Lys Ile Met Gly Phe
245 250 255

Ala Leu Pro Leu Phe His Ala Arg Gly Val Phe Gln Tyr Asn Phe Gly
260 265 270

Leu Met Pro Tyr Arg Lys Ala Ile His Thr Val Val Gly Arg Pro Ile
275 280 285

Pro Val Arg Gln Thr Leu Asn Pro Thr Gln Glu Gln Ile Glu Glu Leu
290 295 300

His Gln Thr Tyr Met Glu Glu Leu Arg Lys Leu Phe Glu Glu His Lys
305 310 315 320

Gly Lys Tyr Gly Ile Pro Glu His Glu Thr Leu Val Leu Lys
325 330

<210> 48
<211> 335
<212> PRT
<213> Pan troglodytes

<400> 48

Met Lys Val Glu Phe Ala Pro Leu Asn Ile Gln Pro Ala Arg Arg Leu
1 5 10 15

Gln Thr Val Ala Val Leu Gln Trp Val Leu Lys Tyr Leu Leu Leu Gly
20 25 30

Pro Met Ser Ile Gly Ile Thr Val Met Leu Ile Ile His Ser Tyr Ser
35 40 45

Phe Leu Tyr Ile Pro Tyr Leu Met Trp Leu Tyr Phe Asp Trp His Thr
50 55 60

Pro Glu Arg Gly Gly Arg Arg Ser Ser Trp Ile Lys Asn Trp Thr Leu
65 70 75 80

Trp Lys His Phe Lys Asp Tyr Phe Pro Ile His Leu Ile Lys Thr Gln
85 90 95

pctau2012001598-seql

Asp Leu Asp Pro Ser His Asn Tyr Ile Phe Gly Phe His Pro His Gly
100 105 110

Ile Met Ala Val Gly Ala Phe Gly Asn Phe Ser Val Asn Tyr Ser Asp
115 120 125

Phe Lys Asp Leu Phe Pro Gly Phe Thr Ser Tyr Leu His Val Leu Pro
130 135 140

Leu Trp Phe Trp Cys Pro Val Phe Arg Glu Tyr Val Leu Ser Val Gly
145 150 155 160

Leu Val Ser Val Ser Lys Lys Ser Val Ser Tyr Met Val Ser Lys Glu
165 170 175

Gly Gly Gly Asn Ile Ser Val Ile Val Leu Gly Gly Ala Lys Glu Ser
180 185 190

Leu Asp Ala His Pro Gly Lys Phe Thr Leu Phe Ile His Gln Arg Lys
195 200 205

Gly Phe Val Lys Thr Ala Leu Thr His Gly Ala Ser Leu Val Pro Val
210 215 220

Val Ser Phe Gly Glu Asn Glu Leu Phe Lys Gln Thr Asp Asn Pro Glu
225 230 235 240

Gly Ser Trp Ile Arg Thr Val Gln Asn Lys Leu Gln Lys Ile Met Gly
245 250 255

Phe Ala Leu Pro Leu Phe His Ala Arg Gly Val Phe Gln Tyr Asn Phe
260 265 270

Gly Leu Met Pro Tyr Arg Lys Ala Ile His Thr Val Val Gly Arg Pro
275 280 285

Ile Pro Val Arg Gln Thr Leu Asn Pro Thr Gln Glu Gln Ile Glu Glu
290 295 300

Leu His Gln Thr Tyr Met Glu Glu Leu Arg Lys Leu Phe Glu Glu His
305 310 315 320

Lys Gly Lys Tyr Gly Ile Pro Glu His Glu Thr Leu Val Leu Lys
325 330 335

<210> 49
<211> 335
<212> PRT
<213> Canis familiaris
<400> 49

pctau2012001598-seql

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

pctau2012001598-seql

Gly Leu Met Pro Tyr Arg Lys Pro Ile His Thr Val Val Gly Arg Pro
275 280 285

Ile Pro Val His Arg Thr Pro His Pro Ser Pro Gly Gln Ile Glu Glu
290 295 300

Leu His Gln Thr Tyr Met Glu Glu Leu Arg Lys Leu Phe Glu Ala His
305 310 315 320

Lys Arg Lys Tyr Gly Ile Pro Glu His Glu Thr Leu Ile Phe Lys
325 330 335

<210> 50

<211> 335

<212> PRT

<213> Bos taurus

<400> 50

Met Lys Val Glu Phe Ala Pro Leu Asn Ile Pro Leu Ala Arg Arg Leu
1 5 10 15

Gln Thr Ala Ala Val Leu His Trp Leu Leu Ser Phe Leu Leu Phe Ala
20 25 30

Gln Val Cys Leu Gly Ile Ile Val Phe Leu Ile Ile Tyr Asn Tyr Trp
35 40 45

Phe Leu Tyr Leu Pro Tyr Leu Thr Trp Leu Tyr Phe Asp Trp Gln Thr
50 55 60

Pro Glu Gln Gly Gly Arg Arg Ser Glu Trp Val Arg Asn Trp Ala Ile
65 70 75 80

Trp Arg Tyr Phe Lys Asp Tyr Phe Pro Ile His Leu Ile Lys Thr Trp
85 90 95

Asp Leu Asp Pro Ser His Asn Tyr Ile Phe Gly Phe His Pro His Gly
100 105 110

Val Leu Val Val Gly Ala Phe Gly Asn Phe Cys Thr Asn Tyr Ser Ala
115 120 125

Phe Lys Glu Leu Phe Pro Gly Phe Thr Ser Tyr Leu His Val Leu Pro
130 135 140

Tyr Trp Phe Arg Cys Pro Leu Phe Arg Glu Tyr Leu Met Ser Ser Gly
145 150 155 160

Pro Val Ser Val Ser Lys Lys Ser Val Cys His Val Leu Ser Lys Glu
165 170 175

Gly Gly Gly Asn Ile Ser Val Ile Val Leu Gly Gly Ala Glu Glu Ser

pctau2012001598-seql

180

185

190

Leu Asp Ala His Pro Gly Lys Phe Thr Leu Phe Ile Arg Gln Arg Lys
 195 200 205

Gly Phe Val Lys Ile Ala Leu Thr His Gly Ala Tyr Leu Val Pro Val
 210 215 220

Phe Ser Phe Gly Glu Asn Glu Leu Phe Lys Gln Val Ser Asn Pro Glu
 225 230 240

Gly Ser Trp Leu Arg Asn Val Gln Glu Lys Leu Gln Lys Ile Met Gly
 245 250 255

Phe Ala Leu Pro Leu Phe His Ala Arg Gly Ile Phe Gln Tyr Asn Phe
 260 265 270

Gly Leu Ile Pro Tyr Arg Lys Pro Ile His Thr Val Val Gly Arg Pro
 275 280 285

Ile Pro Val Arg Gln Thr Leu Asn Pro Thr Ser Glu Gln Ile Glu Glu
 290 295 300

Leu His Gln Thr Tyr Met Glu Glu Leu Arg Lys Leu Phe Glu Glu His
 305 310 315 320

Lys Gly Lys Tyr Gly Ile Pro Glu Asn Glu Thr Leu Ile Phe Arg
 325 330 335

<210> 51

<211> 183

<212> PRT

<213> Rattus norvegicus

<400> 51

Met Ser Lys Glu Gly Gly Gly Asn Ile Ser Val Ile Val Ile Gly Gly
 1 5 10 15

Ala Lys Glu Leu Leu Glu Ser Phe Pro Gly Arg Tyr Ser Leu Cys Leu
 20 25 30

Leu Gln Arg Lys Gly Phe Val Lys Ile Ala Leu Thr His Gly Ala His
 35 40 45

Leu Val Pro Val Phe Ser Phe Gly Glu Asn Glu Leu Tyr Ser Gln Val
 50 55 60

Asp Asn Pro Lys Gly Ser Trp Leu Arg Thr Ala Gln Asp Lys Val Tyr
 65 70 75 80

Asn Leu Thr Gly Leu Ala Leu Pro Leu Phe Tyr Ala Arg Gly Ile Phe
 85 90 95

pctau2012001598-seql

Gln Asn Ser Phe Gly Leu Met Pro Tyr Arg Lys Leu Ile Tyr Thr Val
100 105 110

Ala Ala Val Gly Thr Ala Ala Arg Leu Thr Gln Lys Ser Asn His Gln
115 120 125

Asn Asp Ser Ser Val Ser Ala Glu Lys Gly Asn Ser Phe Ile Cys Cys
130 135 140

Gln Val Gly Trp Cys Asp Arg Asp Arg Gln Pro Leu Ala Ala Pro Ser
145 150 155 160

Leu Phe Arg Arg Leu Cys Thr Arg Pro Gln Ser Arg Leu Met Ser Cys
165 170 175

Ile Arg Arg Thr Trp Arg Ser
180

<210> 52
<211> 287
<212> PRT
<213> Danio rerio

<400> 52

Met Leu Gly Ala Leu Tyr Ala Gly Trp Leu Tyr Leu Asp Arg Asp Thr
1 5 10 15

Pro Ser Cys Gly Gly Arg Arg Ser Gln Trp Val Arg Ser Trp Arg Ile
20 25 30

Trp Met His Phe Arg Asp Tyr Phe Pro Ile Thr Leu Val Lys Thr Val
35 40 45

Asp Leu Asp Pro Arg His Asn Tyr Leu Leu Gly Phe His Pro His Gly
50 55 60

Val Leu Val Ala Gly Gly Phe Gly Asn Phe Cys Thr Glu Ala Ser Gly
65 70 75 80

Phe Ser Gln Met Phe Pro Gly Leu Thr Pro Tyr Leu Leu Met Leu Pro
85 90 95

Phe Trp Phe Arg Val Pro Phe Phe Arg Glu Tyr Ile Met Cys Gly Gly
100 105 110

Leu Val Ser Ser Glu Lys Ala Ser Ala Ser Tyr Leu Leu Gly His Pro
115 120 125

Gly Gly Gly Gln Ala Ala Val Ile Ala Val Gly Gly Ala Pro Glu Ser
130 135 140

pctau2012001598-seql

Leu Glu Ala Arg Pro Gly Ala Leu Thr Leu Gln Leu Leu Gln Arg Lys
145 150 155 160

Gly Phe Ile Lys Leu Ala Leu Lys His Gly Ala Trp Leu Val Pro Val
165 170 175

Phe Ser Phe Gly Glu Asn Glu Leu Phe Asp Gln Met Glu Asn Pro Ala
180 185 190

Gly Ser Ala Leu Arg Arg Met Gln Glu Arg Leu Gln Arg Ile Met Gly
195 200 205

Val Ala Leu Pro Leu Phe His Ala Arg Gly Val Phe Gln Tyr Ser Phe
210 215 220

Gly Leu Leu Pro Tyr Arg Lys Pro Ile His Thr Val Val Gly Arg Pro
225 230 235 240

Ile Pro Val Ser Gln Thr Pro Cys Pro Ser Lys Glu Asp Ile Asp Ala
245 250 255

Leu His Thr Leu Tyr Met Gln Gly Leu Thr Gln Val Phe Glu Glu Asn
260 265 270

Lys Lys His Tyr Gly Ile Ala Asp Asp Lys His Leu Lys Phe Thr
275 280 285

<210> 53
<211> 349
<212> PRT
<213> Caenorhabditis elegans

<400> 53

Met Leu Asp Arg Thr Glu Ser Glu Asp Lys Met Pro His Leu Leu Gly
1 5 10 15

Val Glu Trp Ala Pro Leu Asn Ile Pro Leu Ala Arg Arg Leu Gln Thr
20 25 30

Leu Gly Ala Leu His Phe Phe Phe Ile Thr Leu Phe Thr Pro Val Leu
35 40 45

Val Leu Thr Val Pro Phe Tyr Met Leu Tyr Thr Val Leu Trp Pro Leu
50 55 60

Ile Phe Leu Tyr Gly Leu Trp Met Ile Tyr Asp Trp Asn Ser Pro Lys
65 70 75 80

Lys Gly Ala Tyr Met Ser Asn Trp Phe Gln Arg Gln Arg Ile His Ser
85 90 95

pctau2012001598-seql

Trp Tyr Ala Asn Tyr Phe Pro Val Lys Leu His Thr Thr Ser Asp Met
100 105 110

Pro Glu Glu His Asn Tyr Leu Ile Gly Tyr His Pro His Gly Ile Ile
115 120 125

Ser Met Ala Ala Phe Ile Asn Phe Ala Thr Asn Gly Thr Gly Ile Leu
130 135 140

Asp Thr Leu Pro Arg Ile Arg Phe His Leu Cys Thr Leu Val Gly Gln
145 150 155 160

Phe Trp Thr Pro Trp Arg Arg Glu Trp Gly Leu Leu His Gly Met Ile
165 170 175

Asp Cys Ser Arg Glu Ser Ile Lys His Val Leu Glu His Glu Lys Lys
180 185 190

Gly Lys Ala Val Val Leu Val Val Gly Gly Ala Glu Glu Ala Leu Asp
195 200 205

Ala His Pro Gly Cys His Ile Leu Thr Leu Lys Lys Arg Lys Gly Phe
210 215 220

Val Lys Ile Ala Leu Gln Thr Gly Ala Gln Leu Val Pro Cys Tyr Ser
225 230 235 240

Phe Gly Glu Asn Asp Ile Phe Asn Gln Ala Glu Asn Pro Lys Gly Ser
245 250 255

Thr Ile Arg Gln Phe Gln Thr Ile Met Lys Arg Val Leu Gly Phe Ser
260 265 270

Pro Pro Ala Phe Tyr Gly Arg Gly Val Phe Asn Tyr Thr Phe Gly Leu
275 280 285

Leu Pro Phe Arg Lys Pro Ile Asn Thr Val Leu Gly Ala Pro Ile Ser
290 295 300

Val Thr Lys Thr Val Asn Pro Thr Gln Glu Gln Ile Asp Thr Leu His
305 310 315 320

Gln Thr Tyr Met Asp Arg Leu His Glu Leu Phe Glu Glu His Lys Thr
325 330 335

Lys Tyr Asp Val Ser Pro Thr Thr Gln Leu Val Ile Asn
340 345

<210> 54
<211> 339
<212> PRT
<213> Caenorhabditis elegans

pctau2012001598-seql

<400> 54

Met Pro His Leu Leu Gly Val Glu Trp Ala Pro Leu Asn Ile Pro Leu
1 5 10 15

Ala Arg Arg Leu Gln Thr Leu Gly Ala Leu His Phe Phe Phe Ile Thr
20 25 30

Leu Phe Thr Pro Val Leu Val Leu Thr Val Pro Phe Tyr Met Leu Tyr
35 40 45

Thr Val Leu Trp Pro Leu Ile Phe Leu Tyr Gly Leu Trp Met Ile Tyr
50 55 60

Asp Trp Asn Ser Pro Lys Lys Gly Ala Tyr Met Ser Asn Trp Phe Gln
65 70 75 80

Arg Gln Arg Ile His Ser Trp Tyr Ala Asn Tyr Phe Pro Val Lys Leu
85 90 95

His Thr Thr Ser Asp Met Pro Glu Glu His Asn Tyr Leu Ile Gly Tyr
100 105 110

His Pro His Gly Ile Ile Ser Met Ala Ala Phe Ile Asn Phe Ala Thr
115 120 125

Asn Gly Thr Gly Ile Leu Asp Thr Leu Pro Arg Ile Arg Phe His Leu
130 135 140

Cys Thr Leu Val Gly Gln Phe Trp Thr Pro Trp Arg Arg Glu Trp Gly
145 150 155 160

Leu Leu His Gly Met Ile Asp Cys Ser Arg Glu Ser Ile Lys His Val
165 170 175

Leu Glu His Glu Lys Lys Gly Lys Ala Val Val Leu Val Val Gly Gly
180 185 190

Ala Glu Glu Ala Leu Asp Ala His Pro Gly Cys His Ile Leu Thr Leu
195 200 205

Lys Lys Arg Lys Gly Phe Val Lys Ile Ala Leu Gln Thr Gly Ala Gln
210 215 220

Leu Val Pro Cys Tyr Ser Phe Gly Glu Asn Asp Ile Phe Asn Gln Ala
225 230 235 240

Glu Asn Pro Lys Gly Ser Thr Ile Arg Gln Phe Gln Thr Ile Met Lys
245 250 255

Arg Val Leu Gly Phe Ser Pro Pro Ala Phe Tyr Gly Arg Gly Val Phe
Page 46

pctau2012001598-seql

260

265

270

Asn Tyr Thr Phe Gly Leu Leu Pro Phe Arg Lys Pro Ile Asn Thr Val
 275 280 285

Leu Gly Ala Pro Ile Ser Val Thr Lys Thr Val Asn Pro Thr Gln Glu
 290 295 300

Gln Ile Asp Thr Leu His Gln Thr Tyr Met Asp Arg Leu His Glu Leu
 305 310 315 320

Phe Glu Glu His Lys Thr Lys Tyr Asp Val Ser Pro Thr Thr Gln Leu
 325 330 335

Val Ile Asn

<210> 55

<211> 340

<212> PRT

<213> Caenorhabditis elegans

<400> 55

Met Pro Gln Phe Leu Gly Ile Glu Trp Val Asp Leu Phe Ser Ser Ile
 1 5 10 15

Gln Arg Lys Lys Thr Tyr Leu Gly Val Val Tyr His Phe Met Leu Thr
 20 25 30

Tyr Pro Leu Ala Leu Phe Val Thr Ile Leu Pro Phe Phe Leu Leu Phe
 35 40 45

Thr Phe Gln Trp His Ile Leu Ala Leu Tyr Ala Cys Trp Tyr Phe Tyr
 50 55 60

Asp Met Asp Ser Pro Arg Arg Gly Gly Tyr Ser Ser Asp Trp Val Arg
 65 70 75 80

Lys Trp Arg Val Asn Asp Trp Phe Ala Gln Tyr Phe Pro Ile Asn Leu
 85 90 95

His Lys Thr Ala Glu Leu Ser Thr Asp Lys Asn Tyr Leu Val Gly Ile
 100 105 110

His Pro His Gly Ile Ile Ser Met Ala Ala Trp Ser Asn Phe Ala Thr
 115 120 125

Asn Gly Thr Gly Ile Tyr Glu Lys Phe Pro Gly Ile Arg Trp Asn Leu
 130 135 140

Cys Thr Leu Ala Leu Gln Phe Arg Met Ala Ile Arg Arg Glu Leu Leu
 145 150 155 160

pctau2012001598-seql

Leu Leu Thr Gly Leu Ile Asp Cys Ser Arg Glu Ser Ile Glu Tyr Val
165 170 175

Leu Asp Lys Cys Gly Gln Lys Gly Arg Ala Val Val Leu Val Ile Gly
180 185 190

Gly Ala Glu Glu Ala Leu Asp Ala His Pro Gly Tyr His Thr Leu Thr
195 200 205

Leu Ala Ser Arg Lys Gly Phe Val Arg Glu Ala Leu Ile Thr Gly Ala
210 215 220

Tyr Leu Val Pro Val Tyr Ser Phe Gly Glu Asn Asp Val Phe Glu Gln
225 230 235 240

Met Glu Asn Pro Val Gly Ser Arg Leu Arg Asn Phe Gln Glu Trp Cys
245 250 255

Lys Ser Ile Phe Gly Ile Ser Tyr Pro Ile Phe His Gly Arg Gly Phe
260 265 270

Phe Gln Leu Thr Phe Gly Tyr Leu Pro Phe Arg Lys Pro Ile Asp Thr
275 280 285

Val Val Gly Ala Pro Ile Pro Val Glu Lys Val Glu Asn Pro Thr Lys
290 295 300

Glu Gln Ile Asp Glu Leu His Thr Ile Tyr Cys Gln Lys Leu Thr Glu
305 310 315 320

Leu Phe Asp Glu His Lys Glu Lys Tyr Gly Val Glu Lys Asp Val Pro
325 330 335

Leu Val Leu Arg
340

<210> 56
<211> 345
<212> PRT
<213> Caenorhabditis elegans

<400> 56

Met Leu Asn Tyr Gln Ile His Lys Lys Leu Thr Asp Ile Lys Trp Val
1 5 10 15

Asn Ile Phe Ser Pro Trp Asp Arg Gln Arg Ala Tyr Phe Ala Leu Val
20 25 30

Val Trp Phe Gly Leu Ile Tyr Pro Phe Cys Cys Leu Cys Gln Val Ala
35 40 45

pctau2012001598-seql

```

Pro Phe Val Leu Phe Phe Thr Gly Gln Trp Ile Ile Leu Gly Leu Tyr
 50          55          60

Ala Val Trp Tyr Leu Tyr Asp Arg Glu Ser Pro Arg Arg Gly Gly Tyr
65          70          75          80

Arg Asp Asn Trp Phe Arg Asn Leu Ser Leu His Lys Trp Phe Ala Glu
      85          90          95

Tyr Phe Pro Val Lys Leu His Lys Thr Ala Glu Leu Asp Pro Asn Gln
      100          105          110

Asn Tyr Leu Phe Gly Tyr His Pro His Gly Ile Leu Gly Val Gly Ala
     115          120          125

Trp Ser Cys Phe Gly Phe Asp Ala Cys Asn Val Lys Gln Val Phe Lys
     130          135          140

Gly Ile Arg Phe Asn Ile Cys Thr Leu Pro Gly Asn Phe Thr Ala Met
145          150          155          160

Phe Arg Arg Glu Ile Leu Leu Ser Ile Gly Met Ile Glu Ser Ser Lys
      165          170          175

Glu Ser Ile Glu His Val Leu Asn Ser Glu Glu Lys Gly Arg Ala Val
      180          185          190

Val Ile Val Val Gly Gly Ala Ala Glu Ala Leu Glu Ala His Pro Gly
     195          200          205

Lys His Thr Leu Thr Leu Ala Asn Arg Lys Gly Phe Val Arg Glu Ala
     210          215          220

Val Lys Thr Gly Ala His Leu Val Pro Val Tyr Ala Phe Gly Glu Asn
225          230          235          240

Asp Ile Tyr Lys Gln Ile Asp Asn Pro Glu Gly Ser Lys Leu Arg Lys
      245          250          255

Ile Gln Glu Trp Gly Lys Lys Lys Met Gly Ile Ser Leu Pro Leu Ile
      260          265          270

Tyr Gly Arg Gly Tyr Phe Gln Met Ala Leu Gly Leu Leu Pro Met Ser
     275          280          285

Arg Ala Val Asn Val Val Val Gly Ala Pro Ile Gln Val Glu Lys Glu
     290          295          300

Leu Asp Pro Ser Lys Glu Val Ile Asp Glu Ile His Gly Val Tyr Met
305          310          315          320

```

pctau2012001598-seql

Glu Lys Leu Ala Glu Leu Phe Glu Glu His Lys Ala Lys Phe Gly Val
325 330 335

Ser Lys Asp Thr Arg Leu Val Phe Gln
340 345

<210> 57
<211> 359
<212> PRT
<213> Caenorhabditis elegans
<400> 57

Met Arg Leu Arg Leu Ser Ser Ile Ser Gly Lys Ala Lys Leu Pro Asp
1 5 10 15

Lys Glu Ile Cys Ser Ser Val Ser Arg Ile Leu Ala Pro Leu Leu Val
20 25 30

Pro Trp Lys Arg Arg Leu Glu Thr Leu Ala Val Met Gly Phe Ile Phe
35 40 45

Met Trp Val Ile Leu Pro Ile Met Asp Leu Trp Val Pro Phe His Val
50 55 60

Leu Phe Asn Thr Arg Trp Trp Phe Leu Val Pro Leu Tyr Ala Val Trp
65 70 75 80

Phe Tyr Tyr Asp Phe Asp Thr Pro Lys Lys Ala Ser Arg Arg Trp Asn
85 90 95

Trp Ala Arg Arg His Val Ala Trp Lys Tyr Phe Ala Ser Tyr Phe Pro
100 105 110

Leu Arg Leu Ile Lys Thr Ala Asp Leu Pro Ala Asp Arg Asn Tyr Ile
115 120 125

Ile Gly Ser His Pro His Gly Met Phe Ser Val Gly Gly Phe Thr Ala
130 135 140

Met Ser Thr Asn Ala Thr Gly Phe Glu Asp Lys Phe Pro Gly Ile Lys
145 150 155 160

Ser His Ile Met Thr Leu Asn Gly Gln Phe Tyr Phe Pro Phe Arg Arg
165 170 175

Glu Phe Gly Ile Met Leu Gly Gly Ile Glu Val Ser Lys Glu Ser Leu
180 185 190

Glu Tyr Thr Leu Thr Lys Cys Gly Lys Gly Arg Ala Cys Ala Ile Val
195 200 205

pctau2012001598-seql

Ile Gly Gly Ala Ser Glu Ala Leu Glu Ala His Pro Asn Lys Asn Thr
210 215 220

Leu Thr Leu Ile Asn Arg Arg Gly Phe Cys Lys Tyr Ala Leu Lys Phe
225 230 235 240

Gly Ala Asp Leu Val Pro Met Tyr Asn Phe Gly Glu Asn Asp Leu Tyr
245 250 255

Glu Gln Tyr Glu Asn Pro Lys Gly Ser Arg Leu Arg Glu Val Gln Glu
260 265 270

Lys Ile Lys Asp Met Phe Gly Leu Cys Pro Pro Leu Leu Arg Gly Arg
275 280 285

Ser Leu Phe Asn Gln Tyr Leu Ile Gly Leu Leu Pro Phe Arg Lys Pro
290 295 300

Val Thr Thr Val Met Gly Arg Pro Ile Arg Val Thr Gln Thr Asp Glu
305 310 315 320

Pro Thr Val Glu Gln Ile Asp Glu Leu His Ala Lys Tyr Cys Asp Ala
325 330 335

Leu Tyr Asn Leu Phe Glu Glu Tyr Lys His Leu His Ser Ile Pro Pro
340 345 350

Asp Thr His Leu Ile Phe Gln
355

<210> 58
<211> 412
<212> PRT
<213> Kluyveromyces lactis

<400> 58

Met Gly Lys Glu Gln Ile Arg Lys Arg Gln Thr Lys Lys Val Arg Lys
1 5 10 15

Gly Ser Gln Pro Val Ala Asn Gly Lys Ser Ser Val Val Gly Val Ser
20 25 30

Lys Asp Gly Lys Pro Glu Phe Cys Ala Ile Asp Thr Pro Ile Glu Arg
35 40 45

Arg Leu Gln Thr Leu Ser Val Ala Trp Tyr Val Glu Cys Ile Pro Leu
50 55 60

Met Ile Ile Ile Met Leu Phe Val Trp Val Asn Pro Leu Met Trp Ser
65 70 75 80

Phe Val Ile Pro Tyr Thr Ile Tyr Tyr Phe Ile Asp Arg Thr Ala Ser

pctau2012001598-seql

85

90

95

Asn Gly Asn Ala Val Lys Arg His Ser Lys Trp Phe Arg Ser Leu Lys
 100 105 110

Val Trp Phe Tyr Phe Arg Asp Tyr Phe Pro Ile Ser Met His Lys Ser
 115 120 125

Thr Glu Leu Glu Pro Thr Phe Thr Ser Ile Asp Ser Thr Glu Leu Glu
 130 135 140

Asn Asp Ala Ser Glu Pro Gly Tyr Leu Asp Ser Ser Gln Pro Val Leu
 145 150 155 160

Pro Asp Lys Trp Trp Asn Pro Phe Arg Glu Lys Asp Glu Thr Val Arg
 165 170 175

Pro Thr Gly Pro Arg Tyr Ile Phe Gly Tyr His Pro His Gly Ile Ala
 180 185 190

Ala Phe Gly Ala Phe Gly Ala Phe Ala Thr Glu Ala Cys Asn Trp Ser
 195 200 205

Lys Val Phe Pro Gly Ile Pro Val Cys Leu Leu Thr Leu Val Asn Gln
 210 215 220

Phe Gln Ile Pro Val Tyr Arg Asp Tyr Leu Leu Ala Leu Gly Ile Thr
 225 230 235 240

Ser Val Ala Arg Lys Asn Ala Met Lys Val Leu Glu Lys Asn Tyr Ser
 245 250 255

Ile Ala Ile Val Ile Gly Gly Ala Ser Glu Ser Leu Leu Thr Asn Leu
 260 265 270

Gly Ser Ser Asp Ile Ile Leu Ser Lys Arg Lys Gly Phe Val Lys Leu
 275 280 285

Ala Leu Gln Thr Gly Asn Val Ser Leu Val Pro Val Tyr Gly Phe Gly
 290 295 300

Glu Thr Asp Thr Tyr Lys Ile Leu Lys Leu Lys Asn Asp Ser Ile Ile
 305 310 315 320

Gly Arg Ile Gln Ile Trp Leu Lys Glu Asn Tyr Ser Phe Thr Val Pro
 325 330 335

Leu Phe Phe Ala Arg Gly Val Phe Asn Tyr Asp Phe Gly Leu Leu Pro
 340 345 350

Phe Arg His Pro Val Asn Val Val Val Gly Asn Pro Ile His Ile Lys

355

Glu Lys Ile Asp His Pro Thr Ile Glu Glu Ile Asp His Tyr His Ser
370 375 380

Leu Tyr Ile Glu Glu Leu Lys Arg Leu Tyr Asp Asp Asn Lys Ala Lys
385 390 395 400

Phe Asn Tyr Ser Glu Lys Thr Leu Asn Ile Val Glu
405 410

<210> 59

<211> 461

<212> PRT

<213> Ashbya gossypii

<400> 59

Met Gln Asp Ser Met Asp Asp Ser Leu Arg Glu Ala Glu Gly Arg Gln
1 5 10 15

Asp Asp Ser Glu Val Ser Ser Gly Thr Thr Leu Gly Ser Ser Thr Pro
20 25 30

Glu Asp Ser Gly Val Thr Ala Lys Leu Arg Lys Lys Tyr Gln Met Ala
35 40 45

Ser Ala Leu Leu Arg Arg Glu Leu Glu Glu Leu Ser Val Tyr Asp Ala
50 55 60

Lys Thr Ala Gly Val Ser Gly Arg Ser Ser Gly Ser Gly Ser Gly Gly
65 70 75 80

Leu Ala Leu Leu Gly Gly Arg Phe His Val Ala Pro Leu Arg Ile Pro
85 90 95

Ala Arg Arg Arg Leu Gln Thr Leu Val Val Ala Trp His Thr Ser Ser
100 105 110

Phe Ile Tyr Met Thr Val Leu Val Leu Phe Leu Ala Ala Asn Pro Leu
115 120 125

Met Trp Trp Phe Met Val Pro Tyr Met Val Tyr Tyr Val Trp Asn Arg
130 135 140

Ser Pro Ala Asn Gly Gly Val Val Arg Arg Tyr Ser Pro Arg Leu Arg
145 150 155 160

Ser Leu Ala Leu Trp Arg Tyr Tyr Cys Glu Tyr Tyr Pro Ile Ser Leu
165 170 175

His Lys Ser Glu Asp Leu Ala Pro Thr Phe Val Pro Asp Pro Arg Gly
180 185 190

pctau2012001598-seql

Ala Glu Pro Arg Glu Trp Lys Leu Arg Leu Trp Leu Trp Pro Thr Arg
195 200 205
Val Glu Leu Leu Asn Leu Thr Leu Gln Trp Thr Arg Ala Arg Pro Gln
210 215 220
Val Ala Thr Gly Pro Arg Tyr Ile Phe Gly Tyr His Pro His Gly Val
225 230 235 240
Gly Ala Leu Gly Ala Phe Gly Ala Ile Ala Thr Glu Gly Cys Asn Trp
245 250 255
Ser Lys Val Phe Ala Gly Ile Pro Ala Cys Leu Cys Thr Leu Val Asn
260 265 270
Gln Phe Gln Ile Pro Ile Tyr Arg Asp Tyr Leu Leu Gly Leu Gly Cys
275 280 285
Thr Ser Val Ala Arg Lys Asn Val Leu Lys Val Leu Glu Gln Asn Tyr
290 295 300
Ser Val Cys Ile Val Val Gly Gly Ala Gln Glu Ala Leu Leu Ser Arg
305 310 315 320
Val Gly Ser Thr Glu Leu Val Leu Asn Lys Arg Lys Gly Phe Ile Lys
325 330 335
Leu Ala Leu Glu Thr Gly Asn Val Asn Leu Val Pro Ile Tyr Ala Phe
340 345 350
Gly Glu Thr Asp Cys Phe Asn Val Leu Asp Thr Gly Asn Glu Ser Tyr
355 360 365
Leu Arg Lys Phe Gln Leu Trp Ile Lys Lys Thr Tyr Gly Phe Thr Ile
370 375 380
Pro Phe Phe Phe Ala Arg Gly Val Phe Asn Tyr Asp Phe Gly Phe Leu
385 390 395 400
Pro Phe Arg Asn Pro Ile Asn Val Val Val Gly Lys Pro Val Tyr Val
405 410 415
Asp Lys Arg Arg Thr Asn Pro Thr Met Glu Glu Ile Asp His Tyr His
420 425 430
Asp Leu Tyr Val Gln Glu Leu Arg Asn Val Phe Asp Lys Asn Lys His
435 440 445
Lys Phe Gly Tyr Ala Gly Lys Glu Leu Lys Ile Val Glu
450 455 460

pctau2012001598-seql

<210> 60
 <211> 550
 <212> PRT
 <213> Magnaporthe oryzae

<400> 60

Met Ala Thr Val Asp Asp Thr Gln Asn His Val Gln Leu Asn Ala Glu
 1 5 10 15

Ser Ser Leu Asn Glu Phe Pro Tyr Gln Ala Ser Ala Asp Lys Glu Glu
 20 25 30

Tyr Pro Ser Leu Pro Ser Ala Pro Leu Pro Asp Val Asp Gly Ser Ser
 35 40 45

Gln Asn Gly His Thr Ala Ala Asn Ala Asp Leu Pro Pro Ser Glu Ser
 50 55 60

Thr Glu Gln Pro Trp Asn Lys Val Ala Ser Gly Asp Phe Ile His Glu
 65 70 75 80

Thr Lys Met Pro Ser Glu Asp Arg Asp Asn Leu Ala Arg Pro Gly Ser
 85 90 95

Leu Pro Thr Pro Glu Pro Ser Arg Arg Gly Ser Asn Thr Leu Val Gly
 100 105 110

Ser Pro Arg Gln Gly Asp Gly Leu Gln Met Leu Met Asp Glu Phe Ala
 115 120 125

Asp Asp Glu Ser Ser Asp Thr Ser Trp Asp Ile His Glu Thr Ile Arg
 130 135 140

Ser Arg Thr Gln Ala Leu Ala Glu Lys Asp His Gln Gly Ile Phe Ser
 145 150 155 160

Asn Gln Arg Arg Gly Asn Gln Ala Asp Ala Pro Asp Thr Val Thr Ser
 165 170 175

Ala Asn Ala Trp Leu Pro Ala Gly Phe Glu Asp Asp Leu Asp Glu Val
 180 185 190

Thr Gly Ala Asp Asp Glu Phe Glu Pro Thr Arg Gly Ser Arg Trp Arg
 195 200 205

Val Arg Leu Ala Pro Leu Asp Thr Pro Phe His Arg Arg Met Gln Thr
 210 215 220

Leu Val Val Leu Leu His Val Leu Gly Met Gly Ile Thr Phe Ser Phe
 225 230 235 240

pctau2012001598-seql

Phe Cys Phe Leu Cys Thr Leu Pro Leu Phe Trp Pro Ile Ile Ile Ala
245 250 255
Tyr Leu Val Phe Ile Arg Leu Ser Arg Ala Gly Ser Asp Gly Lys Thr
260 265 270
Asn Arg Arg Ile Glu Trp Leu Arg Arg Ala Arg Ile Trp Lys Tyr Phe
275 280 285
Ala Asp Tyr Phe Pro Val Lys Leu His Lys Thr Ala Asp Leu Pro Pro
290 295 300
Thr Arg Lys Tyr Ile Phe Ala Ile His Pro His Gly Ile Ile Ser His
305 310 315 320
Gly Ala Phe Ala Ser Phe Ala Thr Glu Ala Leu Gly Phe Ser Glu Lys
325 330 335
Phe Pro Gly Ile Thr Asn Ser Val Cys Thr Leu His Gly Asn Phe Lys
340 345 350
Thr Pro Phe Tyr Arg Glu Tyr Leu Met Ala Met Gly Leu Ile Ser Val
355 360 365
Ser Lys Glu Ser Ile Ile Asn Thr Leu Met Thr Gly Gly Ile Asn Gly
370 375 380
Glu Gly Met Gly Arg Ala Val Ser Ile Ile Val Gly Gly Ala Arg Glu
385 390 395 400
Ala Leu Glu Ala Cys Pro Lys Thr Met Arg Leu Ile Leu Lys Arg Arg
405 410 415
Gly Phe Cys Arg Met Ala Leu Arg Thr Gly Ala Asp Leu Val Pro Val
420 425 430
Leu Cys Phe Gly Glu Asn Asp Leu Tyr Gln Gln Trp Gly Pro Gln Asp
435 440 445
His Pro Arg Phe Arg Arg Leu Gln Met Lys Ala Leu Ala Tyr Leu Gly
450 455 460
Phe Ala Val Pro Val Leu Arg Gly Arg Gly Val Phe Asn Tyr Asp Phe
465 470 475 480
Gly Val Leu Pro Gln Arg Arg Pro Ile Asn Val Val Val Gly Glu Pro
485 490 495
Ile Lys Val Glu Gln Phe Arg Gly Ser Gly Asn Ile Glu Pro Arg Val
500 505 510

pctau2012001598-seql

Glu Glu Leu Trp Lys Leu Tyr Cys Gln Lys Leu Gln Glu Leu Tyr Asp
515 520 525

Gln Asn Lys His Val Tyr Phe Lys Asp Arg Gln Glu Asp Met Arg Phe
530 535 540

Val Asp Ser Glu Glu Leu
545 550

<210> 61
<211> 334
<212> PRT
<213> Ciona intestinalis

<400> 61

Met Trp Glu Phe Ala Pro Leu Arg Ile Pro Met Ile Arg Arg Leu Gln
1 5 10 15

Thr Leu Ala Val Ser His Phe Val Phe Cys Phe Leu Ala Leu Ala Ala
20 25 30

Thr Cys Thr Leu Leu Thr Ile Tyr Ile Thr Leu Phe Thr Ser Tyr Trp
35 40 45

Pro Ile Thr Thr Ala Tyr Leu Val Phe Leu Phe Leu Asp Arg Asn Thr
50 55 60

Pro Glu Ser Gly Gly Arg Arg Ser Glu Trp Val Arg Asn Trp Ser Leu
65 70 75 80

Trp Lys Trp Met Ser Asp Tyr Phe Pro Cys Thr Leu His Lys Thr Val
85 90 95

Asp Leu Asp Pro Lys Arg Asn Tyr Ile Phe Gly Ile His Pro His Gly
100 105 110

Val Leu Cys Ile Gly Ser Phe Thr His Phe Ser Thr Asn Gly Ser Gly
115 120 125

Phe Ser His Val Phe Pro Gly Phe Thr Ser Tyr Leu Thr Met Leu Pro
130 135 140

Phe Trp Phe Lys Met Pro Phe Phe Arg Asp Tyr Val Met Ser Gly Gly
145 150 155 160

Leu Thr Pro Ala Thr Arg Lys Ala Ile Lys His Thr Ile Thr Arg Pro
165 170 175

Gly Gly Gly His Ile Cys Cys Ile Ile Pro Gly Gly Ala Pro Glu Ser
180 185 190

pctau2012001598-seql

Leu Asn Ala Arg Pro Gly Asp Val Val Leu Leu Leu Lys Gln Arg Leu
195 200 205

Gly Phe Leu Lys Leu Ala Ile Thr Asn Gly Val Pro Leu Val Pro Val
210 215 220

Phe Ser Phe Gly Asp His Ala Leu Trp Glu Gln Lys Pro Asn Pro Pro
225 230 235 240

Gly Ser Leu Ile Arg Arg Phe Gln Asp Ser Ser Gln Lys Trp Met Gln
245 250 255

Val Ala Leu Pro Val Phe His Ala Arg Gly Ile Phe Gln Tyr Asn Phe
260 265 270

Gly Leu Ile Pro Tyr Arg Arg Ser Val His Thr Val Val Gly Glu Pro
275 280 285

Ile Glu Val Pro Gln Asn Ser Asn Pro Thr Ser Glu Asp Leu Met Ser
290 295 300

Leu Gln Glu Asp Tyr Ile Asn Arg Leu Arg Ala Ile Phe Asp Glu His
305 310 315 320

Lys Ser Lys Tyr Leu Pro Glu Asp Cys Lys Leu Ile Ile Asn
325 330

<210> 62
<211> 334
<212> PRT
<213> Homo sapiens

<400> 62

Met Val Glu Phe Ala Pro Leu Phe Met Pro Trp Glu Arg Arg Leu Gln
1 5 10 15

Thr Leu Ala Val Leu Gln Phe Val Phe Ser Phe Leu Ala Leu Ala Glu
20 25 30

Ile Cys Thr Val Gly Phe Ile Ala Leu Leu Phe Thr Arg Phe Trp Leu
35 40 45

Leu Thr Val Leu Tyr Ala Ala Trp Trp Tyr Leu Asp Arg Asp Lys Pro
50 55 60

Arg Gln Gly Gly Arg His Ile Gln Ala Ile Arg Cys Trp Thr Ile Trp
65 70 75 80

Lys Tyr Met Lys Asp Tyr Phe Pro Ile Ser Leu Val Lys Thr Ala Glu
85 90 95

Leu Asp Pro Ser Arg Asn Tyr Ile Ala Gly Phe His Pro His Gly Val
Page 58

pctau2012001598-seql

100

105

110

Leu Ala Val Gly Ala Phe Ala Asn Leu Cys Thr Glu Ser Thr Gly Phe
 115 120 125

Ser Ser Ile Phe Pro Gly Ile Arg Pro His Leu Met Met Leu Thr Leu
 130 135 140

Trp Phe Arg Ala Pro Phe Phe Arg Asp Tyr Ile Met Ser Ala Gly Leu
 145 150 155 160

Val Thr Ser Glu Lys Glu Ser Ala Ala His Ile Leu Asn Arg Lys Gly
 165 170 175

Gly Gly Asn Leu Leu Gly Ile Ile Val Gly Gly Ala Gln Glu Ala Leu
 180 185 190

Asp Ala Arg Pro Gly Ser Phe Thr Leu Leu Leu Arg Asn Arg Lys Gly
 195 200 205

Phe Val Arg Leu Ala Leu Thr His Gly Ala Pro Leu Val Pro Ile Phe
 210 215 220

Ser Phe Gly Glu Asn Asp Leu Phe Asp Gln Ile Pro Asn Ser Ser Gly
 225 230 235 240

Ser Trp Leu Arg Tyr Ile Gln Asn Arg Leu Gln Lys Ile Met Gly Ile
 245 250 255

Ser Leu Pro Leu Phe His Gly Arg Gly Val Phe Gln Tyr Ser Phe Gly
 260 265 270

Leu Ile Pro Tyr Arg Arg Pro Ile Thr Thr Val Val Gly Lys Pro Ile
 275 280 285

Glu Val Gln Lys Thr Leu His Pro Ser Glu Glu Glu Val Asn Gln Leu
 290 295 300

His Gln Arg Tyr Ile Lys Glu Leu Cys Asn Leu Phe Glu Ala His Lys
 305 310 315 320

Leu Lys Phe Asn Ile Pro Ala Asp Gln His Leu Glu Phe Cys
 325 330

<210> 63

<211> 334

<212> PRT

<213> Mus musculus

<400> 63

Met Val Glu Phe Ala Pro Leu Leu Val Pro Trp Glu Arg Arg Leu Gln
 1 5 10 15

pctau2012001598-seql

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

pctau2012001598-seql

Glu Val Gln Met Thr Pro Gln Pro Ser Arg Glu Glu Val Asp Arg Leu
290 295 300

His Gln Arg Tyr Ile Lys Glu Leu Cys Lys Leu Phe Glu Glu His Lys
305 310 315 320

Leu Lys Phe Asn Val Pro Glu Asp Gln His Leu Glu Phe Cys
325 330

<210> 64
<211> 334
<212> PRT
<213> Pan troglodytes

<400> 64

Met Val Glu Phe Ala Pro Leu Phe Val Pro Trp Glu Arg Arg Leu Gln
1 5 10 15

Thr Leu Ala Val Leu Gln Phe Val Phe Ser Phe Leu Ala Leu Ala Glu
20 25 30

Ile Cys Thr Val Gly Phe Ile Ala Leu Leu Phe Thr Arg Phe Trp Leu
35 40 45

Leu Thr Val Leu Tyr Ala Ala Trp Trp Tyr Leu Asp Arg Asp Lys Pro
50 55 60

Arg Gln Gly Gly Arg His Ile Gln Ala Ile Arg Cys Trp Thr Ile Trp
65 70 75 80

Lys Tyr Met Lys Asp Tyr Phe Pro Ile Ser Leu Val Lys Thr Ala Glu
85 90 95

Leu Asp Pro Ser Arg Asn Tyr Ile Ala Gly Phe His Pro His Gly Val
100 105 110

Leu Ala Val Gly Ala Phe Ala Asn Leu Cys Thr Glu Ser Thr Gly Phe
115 120 125

Ser Ser Ile Phe Pro Gly Ile Arg Pro His Leu Met Met Leu Thr Leu
130 135 140

Trp Phe Arg Ala Pro Phe Phe Arg Asp Tyr Ile Met Ser Ala Gly Leu
145 150 155 160

Val Thr Ser Glu Lys Glu Ser Ala Ala His Ile Leu Asn Arg Lys Gly
165 170 175

Gly Gly Asn Leu Leu Gly Ile Ile Val Gly Gly Ala Gln Glu Ala Leu
180 185 190

pctau2012001598-seql

Asp Ala Arg Pro Gly Ser Phe Thr Leu Leu Leu Arg Asn Arg Lys Gly
195 200 205

Phe Val Arg Leu Ala Leu Thr His Gly Ala Pro Leu Val Pro Ile Phe
210 215 220

Ser Phe Gly Glu Asn Asp Leu Phe Asp Gln Ile Pro Asn Ser Ser Gly
225 230 235 240

Ser Trp Leu Arg Tyr Ile Gln Asn Arg Leu Gln Lys Ile Met Gly Ile
245 250 255

Ser Leu Pro Leu Phe His Gly Arg Gly Val Phe Gln Tyr Ser Phe Gly
260 265 270

Leu Ile Pro Tyr Arg Arg Pro Ile Thr Thr Val Val Gly Lys Pro Ile
275 280 285

Glu Val Gln Lys Thr Leu His Pro Ser Glu Glu Glu Val Asn Gln Leu
290 295 300

His Gln Arg Tyr Ile Lys Glu Leu Cys Asn Leu Phe Glu Ala His Lys
305 310 315 320

Leu Lys Phe Asn Ile Pro Ala Asp Gln His Leu Glu Phe Cys
325 330

<210> 65
<211> 334
<212> PRT
<213> Canis familiaris

<400> 65

Met Val Lys Phe Ala Pro Leu Phe Val Pro Trp Glu Arg Arg Leu Gln
1 5 10 15

Thr Phe Met Val Leu Gln Trp Val Phe Ser Phe Leu Ala Leu Ala Gln
20 25 30

Ile Cys Thr Val Val Phe Val Gly Leu Leu Phe Thr Arg Phe Trp Val
35 40 45

Val Ser Val Leu Tyr Ala Ala Trp Trp Tyr Arg Asp Arg Asp Thr Pro
50 55 60

Arg Gln Gly Gly Arg Pro Val Gln Ala Leu Arg Arg Cys Phe Leu Trp
65 70 75 80

Lys Tyr Met Arg Asp Tyr Phe Pro Val Thr Leu Val Lys Thr Ala Glu
85 90 95

pctau2012001598-seql

Leu Asp Pro Ser Arg Asn Tyr Leu Ala Gly Phe His Pro His Gly Val
100 105 110

Leu Ala Ala Gly Ala Phe Val Asn Leu Cys Thr Glu Ser Thr Gly Phe
115 120 125

Pro Leu Leu Phe Pro Gly Ile Arg Ser His Leu Met Met Leu Thr Leu
130 135 140

Trp Phe Arg Val Pro Phe Phe Arg Asp Tyr Ile Met Ser Gly Gly Leu
145 150 155 160

Val Thr Ser Asp Lys Glu Ser Ala Ala His Ile Leu Ser Arg Lys Glu
165 170 175

Gly Gly Asn Leu Leu Ala Ile Ile Val Gly Gly Ala Gln Glu Ala Leu
180 185 190

Asn Ala Arg Pro Gly Asp Ser Thr Leu Leu Leu Arg Asn Arg Lys Gly
195 200 205

Phe Ile Arg Leu Ala Leu Met His Gly Ala Ala Leu Val Pro Ile Phe
210 215 220

Ser Phe Gly Glu Asn Glu Leu Phe Asp Gln Val Glu Asn Ser Pro Gly
225 230 235 240

Ser Trp Leu Arg Arg Ile Gln Asn Arg Leu Gln Lys Ile Met Gly Ile
245 250 255

Ser Leu Pro Leu Phe His Gly Arg Gly Val Phe Gln Tyr Ser Phe Gly
260 265 270

Phe Ile Pro Tyr Arg Gln Pro Ile Thr Thr Val Val Gly Lys Pro Ile
275 280 285

Glu Val Gln Lys Thr Leu Cys Pro Ser Lys Glu Glu Val Asp Lys Leu
290 295 300

His Gln Arg Tyr Val Lys Glu Leu Cys Glu Leu Phe Glu Thr His Lys
305 310 315 320

Leu Lys Tyr Asn Val Pro Val Asp Gln His Leu Glu Phe Cys
325 330

<210> 66

<211> 332

<212> PRT

<213> Bos taurus

<400> 66

Met Val Glu Phe Ala Pro Leu Phe Val Pro Leu Glu Arg Arg Leu Gln

pctau2012001598-seql

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

275

280

285

Pro Lys Ile Pro His Pro Ser Gln Glu Glu Val Asp Arg Leu His Gln
 290 295 300

His Tyr Leu Lys Glu Leu Ser Asn Leu Phe Glu Thr His Lys Leu Lys
 305 310 315 320

Tyr Asn Ile Pro Lys Asp Gln His Leu Glu Phe Cys
 325 330

<210> 67

<211> 334

<212> PRT

<213> Rattus norvegicus

<400> 67

Met Val Glu Phe Ala Pro Leu Leu Val Pro Trp Glu Arg Arg Leu Gln
 1 5 10 15

Thr Phe Ala Val Leu Gln Trp Val Phe Ser Phe Leu Ala Leu Ala Gln
 20 25 30

Leu Cys Ile Phe Ile Phe Ile Gly Leu Leu Phe Thr Arg Phe Trp Leu
 35 40 45

Phe Ser Val Leu Tyr Ala Thr Trp Trp Tyr Leu Asp Trp Asp Arg Pro
 50 55 60

Arg Gln Gly Gly Arg Pro Ile Gln Phe Phe Arg Arg Met Ala Ile Trp
 65 70 75 80

Lys Tyr Met Lys Asp Phe Phe Pro Val Ser Leu Val Lys Thr Ala Glu
 85 90 95

Leu Asp Pro Ser Arg Asn Tyr Ile Ala Gly Phe His Pro His Gly Val
 100 105 110

Leu Ala Ala Gly Ala Phe Leu Asn Leu Cys Thr Glu Ser Thr Gly Phe
 115 120 125

Thr Ser Leu Phe Pro Gly Ile Arg Ser Tyr Leu Met Met Leu Thr Val
 130 135 140

Trp Phe Arg Ala Pro Ile Phe Arg Asp Tyr Ile Met Ser Gly Gly Leu
 145 150 155 160

Val Ser Ser Glu Lys Val Ser Ala Asp His Ile Leu Ser Arg Lys Gly
 165 170 175

Gly Gly Asn Leu Leu Ala Ile Ile Val Gly Gly Ala Gln Glu Ala Leu
 180 185 190

pctau2012001598-seql

Asp Ala Arg Pro Gly Ala Tyr Arg Leu Leu Leu Lys Asn Arg Lys Gly
195 200 205

Phe Ile Arg Leu Ala Leu Thr His Gly Ala Ala Leu Val Pro Ile Phe
210 215 220

Ser Phe Gly Glu Asn Asn Leu Phe Asn Gln Val Glu Asn Thr Pro Gly
225 230 235 240

Thr Trp Leu Arg Trp Ile Gln Asn Trp Leu Gln Lys Ile Met Gly Ile
245 250 255

Ser Leu Pro Leu Phe His Gly Arg Gly Val Phe Gln Tyr Ser Phe Gly
260 265 270

Leu Val Pro Phe Arg Gln Pro Ile Thr Thr Val Val Gly Lys Pro Ile
275 280 285

Glu Val Gln Met Ile Pro His Pro Ser Glu Glu Glu Val Asn Arg Leu
290 295 300

His Gln Leu Tyr Ile Lys Glu Leu Cys Lys Leu Phe Glu Glu His Lys
305 310 315 320

Leu Lys Phe Asn Val Pro Glu Asp Gln His Leu Glu Phe Cys
325 330

<210> 68
<211> 351
<212> PRT
<213> Gallus gallus

<400> 68

Met Glu Arg Ala Tyr Lys Ile Cys Asp Asp Thr Thr Ser Lys Arg Leu
1 5 10 15

Lys Asn Leu Phe Trp Ala Val Gln Asn Gly Ser Pro Gln Val Ser Arg
20 25 30

Ser Ser Pro Thr Thr Leu Thr Ala His Arg Leu Ala Phe Leu Pro Ala
35 40 45

Gln Cys Cys Thr Ala Ala Phe Ile Gly Leu Leu Phe Thr Arg Phe Trp
50 55 60

Leu Leu Ser Val Leu Tyr Ala Ala Trp Trp Phe Val Asp Arg Glu Ala
65 70 75 80

Pro Leu Arg Gly Gly Arg Arg Ile His Met Val Arg Asn Ser Ala Val
85 90 95

pctau2012001598-seql

Trp Arg His Met Arg Asp Phe Phe Pro Val Thr Leu Val Lys Thr Ala
100 105 110

Glu Leu Asp Pro Arg Gln Asn Tyr Leu Val Gly Phe His Pro His Gly
115 120 125

Val Leu Ala Val Gly Ala Phe Ile Asn Phe Gly Thr Glu Ala Thr Gly
130 135 140

Phe Ser Thr Ile Phe Pro Gly Ile Thr Pro His Leu Met Met Leu Ser
145 150 155 160

Leu Trp Phe Arg Val Pro Phe Leu Arg Asp Tyr Leu Met Ser Gly Gly
165 170 175

Leu Val Ser Ser Asp Lys Glu Ser Ala Tyr His Val Leu Gln Arg Pro
180 185 190

Glu Gly Gly Asn Leu Leu Ala Ile Ile Val Gly Gly Ala Gln Glu Ala
195 200 205

Leu Asp Ala Arg Pro Gly Ser Cys Thr Leu Leu Leu Lys Asn Arg Lys
210 215 220

Gly Phe Val Arg Val Ala Ile Glu Gln Gly Thr Pro Leu Val Pro Ala
225 230 235 240

Phe Ser Phe Gly Glu Asn Glu Leu Phe Asp Gln Val Ser Asn Pro Lys
245 250 255

Gly Ser Trp Leu Arg Trp Ile Gln His Arg Leu Gln Gln Ile Met Gly
260 265 270

Ile Ser Leu Pro Leu Phe His Ala Arg Gly Ile Phe Gln Tyr Ser Phe
275 280 285

Gly Leu Val Pro Tyr Arg Arg Pro Ile Asn Thr Val Ile Gly Lys Pro
290 295 300

Ile Pro Val Leu Lys Lys His Lys Pro Thr Glu Glu Glu Val Asp Arg
305 310 315 320

Val His Lys Lys Tyr Leu Glu Glu Leu Ser Lys Leu Phe Glu Glu His
325 330 335

Lys Ala Lys Tyr Asn Val Pro Glu Asp Ser His Leu Glu Phe Ile
340 345 350

<210> 69
<211> 252
<212> PRT

<213> Danio rerio

<400> 69

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

pctau2012001598-seql

<210> 70

<211> 352

<212> PRT

<213> Drosophila melanogaster

<400> 70

Met Lys Ile Glu Trp Ala Pro Leu Arg Val Pro Leu Glu Arg Arg Leu
1 5 10 15

Gln Ile Leu Val Thr Ala Phe Phe Thr Ser Met Leu Leu Ile Leu Leu
20 25 30

Ser Val Ser Phe Leu Leu Val Ala Gly Ser Leu Ile Tyr Gly Gly Leu
35 40 45

Leu Val Arg Ser Leu Met Val Thr Tyr Leu Ala Tyr Val Phe Val His
50 55 60

His Lys Lys Thr Gln Ser Val Val Asp Gly Asn Gly Trp Met Ile Thr
65 70 75 80

Arg Thr Asn Leu Leu His Arg His Tyr Arg Asp Tyr Phe Pro Val Glu
85 90 95

Leu Val Lys Thr Ala Glu Leu Pro Ala Thr Lys Asn Tyr Ile Leu Ala
100 105 110

Ser Phe Pro His Gly Ile Leu Gly Thr Gly Ile Gly Ile Asn Met Gly
115 120 125

Leu Glu Ile Ser Lys Trp Leu Glu Leu Phe Pro Gln Val Arg Pro Lys
130 135 140

Leu Gly Thr Leu Asp Gln His Phe His Val Pro Phe Met Arg Glu Val
145 150 155 160

Leu Arg Cys Trp Gly Leu Val Ser Val Ser Lys Glu Ala Leu Ile Arg
165 170 175

Met Leu Ser Lys Ser Asn Asp Pro Lys His Lys Asp Asn Arg Asp Gly
180 185 190

Phe Thr Ser Asn Ala Val Ala Ile Leu Val Gly Gly Ala Gln Glu Ala
195 200 205

Met Asp Ser His Pro Gly Gln Tyr Ile Leu Thr Leu Lys Asn Arg Lys
210 215 220

Gly Phe Val Arg Met Ala Ile Arg Thr Gly Ser Ser Ile Val Pro Ser
225 230 235 240

Phe Ser Phe Gly Glu Val Asp Ile Phe Asp Gln Val Ala Asn Pro Pro
Page 69

pctau2012001598-seql

245

250

255

Asn Ser Leu Leu Arg Arg Phe Gln Asp Phe Val Lys Lys Leu Thr Gly
 260 265 270

Val Ser Pro Leu Ile Pro Val Gly Arg Gly Phe Phe Asn Tyr Thr Phe
 275 280 285

Gly Phe Leu Pro Phe Arg Arg Arg Ile Val Gln Val Val Gly Ala Pro
 290 295 300

Ile Asp Val Val Lys Asn Glu His Pro Asp Ser Glu Tyr Val Asp Lys
 305 310 315 320

Val His Gly Gln Val Ile Glu Ser Leu Glu Lys Leu Phe Asp Gln Tyr
 325 330 335

Lys Asp Lys Tyr Leu Glu Asn Ser Lys Ser Ala Thr Leu Val Val His
 340 345 350

<210> 71

<211> 352

<212> PRT

<213> Drosophila melanogaster

<400> 71

Met Lys Ile Glu Trp Ala Pro Lys Gly Val Pro Met Glu Arg Arg Arg
 1 5 10 15

Gln Thr Phe Ala Met Ala Phe Leu Ile Leu Ser Phe Met Ile Leu Ser
 20 25 30

Phe Gly Ser Tyr Phe Phe Val Ala Ala Val Leu Phe Tyr Gly Ser Leu
 35 40 45

Leu Trp Arg Thr Ile Met Val Ile Tyr Leu Val Tyr Val Tyr Ala Asn
 50 55 60

His Lys Arg Thr His Ser Ile Met Asp Gly Asn Gly Trp Lys Ile Asn
 65 70 75 80

Arg Asn Asn Trp Leu Phe Arg His Tyr Arg Asp Tyr Phe Pro Val Gln
 85 90 95

Leu Val Lys Thr Ala Glu Leu Pro Pro Asn Lys Asn Tyr Ile Leu Ala
 100 105 110

Ser Phe Pro His Gly Ile Leu Gly Thr Gly Ile Ser Ile Asn Met Gly
 115 120 125

Leu Asp Ile Ser Lys Trp Leu Gln Leu Phe Pro Gln Val Arg Pro Lys
 130 135 140

pctau2012001598-seql

Val Ala Thr Leu Asp Gln Asn Phe Leu Thr Pro Ile Val Arg Gly Leu
145 150 155 160

Leu Arg Ser Trp Gly Leu Val Ser Val Ser Lys Glu Ala Leu Val Tyr
165 170 175

Leu Leu Thr Lys Ser Asn Asp Pro Lys His Lys Asp Asn Arg Asp Gly
180 185 190

Phe Thr Ser Asn Ala Val Ala Ile Leu Val Gly Gly Ala Gln Glu Ala
195 200 205

Leu Asp Ser His Pro Gly Lys Tyr Ile Leu Thr Leu Lys Asn Arg Lys
210 215 220

Gly Phe Val Lys Met Ala Ile Arg Thr Gly Ser Ser Ile Val Pro Thr
225 230 235 240

Phe Ser Phe Gly Glu Val Asp Ile Leu Asp Gln Val Ala Asn Pro Pro
245 250 255

Asn Ser Arg Val Arg Arg Phe Gln Asp Phe Val Lys Arg Ile Thr Gly
260 265 270

Ile Ser Pro Leu Ile Pro Val Gly Arg Gly Ile Phe Asn Tyr Ser Phe
275 280 285

Gly Phe Leu Pro Asn Arg Arg Arg Ile Val Gln Val Val Gly Ala Pro
290 295 300

Ile Asp Val Val Gln Ser Asp Gln Pro Asp Ala Ala Tyr Val Asp Lys
305 310 315 320

Ile His Lys Gln Val Ile Asp Asp Leu Glu Lys Met Phe Ala Lys Tyr
325 330 335

Lys Asp Gln Tyr Ile Pro Asn Ser Lys Gln Asp Lys Leu Ile Ile His
340 345 350

<210> 72

<211> 349

<212> PRT

<213> Drosophila melanogaster

<400> 72

Met Thr Ile Glu Trp Ala Pro Leu Arg Val Pro Leu Glu Arg Arg Leu
1 5 10 15

Gln Thr Leu Val Thr Ser Phe Phe Thr Tyr Thr Phe Phe Thr Leu Pro
20 25 30

pctau2012001598-seql

```

Ile Ser Ser Cys Leu Ala Val Ala Ile Leu Leu Tyr Tyr Gly Glu Met
    35          40          45

Phe Val Arg Ser Leu Leu Leu Ile Tyr Phe Val Lys Ile Tyr Leu Asp
    50          55          60

Tyr Lys Arg Asn Tyr Gly Ile Met Glu Gly Asn Gly Trp Leu Phe Tyr
    65          70          75          80

Arg Ser Asn Trp Arg Tyr Arg Asn Tyr Phe Pro Val Glu Leu Val Lys
    85          90          95

Thr Ala Glu Leu Pro Pro Asn Arg Asn Tyr Ile Val Ala Ser Phe Pro
    100          105          110

His Gly Ile Leu Gly Thr Gly Thr Cys Ile Asn Met Ser Leu Asp Ile
    115          120          125

Asp Asn Trp Leu Ser Leu Tyr Pro His Val Arg Pro Lys Ile Ala Thr
    130          135          140

Leu Asp His His Phe Lys Thr Pro Phe Leu Arg Asp Ile Leu Arg Trp
    145          150          155          160

Trp Gly Met Val Ser Val Ser Lys Glu Ser Leu Ser Tyr Leu Leu Ser
    165          170          175

Lys Ser Asn Asp Pro Met His Lys Asp Asn Arg Asp Gly Phe Thr Ser
    180          185          190

Asn Ala Val Ala Val Leu Val Gly Gly Ala Lys Glu Ala Met Asp Ser
    195          200          205

His Pro Gly Gln Tyr Ile Leu Thr Leu Lys Asp Arg Lys Gly Phe Val
    210          215          220

Lys Met Ala Val Arg Thr Gly Ser Ser Ile Val Pro Ser Leu Ser Phe
    225          230          235          240

Gly Glu Val Asp Ile Phe Asp Gln Val Ala Asn Pro Pro Asp Ser Ser
    245          250          255

Leu Arg Arg Phe Gln Asn Val Val Lys Lys Phe Thr Gly Ile Ser Pro
    260          265          270

Leu Leu Pro Lys Gly Arg Gly Ile Phe Asn Tyr Asn Tyr Gly Ile Leu
    275          280          285

Pro His Arg Arg Arg Ile Val Gln Val Val Gly Ser Pro Ile Asp Val
    290          295          300

```

pctau2012001598-seql

Glu Arg Cys Glu Thr Pro Asp Pro Glu Tyr Val Asp Lys Ile His Gly
305 310 315 320

Gln Val Ile Asp Ala Leu Ala Arg Met Phe Asp Glu Tyr Lys Glu Lys
325 330 335

Tyr Thr Pro Asn Ser Lys His Ile Lys Leu Ile Ile Gln
340 345

<210> 73

<211> 353

<212> PRT

<213> Ashbya gossypii

<400> 73

Met Gly Lys Ile Glu Trp Ala Pro Leu Asn Val Pro Met Arg Arg Arg
1 5 10 15

Ile Glu Thr Leu Ser Thr Ala Leu Trp Met Trp Leu Ile Leu Phe Gly
20 25 30

Glu Leu Gly Met Leu Ile Ser Tyr Phe Leu Leu Leu Ile Tyr Gly Asn
35 40 45

Leu Phe Ile Lys Ser Leu Cys Val Ile Tyr Gly Tyr Phe Ile Tyr Thr
50 55 60

Asp Arg Lys Val Thr Leu Asn Gly Gly Arg Gly Gln Gly Val Thr Trp
65 70 75 80

Trp Arg Glu Leu Phe Trp Trp Lys Leu Tyr Gln Ser Tyr Phe Pro Ala
85 90 95

Lys Leu His Lys Thr Val Asp Leu Asp Pro Asn Arg Asn Tyr Leu Phe
100 105 110

Ala Ala Phe Pro His Gly Val Leu Gly Leu Gly Ala Phe Ile Asn Phe
115 120 125

Ala Thr Asn Ala Thr Gly Phe His Asp Lys Phe Pro Lys Ile Arg Ser
130 135 140

Arg Pro Val Thr Leu Asn Phe His Phe Val Ile Pro Phe Phe Arg Glu
145 150 155 160

Leu Leu Leu Ser Trp Gly Leu Val Ser Ala Asn Pro Asn Ser Ile Leu
165 170 175

Ser Leu Leu Lys Ala Pro Asn Lys Pro Asp His Pro Leu Asn Asp Asp
180 185 190

pctau2012001598-seql

Gly Tyr Thr Ala Asn Ala Val Val Ile Val Val Gly Gly Ala Ala Glu
195 200 205

Ser Leu His Cys Arg Pro Asn Asn Tyr Thr Leu Val Leu Arg Lys Arg
210 215 220

Lys Gly Phe Cys Lys Leu Ala Ile Lys Ala Gly Thr Pro Leu Val Pro
225 230 235 240

Val Met Thr Phe Gly Glu Val Asp Leu Phe Asp Gln Pro Pro Asn Pro
245 250 255

Pro Gly Ser Arg Leu Arg Arg Phe Gln Glu Phe Val Lys Asn Thr Thr
260 265 270

Gly Ile Ala Pro Ala Ala Phe Val Gly Arg Gly Phe Phe Gln Tyr Ser
275 280 285

Tyr Gly Leu Ile Pro Arg Arg Lys Pro Leu Asn Thr Val Val Gly Ala
290 295 300

Pro Val Glu Val Thr Gln Ile Asp Asn Pro Thr Gln Glu Gln Val Asp
305 310 315 320

Glu Val His Glu Arg Phe Cys Arg Ala Leu Asp Asn Leu Phe Glu Thr
325 330 335

Asn Lys Ser Arg Phe Ile Ala Asp Tyr Lys Asn Val Lys Leu Val Met
340 345 350

Glu

<210> 74

<211> 353

<212> PRT

<213> Drosophila melanogaster

<400> 74

Met Gly Lys Ile Glu Trp Ala Pro Leu Asn Val Pro Met Arg Arg Arg
1 5 10 15

Ile Glu Thr Leu Ser Thr Ala Leu Trp Met Trp Leu Ile Leu Phe Gly
20 25 30

Glu Leu Gly Met Leu Ile Ser Tyr Phe Leu Leu Leu Ile Tyr Gly Asn
35 40 45

Leu Phe Ile Lys Ser Leu Cys Val Ile Tyr Gly Tyr Phe Ile Tyr Thr
50 55 60

Asp Arg Lys Val Thr Leu Asn Gly Gly Arg Gly Gln Gly Val Thr Trp
Page 74

pctau2012001598-seql

```

65              70              75              80
Trp Arg Glu Leu Phe Trp Trp Lys Leu Tyr Gln Ser Tyr Phe Pro Ala
85              90              95
Lys Leu His Lys Thr Val Asp Leu Asp Pro Asn Arg Asn Tyr Leu Phe
100             105             110
Ala Ala Phe Pro His Gly Val Leu Gly Leu Gly Ala Phe Ile Asn Phe
115             120             125
Ala Thr Asn Ala Thr Gly Phe His Asp Lys Phe Pro Lys Ile Arg Ser
130             135             140
Arg Pro Val Thr Leu Asn Phe His Phe Val Ile Pro Phe Phe Arg Glu
145             150             155
Leu Leu Leu Ser Trp Gly Leu Val Ser Ala Asn Pro Asn Ser Ile Leu
165             170             175
Ser Leu Leu Lys Ala Pro Asn Lys Pro Asp His Pro Leu Asn Asp Asp
180             185             190
Gly Tyr Thr Ala Asn Ala Val Val Ile Val Val Gly Gly Ala Ala Glu
195             200             205
Ser Leu His Cys Arg Pro Asn Asn Tyr Thr Leu Val Leu Arg Lys Arg
210             215             220
Lys Gly Phe Cys Lys Leu Ala Ile Lys Ala Gly Thr Pro Leu Val Pro
225             230             235
Val Met Thr Phe Gly Glu Val Asp Leu Phe Asp Gln Pro Pro Asn Pro
245             250             255
Pro Gly Ser Arg Leu Arg Arg Phe Gln Glu Phe Val Lys Asn Thr Thr
260             265             270
Gly Ile Ala Pro Ala Ala Phe Val Gly Arg Gly Phe Phe Gln Tyr Ser
275             280             285
Tyr Gly Leu Ile Pro Arg Arg Lys Pro Leu Asn Thr Val Val Gly Ala
290             295             300
Pro Val Glu Val Thr Gln Ile Asp Asn Pro Thr Gln Glu Gln Val Asp
305             310             315
Glu Val His Glu Arg Phe Cys Arg Ala Leu Asp Asn Leu Phe Glu Thr
325             330             335
Asn Lys Ser Arg Phe Ile Ala Asp Tyr Lys Asn Val Lys Leu Val Met

```


340

345

350

Glu

<210> 75

<211> 338

<212> PRT

<213> Tribolium castaneum

<400> 75

Met Lys Ile Leu Gly Ile Lys Phe Ala Pro Leu His Ile Pro Leu Glu
 1 5 10 15

Arg Arg Leu Gln Thr Leu Ala Ala Gly Cys Trp Phe Thr Thr Leu Ala
 20 25 30

Phe Gly Thr Phe Ile Gly Thr Phe Ile Trp Val Tyr Val Phe Phe Thr
 35 40 45

Arg Phe Trp Tyr Leu Ser Val Leu Tyr Ala Thr Ile Ile Tyr Leu Glu
 50 55 60

Lys Ser Lys Cys Glu Lys Gly Gly Arg Pro Ile Glu Trp Ile Arg His
 65 70 75 80

Trp Gly Trp Trp Tyr Tyr Leu Lys Asn Tyr Phe Pro Cys Lys Leu Asp
 85 90 95

Phe Val Pro Gly Leu Thr Phe Asp Pro Lys Arg Asn Tyr Leu Phe Ala
 100 105 110

Cys Tyr Pro His Gly Ile Leu Pro Ala Gly Pro Phe Asn Thr Ile Gly
 115 120 125

Ser Pro Tyr Ser Glu Phe Ser Lys Leu Phe Pro Lys Phe Arg Val Arg
 130 135 140

Leu Val Ile Leu His Gln His Phe Phe Ile Pro Phe Leu Arg Glu Ile
 145 150 155 160

Ala Tyr Gly Thr Gly Gly Ile Ser Ala Ser Ala Lys Ser Leu Asn His
 165 170 175

Val Leu Gly Ser Pro Glu Gly Gly Tyr Ile Ala Val Leu Met Pro Gly
 180 185 190

Gly Ala Val Glu Ala Tyr Asn Ser Arg Pro Gly Gln Tyr Arg Ile Ile
 195 200 205

Leu Lys Asn Arg Lys Gly Phe Val Lys Leu Ala Leu Arg Asn Gly Ser
 210 215 220

pctau2012001598-seql

Pro Leu Val Pro Val Ile Ser Phe Gly Glu Pro Glu Leu Phe Asp Gln
225 230 235 240

Val Glu Gly Arg Thr Leu Arg Lys Ile Gln Glu Ser Ile Arg Lys Tyr
245 250 255

Leu Gly Leu Ala Pro Val Ile Phe Ser Gly Arg Gly Phe Phe Gln Tyr
260 265 270

Ser Phe Gly Val Ile Pro Gln Arg Arg Pro Ile Thr Thr Val Val Gly
275 280 285

His Pro Ile Glu Val Thr Lys Ile Glu Lys Pro Thr Asn Glu Glu Val
290 295 300

Asp Glu Leu His Lys Lys Tyr Met Gln Glu Leu Glu Asn Leu Phe Glu
305 310 315 320

Glu Tyr Lys Phe Lys Tyr Leu Glu Asn Pro Lys Asp Ile His Leu Glu
325 330 335

Phe Glu

<210> 76
<211> 341
<212> PRT
<213> Homo sapiens

<400> 76

Met Gly Val Ala Thr Thr Leu Gln Pro Pro Thr Thr Ser Lys Thr Leu
1 5 10 15

Gln Lys Gln His Leu Glu Ala Val Gly Ala Tyr Gln Tyr Val Leu Thr
20 25 30

Phe Leu Phe Met Gly Pro Phe Phe Ser Leu Leu Val Phe Val Leu Leu
35 40 45

Phe Thr Ser Leu Trp Pro Phe Ser Val Phe Tyr Leu Val Trp Leu Tyr
50 55 60

Val Asp Trp Asp Thr Pro Asn Gln Gly Gly Arg Arg Ser Glu Trp Ile
65 70 75 80

Arg Asn Arg Ala Ile Trp Arg Gln Leu Arg Asp Tyr Tyr Pro Val Lys
85 90 95

Leu Val Lys Thr Ala Glu Leu Pro Pro Asp Arg Asn Tyr Val Leu Gly
100 105 110

pctau2012001598-seql

Ala His Pro His Gly Ile Met Cys Thr Gly Phe Leu Cys Asn Phe Ser
115 120 125

Thr Glu Ser Asn Gly Phe Ser Gln Leu Phe Pro Gly Leu Arg Pro Trp
130 135 140

Leu Ala Val Leu Ala Gly Leu Phe Tyr Leu Pro Val Tyr Arg Asp Tyr
145 150 155 160

Ile Met Ser Phe Gly Leu Cys Pro Val Ser Arg Gln Ser Leu Asp Phe
165 170 175

Ile Leu Ser Gln Pro Gln Leu Gly Gln Ala Val Val Ile Met Val Gly
180 185 190

Gly Ala His Glu Ala Leu Tyr Ser Val Pro Gly Glu His Cys Leu Thr
195 200 205

Leu Gln Lys Arg Lys Gly Phe Val Arg Leu Ala Leu Arg His Gly Ala
210 215 220

Ser Leu Val Pro Val Tyr Ser Phe Gly Glu Asn Asp Ile Phe Arg Leu
225 230 235 240

Lys Ala Phe Ala Thr Gly Ser Trp Gln His Trp Cys Gln Leu Thr Phe
245 250 255

Lys Lys Leu Met Gly Phe Ser Pro Cys Ile Phe Trp Gly Arg Gly Leu
260 265 270

Phe Ser Ala Thr Ser Trp Gly Leu Leu Pro Phe Ala Val Pro Ile Thr
275 280 285

Thr Val Val Gly Arg Pro Ile Pro Val Pro Gln Arg Leu His Pro Thr
290 295 300

Glu Glu Glu Val Asn His Tyr His Ala Leu Tyr Met Thr Ala Leu Glu
305 310 315 320

Gln Leu Phe Glu Glu His Lys Glu Ser Cys Gly Val Pro Ala Ser Thr
325 330 335

Cys Leu Thr Phe Ile
340

<210> 77
<211> 281
<212> PRT
<213> Pan troglodytes
<400> 77

pctau2012001598-seql

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

pctau2012001598-seql

His Leu Gly Leu Ala Ala Ala Phe Arg
275 280

<210> 78
<211> 344
<212> PRT
<213> Pan troglodytes

<400> 78

Met Gly Val Ala Thr Thr Leu Leu Pro Pro Thr Thr Ser Lys Thr Leu
1 5 10 15

Gln Lys Gln His Leu Glu Ala Val Gly Ala Tyr Gln Tyr Val Leu Thr
20 25 30

Phe Leu Phe Met Gly Pro Phe Phe Ser Leu Leu Val Phe Val Leu Leu
35 40 45

Phe Thr Ser Leu Trp Pro Phe Ser Val Phe Tyr Leu Val Trp Leu Tyr
50 55 60

Val Asp Trp Asp Thr Pro Asn Gln Gly Gly Arg Arg Ser Glu Trp Ile
65 70 75 80

Arg Asn Trp Ala Ile Trp Arg Gln Leu Arg Asp Tyr Tyr Pro Val Lys
85 90 95

Leu Val Lys Thr Ala Glu Leu Pro Pro Asp Arg Asn Tyr Val Leu Gly
100 105 110

Ala His Pro His Gly Ile Met Cys Thr Gly Phe Leu Cys Asn Phe Ser
115 120 125

Thr Glu Ser Asn Gly Phe Ser Gln Leu Phe Pro Gly Leu Arg Pro Trp
130 135 140

Leu Ala Val Leu Ala Gly Leu Phe Tyr Leu Pro Val Tyr Arg Asp Tyr
145 150 155 160

Ile Met Ser Phe Gly Leu Cys Pro Val Ser Arg Gln Ser Leu Asp Phe
165 170 175

Ile Leu Ser Gln Pro Gln Leu Gly Gln Ala Val Val Ile Met Val Gly
180 185 190

Gly Ala His Glu Ala Leu Tyr Ser Val Pro Gly Glu His Cys Leu Thr
195 200 205

Leu Gln Lys Arg Lys Gly Phe Val Arg Leu Ala Leu Arg His Gly Ala
210 215 220

Ser Leu Val Pro Val Tyr Ser Phe Gly Glu Asn Asp Ile Phe Arg Leu
Page 80

pctau2012001598-seql

225 230 235 240

Lys Ala Phe Ala Thr Gly Ser Trp Gln His Trp Cys Gln Leu Thr Phe
245 250 255

Lys Lys Phe Met Gly Phe Ser Pro Cys Ile Phe Trp Gly Arg Gly Leu
260 265 270

Phe Ser Ala Thr Ser Trp Gly Leu Leu Pro Phe Ala Val Pro Ile Thr
275 280 285

Thr Val Gly Glu Cys Pro Pro Pro Gly Gly Arg Pro Pro Ala Ala Ala
290 295 300

Trp Ala Ser Gly Ile Pro Arg Pro Pro Val Ser Leu Ser Leu Gln Trp
305 310 315 320

Ala Ala Pro Ser Leu Ser Pro Ser Ala Ser Thr Pro Pro Arg Arg Lys
325 330 335

Ser Ile Thr Ile Thr Pro Ser Thr
340

<210> 79

<211> 341

<212> PRT

<213> Pan troglodytes

<400> 79

Met Gly Val Ala Thr Thr Leu Leu Pro Pro Thr Thr Ser Lys Thr Leu
1 5 10 15

Gln Lys Gln His Leu Glu Ala Val Gly Ala Tyr Gln Tyr Val Leu Thr
20 25 30

Phe Leu Phe Met Gly Pro Phe Phe Ser Leu Leu Val Phe Val Leu Leu
35 40 45

Phe Thr Ser Leu Trp Pro Phe Ser Val Phe Tyr Leu Val Trp Leu Tyr
50 55 60

Val Asp Trp Asp Thr Pro Asn Gln Gly Gly Arg Arg Ser Glu Trp Ile
65 70 75 80

Arg Asn Trp Ala Ile Trp Arg Gln Leu Arg Asp Tyr Tyr Pro Val Lys
85 90 95

Leu Val Lys Thr Ala Glu Leu Pro Pro Asp Arg Asn Tyr Val Leu Gly
100 105 110

Ala His Pro His Gly Ile Met Cys Thr Gly Phe Leu Cys Asn Phe Ser
115 120 125

pctau2012001598-seql

Thr Glu Ser Asn Gly Phe Ser Gln Leu Phe Pro Gly Leu Arg Pro Trp
130 135 140

Leu Ala Val Leu Ala Gly Leu Phe Tyr Leu Pro Val Tyr Arg Asp Tyr
145 150 155 160

Ile Met Ser Phe Gly Leu Cys Pro Val Ser Arg Gln Ser Leu Asp Phe
165 170 175

Ile Leu Ser Gln Pro Gln Leu Gly Gln Ala Val Val Ile Met Val Gly
180 185 190

Gly Ala His Glu Ala Leu Tyr Ser Val Pro Gly Glu His Cys Leu Thr
195 200 205

Leu Gln Lys Arg Lys Gly Phe Val Arg Leu Ala Leu Arg His Gly Ala
210 215 220

Ser Leu Val Pro Val Tyr Ser Phe Gly Glu Asn Asp Ile Phe Arg Leu
225 230 235 240

Lys Ala Phe Ala Thr Gly Ser Trp Gln His Trp Cys Gln Leu Thr Phe
245 250 255

Lys Lys Phe Met Gly Phe Ser Pro Cys Ile Phe Trp Gly Arg Gly Leu
260 265 270

Phe Ser Ala Thr Ser Trp Gly Leu Leu Pro Phe Ala Val Pro Ile Thr
275 280 285

Thr Val Val Gly Arg Pro Ile Pro Val Pro Gln Arg Leu His Pro Thr
290 295 300

Glu Glu Glu Val Asn His Tyr His Ala Leu Tyr Met Met Ala Leu Gln
305 310 315 320

Gln Leu Phe Glu Glu His Lys Glu Ser Cys Gly Val Pro Ala Ser Thr
325 330 335

Cys Leu Thr Phe Ile
340

<210> 80
<211> 582
<212> PRT
<213> Canis familiaris

<400> 80

Met Thr Phe Pro Asp Pro Lys Thr Gly Gln Cys Cys Leu Ser Pro Val
1 5 10 15

pctau2012001598-seql

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

pctau2012001598-seql

Leu Phe Thr Ser Phe Trp Ser Ile Ser Val Leu Tyr Leu Val Trp Leu
 290 295 300
 Phe Leu Asp Trp Asp Thr Pro Asn Gln Gly Gly Arg Cys Phe Glu Cys
 305 310 315 320
 Asn Arg Ser Cys Thr Ile Trp Lys His Leu Lys Asp Tyr Phe Pro Ile
 325 330 335
 Lys Leu Val Lys Thr Ala Glu Leu Pro Pro Asp Arg Asn Tyr Val Val
 340 345 350
 Gly Cys His Pro His Gly Ile Met Cys Met Gly Thr Phe Cys Asn Phe
 355 360 365
 Phe Thr Glu Ala Asn Asn Phe Ser Lys Gln Phe Pro Gly Ile Gln Thr
 370 375 380
 Ser Pro Val Thr Leu Ala Phe Leu Leu His Leu Pro Val Tyr Arg Asp
 385 390 395 400
 Tyr Leu Met Tyr Leu Gly Leu Cys Ser Val Asn Arg Arg Ser Leu Asp
 405 410 415
 Phe Ile Leu Ser Lys Pro Gln Pro Gly Gln Ala Val Ala Ile Val Val
 420 425 430
 Gly Gly Ala His Glu Ser Leu Phe Ala Ile Pro Gly Met His Cys Leu
 435 440 445
 Val Leu Arg Asn Arg Lys Gly Phe Val Arg Leu Ala Leu Arg His Gly
 450 455 460
 Ala Ser Leu Val Pro Val Tyr Ser Phe Gly Glu Asn Asp Ile Phe Asn
 465 470 475 480
 Phe Lys Ala Phe Pro Thr Asn Ser Trp Gln Tyr Leu Cys Gln Ile Thr
 485 490 495
 Ile Lys Lys Ile Met Lys Phe Ser Pro Cys Ile Phe Trp Gly Arg Gly
 500 505 510
 Leu Phe Ser Ala Asp Ser Trp Gly Leu Leu Pro Phe Ala Lys Pro Ile
 515 520 525
 Thr Thr Val Val Gly Arg Pro Ile Pro Val Pro Gln Arg Leu Asn Pro
 530 535 540
 Thr Glu Glu Glu Val Asp His Tyr His Met Leu Tyr Met Glu Ala Leu
 545 550 555 560

pctau2012001598-seql

Glu Gln Leu Phe Glu Glu His Lys Glu Ser Cys Gly Val Pro Ala Ser
565 570 575

Thr His Leu Ile Phe Lys
580

<210> 81
<211> 406
<212> PRT
<213> Bos taurus

<400> 81

Met Ile Ser Glu Ser Arg Leu Glu Thr Ser Gly Ala Leu Met Arg Glu
1 5 10 15

Ser Pro His Ala His Ser Pro Leu Gln Ala Ile Leu Gly Gly Ser Gln
20 25 30

Ala Leu Gly His Ser Asp Gln Lys Asp Phe Leu Arg Ser Asn Gln Ser
35 40 45

Leu His Pro Ile Phe Pro Gly Ser Glu Ala Leu Leu Ala Glu Gly Glu
50 55 60

His Leu Gly Val Ser Thr Thr Pro Pro Pro Thr Pro Ser Met Lys Thr
65 70 75 80

Leu Lys Lys Gln Trp Leu Glu Val Leu Ser Thr Tyr Gln Tyr Val Leu
85 90 95

Cys Phe Leu Phe Leu Gly Pro Phe Phe Ser Leu Val Gly Phe Phe Leu
100 105 110

Leu Phe Thr Ser Leu Trp Tyr Leu Ser Val Leu Tyr Leu Val Trp Leu
115 120 125

Phe Leu Asp Trp Asp Thr Pro Gln Gln Gly Gly Arg Arg Asn Gln Trp
130 135 140

Leu Lys Asn Cys Thr Val Trp Lys His Leu Ser Asp Tyr Phe Pro Ile
145 150 155 160

Lys Leu Val Lys Thr Val Glu Leu Pro Pro Asp Arg Asn Tyr Val Leu
165 170 175

Leu Ser His Pro His Gly Ile Met Gly Phe Gly Thr Val Cys Asn Phe
180 185 190

Ser Thr Glu Gly Thr Gly Cys Ser Gln Leu Phe Pro Gly Leu Arg Phe
195 200 205

pctau2012001598-seql

Ser Leu Ala Val Leu Asn Cys Leu Leu Tyr Val Pro Gly Cys Arg Glu
210 215 220

Tyr Ile Met Ser Cys Gly Thr Cys Ser Val Asn Arg Gln Ser Leu Asp
225 230 235 240

Tyr Val Leu Ser Gln Pro Gln Leu Gly Arg Ala Val Val Ile Met Val
245 250 255

Gly Gly Ala Asn Glu Ala Leu His Thr Val Pro Gly Glu His Cys Leu
260 265 270

Thr Leu Arg Asn Arg Lys Gly Phe Val Arg Leu Ala Leu Arg His Gly
275 280 285

Ala Ser Leu Val Pro Val Tyr Ser Phe Gly Glu Asn Asp Val Phe Arg
290 295 300

Val Lys Ala Phe Ala Pro Asp Ser Trp Gln Arg Leu Phe Gln Val Thr
305 310 315 320

Ile Lys Arg Leu Leu Ser Phe Ser Pro Cys Ile Phe Trp Gly Arg Gly
325 330 335

Leu Phe Ser Ala Lys Ser Trp Gly Leu Met Pro Leu Ala Arg Pro Ile
340 345 350

Thr Thr Val Val Gly Arg Pro Ile Pro Val Pro Gln Cys Pro Gln Pro
355 360 365

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

Glu Gln Leu Phe Glu Glu His Lys Lys Ser Cys Gly Leu Pro Ala Ser
385 390 395 400

Thr His Leu Thr Phe Ile
405

<210> 82
<211> 336
<212> PRT
<213> Danio rerio

<400> 82

Met Gly Thr Asp Asn Val Ser Glu Val Lys Glu Lys Gly Asp Arg Ser
1 5 10 15

Pro Trp Lys Asp Ile Ile Glu Asp Ile Ser Val Leu Gln Leu Val Leu
20 25 30

Ser Phe Leu Phe Leu Gly Val Ala Cys Leu Leu Leu Met Ile Tyr Leu

35

40

45

Met Phe Thr Ser Leu Trp Ile Phe Pro Thr Leu Tyr Phe Thr Trp Gln
 50 55 60
 Ile Tyr Asp Trp His Thr Pro Glu Arg Gly Gly Arg Arg Thr Lys Phe
 65 70 75 80
 Val Arg Gly Trp Glu Val Trp Lys His Leu Arg Asp Tyr Phe Pro Val
 85 90 95
 Lys Leu Val Lys Thr Ala Glu Leu Asn Pro Asn Lys Asn Tyr Ile Met
 100 105 110
 Gly Cys His Pro His Gly Ile Met Cys Phe Gly Ala Phe Ser Cys Phe
 115 120 125
 Ser Thr Asp Arg Asn Gly Phe Ala Glu Thr Phe Pro Gly Ile Arg Ser
 130 135 140
 Thr Leu Ala Ile Leu Ala Gly Leu Phe Arg Leu Pro Leu Phe Arg Glu
 145 150 155 160
 Tyr Ile Leu Ala Ala Gly Leu Leu Pro Val Ser Lys Ala Ser Leu Asp
 165 170 175
 Tyr Leu Leu Ser Gln Thr Gly Val Gly Asn Ala Val Val Ile Ile Ile
 180 185 190
 Gly Gly Ala Glu Glu Ser Leu Thr Ser Ser Thr Gly Val Asn Thr Val
 195 200 205
 Val Ile Lys His Arg Lys Gly Phe Val Arg Leu Ala Leu Glu His Gly
 210 215 220
 Ala Asp Leu Val Pro Val Tyr Ser Phe Gly Glu Asn Glu Leu Phe Pro
 225 230 235 240
 Gln Val Val Leu Ser Glu Gly Ser Val Gly Arg Arg Leu Gln Val Leu
 245 250 255
 Phe Lys Gln Ile Met Gly Phe Ala Pro Cys Ile Phe Thr Gly Gly Arg
 260 265 270
 Trp Leu Leu Leu Pro Tyr Lys Leu Pro Val Thr Thr Val Val Gly Cys
 275 280 285
 Pro Ile Asn Val Pro Leu Val Lys Ile Pro Thr Gln Glu Gln Val Asp
 290 295 300
 His Tyr His Gly Leu Tyr Met Ala Ser Leu Ala Asp Leu Phe His Lys

305 310 315 320

His Lys Thr Ser Tyr Gly Leu Ala Glu Thr His Glu Leu His Phe Ile
325 330 335

```
<210> 83
<211> 520
<212> PRT
<213> Arabidopsis thaliana
```

<400> 83

Met Ala Ile Leu Asp Ser Ala Gly Val Thr Thr Val Thr Glu Asn Gly
1 5 10 15

Gly Gly Glu Phe Val Asp Leu Asp Arg Leu Arg Arg Arg Lys Ser Arg
20 25 30

Ser Asp Ser Ser Asn Gly Leu Leu Leu Ser Gly Ser Asp Asn Asn Ser
35 40 45

Pro Ser Asp Asp Val Gly Ala Pro Ala Asp Val Arg Asp Arg Ile Asp
50 55 60

Ser Val Val Asn Asp Asp Ala Gln Gly Thr Ala Asn Leu Ala Gly Asp
65 70 75 80

Asn Asn Gly Gly Gly Asp Asn Asn Gly Gly Gly Arg Gly Gly Gly Glu
85 90 95

Gly Arg Gly Asn Ala Asp Ala Thr Phe Thr Tyr Arg Pro Ser Val Pro
100 105 110

Ala His Arg Arg Ala Arg Glu Ser Pro Leu Ser Ser Asp Ala Ile Phe
115 120 125

Lys Gln Ser His Ala Gly Leu Phe Asn Leu Cys Val Val Val Leu Ile
130 135 140

Ala Val Asn Ser Arg Leu Ile Ile Glu Asn Leu Met Lys Tyr Gly Trp
145 150 155 160

Leu Ile Arg Thr Asp Phe Trp Phe Ser Ser Arg Ser Leu Arg Asp Trp
165 170 175

Pro Leu Phe Met Cys Cys Ile Ser Leu Ser Ile Phe Pro Leu Ala Ala
180 185 190

Phe Thr Val Glu Lys Leu Val Leu Gln Lys Tyr Ile Ser Glu Pro Val
195 200 205

Val Ile Phe Leu His Ile Ile Ile Thr Met Thr Glu Val Leu Tyr Pro
210 215 220

pctau2012001598-seql

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

pctau2012001598-seql

Cys Ile Phe Gly Gln Pro Met Cys Val Leu Leu Tyr Tyr His Asp Leu
500 505 510

Met Asn Arg Lys Gly Ser Met Ser
515 520

<210> 84
<211> 1856
<212> DNA
<213> Arabidopsis thaliana

<400> 84
aaggaccttc cctctctttc tttccaaatc tctctttatc tctctttctc tgtttactaa 60
atcgccagta ttcgccgcaa gaaagatgtc tccggcgaag aagagccgaa gttttcctcc 120
gataagcgag tgtaaaagca gagagtatga ttcgatcgct gcggatctcg acgggactct 180
gcttctgtca agaagctctt tcccttactt catgctggtg gccattgaag caggtagtct 240
cttccgtgga ttgatccttc ttctttctct tcccatcgtc atcatcgctt atctcttcgt 300
ctccgaatct ctcggaatcc aaatccttat cttcatctcc ttcgccggta tcaagatcaa 360
gaacatcgaa ctgctctctc gcgccgttct tacacggttt tacgcggcgg atgtgaggaa 420
agatagtttt gaggtgtttg ataaatgtaa gaagaggaaa gtagtggtga ctgcgaatcc 480
gatagtgatg gttgagccat tcgtcaaaga ctatttggga ggagataaag ttttgggaac 540
agagattgaa gttaacccca aaacgatgaa agctactggt tttgtgaaaa agcctggtgt 600
tcttgttggg gatctcaaga gattagccat cttgaaagag ttcggcgatg actcaccgga 660
tcttggcctc ggtgaccgca cctccgatca cgatttcatg tccatttgca aggaaggtta 720
catggtgcat gagaccaa at cagccacaac agtccctata gagagtctca agaaccgcat 780
aatcttccat gatggccgtc ttgtccaacg tccgacccca ttaaaccgct taatcattta 840
cctttggctt cctttcgggt tcatgctctc tgtcttccgc gtctacttca acctcccttt 900
acccgaacgc ttcgtccgct acacttacga gatcctcggc attcacctca caatccgtgg 960
ccaccgtcct ccacctcctt cccccgaaa acctggaac ctctacgttc tcaaccaccg 1020
caccgccctt gacccatta tcatcgccat tgctctcggc cgtaagatca catgtgtcac 1080
ttatagtgtc tctcgctctt ccctgatgct ctcaccatt cctgctgttg ctctgaccgg 1140
agaccgtgtc gctgacgcag cccgatgag gcaactcctc gagaaagggtg atttggtgat 1200
ctgtcctgaa ggcacaacgt gtagagaacc atatctacta aggttcagtg ctctatttgc 1260
agagctgagc gaccggatcg tgctgtggc catgaactgc aagcaaggga tgttcaacgg 1320
gacgacagtt agagggtgta aattctggga tccgtacttc ttcttcatga acccccgacc 1380
tagctacgag gccactttct tggaccgttt gccgaagaa atgacggtaa acggtggtgg 1440
caagactcct tttgaggtgg ctaattacgt tcagaagggtg atcgggtggag ttttgggggtt 1500
cgaatgtact gaacttaca ggaaggataa gtatctgttg cttggaggca acgacggaaa 1560

pctau2012001598-seq1

ggttgagtct atcaataaga ccaagtccat ggagtaataa tccagctagt ctggttgctg	1620
ctattttgct tgggagtga gttaggtctt tcgggttttc ccaattactt tttttttgt	1680
tttttttcat gtattgttcg aggaaaatgt tttgttgaat atgcactgga taatgaaact	1740
acatttgaat caaagtacgt acatgtagtg taccatttcc ccttgcctta actatgaatg	1800
atttgtaaac tcaccttgta ttataagata tcatagagaa aaacacattc tcgact	1856

<210> 85
 <211> 1845
 <212> DNA
 <213> Arabidopsis thaliana

<400> 85	
atatccattc tcaccttcca aatctcccca ctctctctct ctctctctct ctctctctct	60
ctctctctct ctctatctct ctctatctct ctctatcttc tcttttctcc tctctccaac	120
ttcttcacta atgggagctc aggagaaacg gcgccgtttc gagcagatat caaagtgcga	180
tgtaaggac cggccaacc ataccgtggc cgctgatcta gacggaacac tactaatctc	240
tcgtagcgcc ttcccttact atttcctcgt agccctcgag gcaggagagct tgctccgagc	300
gttgatccta cttgtgtccg taccattcgt ttatcttacg tacttgacca tctccgagac	360
tttagccatc aacgtatttg tcttcatcac gttcgcgggt ctcaagatcc gagacgttga	420
gctagtggtc cgttccgtcc tcccagaggt ctatgcggag gacgtgaggc ccgataacctg	480
gcgtatcttc aacacgttcg ggaaacggta cataataact gcgagccctc gaattatggt	540
cgagccattc gtgaaaacat tcctaggagt tgataaagtt cttggaacag agctagaggt	600
ctccaaatcg ggtcgggcaa ccgggttcac cagaaaacca ggtattctcg tcggtcagta	660
caaacgtgac gtcgttttga gagagtttgg tggcctagcg tctgatttac ctgatttggg	720
gctcggcgat agcaagacgg accacgactt catgtccatc tgcaaggaag gttacatggt	780
gccacgtacg aaatgcgaac cattaccaag aaacaaactc ttaagcccca taatattcca	840
cgagggcaga ttagtccaac gcccaacgcc gttagttgct ctgttaactt tcctctggct	900
tcccgtcggt ttcgtcctct ctatcatccg cgtctacacg aatattccgt taccggaacg	960
tatcgcccgt tacaactaca agcttactgg catcaagcta gtcgtcaacg gccaccctcc	1020
tccgcccga aaacctggcc agccaggcca tcttttggtc tgcaaccacc gcaccgttct	1080
cgatcctgtg gtcacagctg tcgcactcgg ccggaaaatc agctgcgta cttacagcat	1140
cagcaagttc tctgagctaa tctaccaat caaagccgtt gcgttgactc gtcaacgtga	1200
gaaagacgca gcgaacatca agcgtctttt ggaggaaggc gatctcgtga tatgtcccga	1260
gggaaccacg tgccgtgagc ctttccttct ccggtttagt gctcttttcg ctgagctcac	1320
ggaccggatc gttcccgtgg cgatcaacac aaagcagagc atgttcaatg gtaccaccac	1380
acgtggatac aagcttcttg atccttactt tgcgttcacg aacccgaggc cgacgtatga	1440
gatcacgttc ctcaaacaga ttccagctga gctgacgtgt aaaggaggca aatctccgat	1500
agaggttgcg aattacatac agagggtttt gggaggaacc ttaggttttg agtgcaccaa	1560

pctau2012001598-seq1

tttcacaaga aaggataagt acgcaatgct tgctggtact gacggtaggg ttccggtgaa	1620
gaaggagaag acgtgattat tgatcgagg atggagaaaa caaagaaata atacaaagat	1680
tggaatcttt ttcatttgga tggttaatta tcaatactta aggggcaaag aattgataag	1740
tttgtttttg ttgttacctt tatcttctct gcttcggtga atgtttcttc aagttttaca	1800
tttagtacia taacggttta agaaaagaaa gaaaaaagta tcggt	1845

<210> 86
 <211> 1216
 <212> DNA
 <213> Arabidopsis thaliana

<400> 86	
tcacttcttg gtgttggtga tagactccac cttgccgtca ttacctcaa gcaaaagata	60
tttatccttg cgagtaagtt cgggtgcattc gaagcccaaa accgcgccga taactttctg	120
gacgtaatta gccacctcta taggagtctt gccaccaccg ttgacagtca tttcttcagg	180
caaacgatcc aagaaagtgg cttcatagct tggctctggg ttcatagaaga agaagtaagg	240
gtcccaaaaac ttcacacccc taactgtggt cccgttgaac attccttggt tacagttcat	300
cgctactggc acaatccggt cgcttagctc tgcgaataga gcgctaaatc tcagtagata	360
ctcttctcta cacgtcgtgc cttccggaca tatcaccaag tcgcctgtta tacacaaaaa	420
acatatcatt agtcaacaa ctccattgat ttttaatttc tcaactaaag attaattgtga	480
atgtaagtca atattcaaaa gaagtaaaca atattatatt atgttcattt atataaattt	540
ggtagaaga tgcgttaatt aggtaccaa gtaaacatca aagaccatac caaagtaata	600
tttactatat tgagatttgc atgggctact atatatgtat tcatcgaaaa ataaaaagaa	660
agaaatagaa actcctattg aaccccctgg acagtgaac ttgttatagg tatgctctag	720
taatatatga tgagaagatg tgtaaagtac ctttctcgag aagttttctc atgttggcag	780
catcgggtggc acggtcacgg gtgagggcaa cagcaggaat aggagaaagc ataagggaga	840
gacgagagac actgtaagt acgcaacaga tcttacgtcc aagagcaata gcaacgatga	900
tgggatcaag cgcggtacgg tggtaagga catagagggt gccaaagagt ccaggggaag	960
gaggtggagg acgatgacca cgaatgggtta agtggatccc gagcatctcg taagtgtaac	1020
ggacaaatct ttcaggtaaa gggagggttg agtagacgcg aatgatggag aggatgaaac	1080
caaaaggaag ccataggat gtgataatgg cgtttaacgg agttggacgt tgcgctaaac	1140
gccatcatg gaagactatg cggttcttta agcgttcttt tggaatcggt gtggctgact	1200
tggtcgcagt aacct	1216

<210> 87
 <211> 1736
 <212> DNA
 <213> Arabidopsis thaliana

<400> 87	
atctctcttt ctctgtttac taaatcgcca gtattcgccg caagaaagat gtctccggcg	60

pctau2012001598-seq1

aagaagagcc	gaagttttcc	tccgataagc	gagtgtaaaa	gcagagagta	tgattcgcgc	120
gctgcggatc	tcgacgggac	tctgcttctg	tcaagaagct	ctttccctta	cttcatgctg	180
gtggccattg	aagcaggtag	tctcttccgt	ggattgatcc	ttcttctttc	tcttcccatc	240
gtcatcatcg	cttatctctt	cgtctccgaa	tctctcggaa	tccaaatcct	tatcttcatc	300
tccttcgccc	gtatcaagat	caagaacatc	gaactcgtct	ctcgcgccgt	tcttacacgg	360
ttttacgcgg	cggatgtgag	gaaagatagt	tttgaggtgt	ttgataaatg	taagaagagg	420
aaagtagtgg	tgactgcgaa	tccgatagtg	atggttgagc	cattcgtcaa	agactatttg	480
ggaggagata	aagttttggg	aacagagatt	gaagttaacc	ccaaaacgat	gaaagctact	540
ggttttgtga	aaaagcctgg	tgttcttgtt	ggtgatctca	agagattagc	catcttgaaa	600
gagttcggcg	atgactcacc	ggatcttggc	ctcggtgacc	gcacctccga	tcacgatttc	660
aagtcatttt	gcaaggaagg	ttacatggtg	catgagacca	aatcagccac	aacagtccct	720
atagagagtc	tcaagaaccg	cataatcttc	catgatggcc	gtcttgtcca	acgtccgacc	780
ccattaaacg	ccttaatcat	ttacctttgg	cttcctttcg	gcttcatgct	ctctgtcttc	840
cgcgtctact	tcaacctccc	tttaccggaa	cgcttcgtcc	gctacactta	cgagatcctc	900
ggcattcacc	tcacaatccg	tggccaccgt	cctccacctc	cttcccccg	aaaacctgga	960
aacctctacg	ttctcaacca	ccgcaccgcc	cttgacccca	ttatcatcgc	cattgctctc	1020
ggccgtaaga	tcacatgtgt	cacttatagt	gtctctcgcc	tctccctgat	gctctcacc	1080
attcctgctg	ttgctctgac	ccgagaccgt	gtcgtgacg	cagcccgcac	gaggcaactc	1140
ctcgagaaa	gtgatttgg	gatctgtcct	gaaggcaca	cgtgtagaga	accatatcta	1200
ctaaggttca	gtgctctatt	tgcagagctg	agcgaccgga	tcgtgcctgt	ggccatgaac	1260
tgcaagcaag	ggatgttcaa	cgggacgaca	gtagaggtg	tgaaattctg	ggatccgtac	1320
ttcttcttca	tgaacccccg	acctagctac	gaggccactt	tcttgaccg	tttgcccga	1380
gaaatgacgg	taaacggtg	tggcaagact	ccttttgagg	tggctaatta	cgttcagaag	1440
gtgatcggtg	gagttttggg	gttcgaatgt	actgaactta	caaggaagga	taagtatctg	1500
ttgcttggag	gcaacgacgg	aaaggttgag	tctatcaata	agaccaagtc	catggagtaa	1560
taatccagct	agtctggttg	ctgctatctt	gcttgggagt	gaagttaggt	ctttcgggtt	1620
ttcccaatta	cttttttttt	tgtttttttt	catgtattgt	tcgaggaaaa	tgttttggtg	1680
aatatgcact	ggataatgaa	actacatttg	aatcaaagta	cgtacatgta	gtgtac	1736

<210> 88
 <211> 1817
 <212> DNA
 <213> Oryza sativa

<400> 88	
ttaacatccc	atggccttct tctcgtcgat gttgggcttg gacacgacgg tgccgtcggt 60
gccggcgagc	gccctgtact tgtccttcct ggtgaagctg gtgcactcgt aggagagcgt 120

pctau2012001598-seq1

ggaggcgatg agcctctgga tgtagttggc cacctcgtgg ctgctccggc cgccgccgcc	180
gccgccaccg ccgttgagg tgagctcgcc ggggagcttg ctgaggaagg tgaccacgta	240
tccggggcta ggggtcatga agaagtagaa ggggtccagc gccttcacc ccctcgccgt	300
cgtgccgtgg aacatgctca tctggttctc catcgccacc ggcacgatct cgtcgggtgag	360
ctccgcgaac agcgccgaga accggaggag gaacgggtcc cggcacgtcg tcccctccgg	420
gcagatcacc aggtcgccct ccgccagcag ccgccggatc atcgccgcgt ccgccgccct	480
gtcccgctc agccgcaccg tcctgatcgg cgacaatata tccgacagcc gcgacacctg	540
ttcagacacg caccaagatt tttattttta aaattttgaa caacttaaat tttaaaattt	600
aaaaataaac ctttctaaga tttattatca acatctaaac ttttttgacc atgttagccg	660
agttggcatc atacaggtct ggctagccta acagcgttgt cctgccacac caccctaggt	720
aacgtagcag aacaaaactg tcacgcttga cagcacagtg tcattttgcc atgttagaca	780
caggctggcg tgattaaaat gttcaacttt ctaaaataag ctatagaata gtttgTTTTT	840
aatataaaaa aagaaaaatt caaaagttaa tataaattcc atgcaccggc gagttagctg	900
gtgagctagc tgctcgccgg agttggtggg ggtgtgatca tgtgtactac cgagtaggtg	960
acggcggtga tggggcggcc gagggcggtg gagaggaaga tggggtcgag gagggcgcg	1020
tgggagcaga tgaagagcac gccggactgg ccggtctcgc ggctggccgg tggcgccggg	1080
gtgcccttga cggtgacgag cacgccgagg gcgcggaagg cgtggtacac cagcgcatc	1140
gggaggagcg cgccggcgcc gatgcggagg caggcgagca cgaagccgat ggggatccag	1200
agcacggtga gcagcgag cgccggcgac ggcttctgga ccaggcgcc gtcgtggaac	1260
accaccggct tcgggagctc ctcccgccg acgggcctcg gcgtcggcga caccggcacc	1320
acgtaccct cttgcatag cctcatgaac gggtagtccg tcctccggtc gccgatgccg	1380
atctccggcg aggcgtcgcc gaacgcctcc cgagcgcgcg ccgccttggt ctcgccgacg	1440
agcacgccgg ggctccggac gagccccgtg gcgcggccac gccagaccac gagctccgtg	1500
cccacgacga tgtcggcgcc gaggtactcc ttcaagaacg cctccaccat catcctcggg	1560
ttcgcggtga gcacgcaccg gcgcccgcac gccgagaaca cgccccacga ctccgggtgg	1620
atgtcggagc agtagaactt gggcagcacg gcgcgcgcca ccgcctcaac gtcggcgacc	1680
ctcgcgccgg ccatggacgc gaagatgagc acctggatgc cggccgactc ggacaccagg	1740
tagtagagca cgccggcgag cggcgccagc aggatcagcg ccagcagccg gaggacgccc	1800
cccgtctcga acgcat	1817

<210> 89
 <211> 1924
 <212> DNA
 <213> Picea sitchensis

<400> 89	
cccccggtgta cgTTTTgttt gagaaggaat atggcgaagg gccagagtgg cctcgactgt	60
caagggctct ttgatgtgat tagcaaatgc agtccaagg gccgagaaaa tcagaccgtg	120

pctau2012001598-seq1

gctgcggatc ttgatggcac catgctcaga tccaggagct cgttccctta tttcatgctc	180
gtggccatgg agggcgggag cctgttacgc ggcattggtgc tcctcgctg cgctcccatg	240
gcctggtttc tctacaactt catctccgaa gctgcgggga ttcagcttct catctttatt	300
tccttcgctg gactcaagat ccgggatatc gagctcgttt cccgcgccgt gttgccaag	360
ttctatgccg aggacgtgca cccggagtca tggaggttct tcagcgctt tgggaagcgg	420
tacatcgtga ctgcgaatcc aagaatcatg gtggagccct tcgtgaaggg ctatttgggg	480
gctgacaagg ttcttgaac ggagctgcag gtttccagga gcgggcgggc taccgggttc	540
gtgaagaagc ctggtgtgct cgttggggat ttgaagaagg cagctgtgga agcggaattc	600
ggtgataaat tgcctgagct tggccttggg gatcgagaga cagatcatcc tttcatgtcc	660
ctctgcaagg aggggtacat ggtgccaag atgaaggtag acgaagtgcc cacaacaaa	720
ctaagagcc caatcgtctt ccacgacggg cgtttggtgc agagaccgaa cccgggggcc	780
gcccttctca ctttctatg gctgcccata ggcttctttt tggctcagtt tcgagtcttc	840
ggcaacattc ccatcccaga gaaatacgtg aaaatctcct acaagataat gggaatcaag	900
ctggtcgtca agggaacgcc tccgccggcc ccgaagaaga aaggcgagcg gggcgtgctg	960
ttcgtatgctg atcacagaac gttgctggat cccgtgatcg tggccgtggc cctgggaagg	1020
aaagttagcg ccgtcacgta cagcatcagt cggttctcgg agattatctc gccattaaa	1080
accgtccggc tcactcggga ccgcgagagg gacgccgcta acatcaagcg cctgctggag	1140
gagggcgacc tggatgatctg cccggagggc accacgtgca gggagccctt cctcctgagg	1200
ttcagtgtc tattcgcgga gtcacggac aggatagtgc ccgtggccat ctgcaacaag	1260
atgagcatgt tccacggaac aacggtgcgc ggggtggaagg gcttcgatcc tttttcttt	1320
ttcatgaatc cgggtgccac ctacgaggta acgttcttgg atcagttgcc tacggagctc	1380
acgtgcgccg gaggcaaatt tcccatcgaa gtcgctaatt acattcagag ggtgctggcg	1440
gccactcttg gctacgagt caccaatttc accaggaagg acaagtaccg gatgctcgca	1500
ggcaacgatg gcattgtttc ctttaataag actgctactg cccgcagcaa caaccacaaa	1560
cagaattgat ccaccgaatt ccaccgaatt ccaggccggg tattcttaat catgataaat	1620
tttaataatc tcttattaat ctatgattcc aattattaca tactgcaa at ttgaacgcgt	1680
agttaagcag aatcttccca cgaatgaatg tatacacctc ataattggta tatctgtatt	1740
gcaaattaga tcagatcgta gttaaagaag aatccacta atgggtgctg cataattggc	1800
acatctggtc attatattgc aaattattcc agtaatttgc ctggactgta ttatgctcaa	1860
ggatcagaaa taatataatg cccaaggatt tttttgtcca aaaacaagaa aaaaaaaaaa	1920
aaaa	1924

<210> 90
 <211> 1765
 <212> DNA
 <213> Zea mays

pctau2012001598-seq1

<400> 90
 cgacgcagc agtacgacgt ctctcctctg ggtctggggc cgagacaccg agcacgtact 60
 accagcaaga tgggtggcgtc tcccagattc aagcccatcg aggagtgtg ctcggagggg 120
 cggtcggagc agacgggtggc cgccgacctg gacggcacgc tgctcatctc caggagcgcg 180
 ttcccctact acctcctcgt ggctctcgag gccggcagcg tcctccgcgc cgcgtgtgtg 240
 ctctgttccg tgccgttcgt ctacgtcacc tacgccttct tctccgagtc gctggccatc 300
 agcacgctgg tgtacatctc cgtggcgggg ctcaagggtg gcagcatcga gatggtggcg 360
 cgctccgtgc tccccagggt ctacgcgggc gacgtgcacc cggagagctg gaggggtgttc 420
 agtccttcg gcaggcgcta cgtcgtcacc gccagcccca gggcatgtgt cgagcccttc 480
 gccagggcct tcctcggcgc cgacaaggtc gtcgggaccg agctggagggt cggcagggac 540
 ggcaaggcca cgggcttcat ggccaggcca ggcgtgtctg tcggcgacca caagaagaag 600
 gccgtcgtca aggagctcgg cgacgcgctg cccgacgtcg gcatggggga tagggagacc 660
 gacttcgact tcatgtccat ctgcaaggag gcctacctgg tgacgtcaag gaagtacagc 720
 ccggtgcca ggaaccagct gctgagcccg ctgatcgtgc acgacggccg cctcgtgcag 780
 cgcccgacgc cgctcgtcgc gctcgtcacc ttctcttgga tgccgttcgg cttcgcgctg 840
 gcgctcatgc gcgtgtacat caacctgccg ctgcccagac gcatcgtcta ctacacctac 900
 aagctcatgg gcatcaggct cgtcgtcaag ggcacccgc cgccgccgcc caagaagggc 960
 caccggggcg tcctcttcgt ctgcaaccac cgcaccgtgc tcgaccccg ctaggtggcc 1020
 gtggcgctgc gccgaagggt cagctgcgtc acctacagca tctccaagtt ctccgagctc 1080
 atctcgcca tcaaggccgt cgcgtgtctg cgggagcgcg acaaggacgc cgagaacatc 1140
 cgccgcctgc tggaggaggg cgacctggtc atctgccccg agggcaccac ctgccgcgag 1200
 cccttcctgc tgcgttcag cgcgtcttc gccgagctca ccgaccgat cgtgcccggtg 1260
 gccatcaaca ccaaggagac catgttccac ggctccaccg tgcgcggctt caagctcatg 1320
 gatccctact tcttcttcat gaaccgcgc ccgacgtacg agatcacgtt cctcaccag 1380
 ctcccaaag acctcacgtg cagcggcggc aagtcgcca tcgaggtggc caactatata 1440
 cagaagacgc tcagcggaca gcttggcttc gactgcacgt ccatcacgcg caaggagaag 1500
 tacggcatgc tcgctggcac cgacggccgt gtcccgtcca agaacaagga gaaggagaag 1560
 gagaaggata agaactaacg gccaacgtcg ctctcttgct gctgccagc agactgccgt 1620
 ttaattactg ctcaaattta ttccaattta ttgttccttg tcgacaagtt ggccagtgtg 1680
 atgtgggccg tgtactagta caagcttgat gatgattagt atatctaaa ttcttccttt 1740
 ccggaagaag ggcggaagag ttgtt 1765

<210> 91
 <211> 1778
 <212> DNA
 <213> Arabidopsis thaliana

pctau2012001598-seq1

```

<400> 91
agagacacac acacagcgag tgaaaatgtc tccggagaag aagagtcaaa acttccctcc      60
aataacggaa tgcagagacg gagagtacga ttcgatagcc gccgatcttg acgggactct      120
gcttctctca agaagctcct tcccttactt catgctcgtc gctgttgaag ctggaagcct      180
tttacgtgga ctaatcctcc ttctctcggt accattcgtc attatctctt acctattcgt      240
atccgaatct cttggtatcc agatcctcat cttcatctca ttcgctggtc tcaaaatccg      300
cgatatcgaa cttgtctctc gcgcagttct tccacggttt tacgcggcgg atgtgaggaa      360
agacagtttt gaggtgtttg ataagtgtaa gagaaaagtg gtagtgacgg cgaatccgat      420
tgtgatgggt gaggcgtttg tgaaggatta tcttggagggt gataaagttt tgggaacaga      480
gattgaagtt aaccctaaaa ccaatagagc cactggattt gtgaagaagc ctggtgttct      540
tgttggatgat cttagagggt tagccatttt aaaagagttt ggtaacgaat cacctgatct      600
cggcctcggt gatcgaacct ctgatcatga tttcatgtct ctctgcaaga aaggttacat      660
ggttcatgcg accaagtcag ccacaacgat tccaaaagaa cgcttaaaga accgcatagt      720
cttccatgat gggcgtttag cgcaacgtcc aactccgtta aacgccatta tcacatacct      780
atggccttcc tttggtttca tcctctccat cattcgcgtc tacttcaacc tccctttacc      840
tgaaagattt gtccgttaca cttacgagat gctcgggatc cacttaacca ttcgtgggtca      900
tcgtcctcca cctccttccc ctggaactct tggcaacctc tatgtcctta accaccgtac      960
cgcgcttgat cccatcatcg ttgctattgc tcttgacgtt aagatctgtt gcgtcactta     1020
cagtgtctct cgtctctccc ttatgctttc tcctattcct gctgttgccc tcacccgtga     1080
ccgtgccacc gatgctgcca acatgagaaa acttctcgag aaaggcgact tggatgatg      1140
tccggaaggc acgacgtgta gagaagagta tctactgaga tctagcgctc tattcgcaga     1200
gctaagcgac cggattgtgc cagtagcgat gaactgtaaa caaggaatgt tcaacgggac     1260
cacagttagg ggtgtgaagt tttgggaccc ttacttcttc ttcattgaacc caagaccaag     1320
ctatgaagcc actttcttgg atcgtttgcc tgaagaaatg actgtcaacg gtggtggcaa     1380
gactcctata gaggtggcta attacgtcca gaaagttatc ggcgcgggtt tgggcttcga     1440
atgcaccgaa cttactcgca aggataaata tcttttgctt ggaggtaatg acggcaaggt     1500
ggagtctatc aacaacacca agaagtgaat atcttattct tgttataact tggcattgac     1560
ttattgcaag taaggttgga tcatatacga attttcatgt gtgtaattgc ttgagctctt     1620
gtcatcatta tggtagctta ttgttatgtt tattctattt tgttcttaat gtcgaacggg     1680
taactatatc cttaaaaact attttgtgat cccgtttcta tacggatgtg tatagagact     1740
ttctatttaa agtgcaaaat tttattggag aaaaaaaaaa                        1778

```

```

<210> 92
<211> 1832
<212> DNA
<213> Oryza sativa

```

```
<400> 92
```

pctau2012001598-seq1

ggctagctgc	tccgtgtctc	cctcgttgca	tggttaagcta	gcatgtcgcc	gttcaagccg	60
atcgagcagt	gctcgacgga	ggggcggtcg	cagcagacgg	tggcctccga	cctcgacggc	120
acgctgtctc	tgtcccgag	cgccttcccg	tactacctcc	tcgtcgcgct	cgaggccggc	180
ggccccctcc	gcgcggtcgc	gctgctcatg	tccgtgccct	tcgtctacct	cacgtacgtc	240
accatctccg	agtccctcgc	cgtgcgcgcg	ctgctgtaca	tcgccgtggc	gggtctcgag	300
gtgagggacg	tcgagtcggt	ggcgcggtcg	gtgctgccca	ggttctacgc	cggcgacgtc	360
cacccggagg	ggtggcgggg	gttcagctcg	ttcggcagga	ggtgcgtcgt	gacggcgagc	420
cccaggggtga	tgggtggagcc	gttcgcgagg	gcgttcctcg	gcgcggacag	ggcatcggc	480
acggagctgg	aggtcggcga	ggacggcagg	gccaccggct	tcgtcgccaa	gcccggcgtg	540
ctcatccgcg	agcacaagag	gaacgcggtg	gtgagggagt	tcggcgacgc	gctgccggac	600
gtcggcatgg	gggacagga	gagcgacttc	gacttcatgg	ccatctgcaa	ggacgcgtac	660
gtcgtgacca	cctcgcgga	gcaccggccg	gtgccggaga	gccagctgct	gagaaccgtc	720
gtcctccacg	acggccgcct	cgcgcagcgc	ccgacggcga	tcaacacgct	gctcgtcttc	780
ctctggatgc	cggtcggctt	cgcgctcgcg	ctcctccgcg	cgtgcctcag	cctgctcctc	840
cccagcgcg	tcctctccta	cgcctacaag	ctcacgggcg	tcgggctcgt	ggtgcgcggc	900
cgcgcgcgc	cgcggacgg	gagccccggc	gtgctcttcg	tgtgcaacca	ccgcaccgtg	960
ctcgaccagg	tcgccgtcgc	cgcgcgcgtg	ggccgcaagg	tgatctgcgt	cacgtacagc	1020
gtccctcgga	aaacgtacgg	catgagctcc	aggctcccgg	aggcgctcac	ggcctcgccg	1080
gtcaaggccg	ccgtcgcgct	gtgcaggag	cgcgacaggg	acgcggaccg	tgtccggcgg	1140
ctgctggagg	agggcgctga	catcgtcgcc	ttccccgagg	gcaccacctg	ccgaggggcg	1200
ttcctgctgc	ggttcagctc	gctgttcgcc	gagctcaccg	accgcatcgt	gccggtggcc	1260
atcgccacca	gggagaccat	gttcacggc	tccacggcgc	gcgggtttta	gggcatggac	1320
ccttacttct	tcttcatgaa	cccgcggccg	gcgtacgagg	tcacgttcct	gagccagctg	1380
ccgagtgagc	tcacctccgg	cggcgggcgg	aagtcgccgg	tcgaggtggc	caactacgtg	1440
cagaaggcgc	tcgccgggca	gctggggctc	gagcacatcg	gcatcacgcg	caaggagaag	1500
tagcgaatgc	ttgccggcac	cgacggccgt	gtggagtcga	atgaggacta	gaaccagaag	1560
ccacacgccc	actattattg	gttccaatat	atatggccag	gagcctcggt	tcttcgttat	1620
ttctgtcaca	taacaagctt	caagtgttta	tacgagaact	gaagatgcaa	catagtagaa	1680
ctcatgttgt	tcgttgtgga	cttgtggtat	aacgactttg	atcatttctt	tgtcctttta	1740
tttttacaaa	aatccaagaa	atatagtcac	ataagtttca	tactgatgta	aattcttttg	1800
ctatataata	gatttggtga	tgtttttttt	tc			1832

<210> 93
 <211> 2529
 <212> DNA
 <213> Oryza sativa

pctau2012001598-seql

<400> 93
tcaattggcc ttgggcttgt cggcggcggc ggcgacccgg ccgtcgttgc cggcgagctt 60
catgtacttg tccttccttg tgagcgtggt gcattggaac cctagctcct tggcgatcac 120
cctctgcacg tggttcgcca cctccaccgc gtcctctccg ccggcgacgc acgtctcctc 180
cggcctcagc gccggcagga acgtcacctc gtaccccggc cgcgggttca tgtagaagaa 240
gtaaggggtcc aggaacttcc accccctcgc cgtcgaccca tagtacgtcc cctgcctcgc 300
ctccaccgcc accggcacga tcctcgccgt cagctccgcg aacagcgccg agaaccgcag 360
caggtacggc tcccggcacg tcgtcccctc cgggcacacc accacgtccc cctcctccag 420
cagcgccgcg atgcgcgccg cgtccgccgc gcggtccggc gtcagcgccg ccgcgcggat 480
cggcgagatc gccgtcgaca gccggctaac gctgtacgtc acgcacgtca cggggcggcc 540
cagcgcgatg gacacgatga tgggggtccag cgcggtgcgg tggttgcaca ccaggaggga 600
gcccgccgctc cccgggcgcg gcggcgccgg cggcgccccg cgcacggcga ggcggatgcc 660
ggtgaggcgg taggtgtggc ggacgaggcg cgccgggacg gggagggttg ggaacacgcg 720
gaggagcgcc acggcgaagc ccacggggag gtaggccagc gcgaacagcg cgctcgctgg 780
ctccgggcgg cggacgaggc ggccgtcgtg gaagatggcg cgcgacagca gctcgtcggc 840
ggcggcgcg cgcgcgcgct tgttcttggg caccatgtag gcttctcgcg cattagcaat 900
ttggcgccaa aacagagtta gcacactaag gctgtgttta gtttcacatc aaaattgaaa 960
gtttggagaa attagaacgg tgtaacagaa aagttggaag tttgtgtgtg tagaaaaatt 1020
cgatgtgaca gaaaagttaa aagtttaag aaaaaagtta aatctaaaca aggcctaatt 1080
tttacaataa tctggctttg gagtatattc atatgtacat tttgaacagt tgtcttcttt 1140
ttttttgtta aataaaaaa aatattgatg atggactagg ttgcattctg aatgtaagca 1200
tgaaatttta gtttatgata gtctaacc aa ttaaaaattg ctagtgggtca aagttgaaac 1260
agtagagtca atattcagag taaatgcaaa tggagagagt actaaatcca cattaattga 1320
tgcaaagaaa acaatactat atgagtaaag tgatcaatga taatgatatt actataacaa 1380
gtcagcatga gcaacaata tttgaatata ggagtacata ctaatttggg gttttggttt 1440
taaaaaggac gaaataaata ggaagcaact gctaactgcc tactaaagga ggtgatggta 1500
ccaacagtat tgagtttttg ccacaagttt tgcataacca taatgtaaat tctagagcct 1560
tcttgatctt tcctatgaaa agaggggaaa aaaaaggcaa ttcaacaaaa agaaaaggca 1620
aaaagctaag aagcgtacat attcttgcta ctcgtgctac cacaaatgag ccaatcggat 1680
cgaagccatg aacactcgaa gaacaatcga gcgtgcaatt cctaccacaa ttaatttcac 1740
aatattttct ttttttattt ttggcgttct aaatgaattg ttaaattatt agtggccaca 1800
tgatgcaaaa cttcaatttt gttcatgtgt gattgattca attcaagaaa tgatgaaaat 1860
ttcgagcttt ggatcgtagg ccatgaagtg taccttgtag atggccatga agtcgtggtc 1920
gctctcgcgg tcgccgaggc cgacgtcggg catgtcgccg ccggcgaaga gcctctcgac 1980
gacctcgcgc ttcttctcgc cgacgaggac ggccttgatg cggccggtga accgcttgcc 2040

pctau2012001598-seq1

ggaggcgaac	gtctcgagct	cggtgccggc	cacctcggcg	ccgaggaact	cgcgcacgaa	2100
ctccccacc	atgacggccg	gcgacgcggt	gaccacgacg	cggcgctcgg	cgagccggtg	2160
gaacaccgcc	cacgtgtcgg	cgcgcacccc	ggcggcgtag	tggcgcgga	gcacggcgcg	2220
cgccacggcc	tcgacgtcgc	ggacgcgcag	cccggcgaac	gtgatgaaca	cgagcagggc	2280
gatggcggcg	gcctcggaga	cgccgacgta	gagcgcgagg	agccacggcg	cggcgagcag	2340
cagcgccagc	gcgcggaggt	agctgccggc	ctcgagagcc	acgaggaagt	agtacgggaa	2400
cgccgacgac	gacaccagca	gcgtcccgtc	gaggtccgcc	gccaccgtcc	gcctcgcgcc	2460
cgccgccgcg	tcgtacgccg	tcaccggcgg	aaagaacctc	ctcctccccg	ccgccgccgc	2520
cgccgccat						2529

<210> 94
 <211> 3320
 <212> DNA
 <213> Oryza sativa

<400> 94						
atgtcgccgt	tcaagccgat	caagcagtgc	tcgacggagg	ggcggtcgca	gcagacgggtg	60
gcctccgacc	tcgacggcac	gctgctcctg	tcccgcagcg	ccttcccgtg	ctacctcctc	120
gtcgcgctcg	aggccggcgg	ccccctccgc	gcggtcgcgc	tgctcatgtc	cgtgcccttc	180
gtctacctca	cgtacgtcac	catctccgag	tccctcgccg	tgcgcgcgct	gctgtacatc	240
gccgtggcgg	gtctcgaggt	gagggacgtc	gagtcggtgg	cgcggtcggt	gctgcccagg	300
ttctacgccg	gcgacgtcca	cccggagggg	tggcgggtgt	tcagctcggt	cggcaggagg	360
tgcgtcgtga	cggcgagccc	caggggtgatg	gtggagccgt	tcgcgagggc	gttcctcggc	420
gcggacaggg	tcacgggcac	ggagctggag	gtcggcgagg	acggcagggc	caccggcttc	480
gtcgccaagc	ccggcgtgct	catccgcgag	cacaagagga	acgcggtggt	gagggagttc	540
ggcgacgcgc	tgccggacgt	cggcatgggg	gacagggaga	gcgacttcga	cttcattggcc	600
atctgcaagg	tacggacaaa	gaatattgcc	tattccctta	actacttact	actacctaag	660
ttccaaaaat	agtttgcttt	tacacatgga	caactgcctg	tctaaatgta	gaaaaatgag	720
ttctttttta	aaaagtagag	gtaggattgt	actagcacac	atgttcatat	tacgtgtgtc	780
ctagtatttg	gtacagtaca	tttataacgg	caagtagctt	atagtgaac	ctgcaggcag	840
catacttgca	tatggaattt	tgcgatgtgg	ccgcccaatt	atgtacttag	tatttttttt	900
aaaagagatg	tacttagtat	ttcgcataca	atatttgctat	atatttgacg	acaaatttta	960
ccacaaaaat	ttgacagata	taattacagt	acaatcgtag	tgtaattaca	ctgtaacttg	1020
catgtaatta	cactgtaact	atagtgtaac	ttgtatgtaa	ttttcaaaaa	tatctccgta	1080
acatgttatt	tcgataaaat	gaggttgtgg	gaataaatcc	tttcatatat	gtgtgctatg	1140
atgtgttttc	ttcctcacca	aaacaaatct	tgtaatagat	ctaacgattc	aaaattatgt	1200
gaaacttaca	tataagttac	actgtagtta	catgcaagtt	acagtgtaat	tacatataag	1260

pctau2012001598-seq1

ttacagtgtgta	attacactac	cactgtacta	taattacatc	tgtcaaattt	ttaggagaaa	1320
atttgtcgac	aaatgtatag	gtgggtccctt	tggcatatat	aagtaggtta	attgcaacag	1380
tacaaacaat	atattattaa	ttccttgtac	tatgtaatac	actgaatagt	acaacgagta	1440
tgcattttgt	tcgttgccat	ttatgctgag	gataattgta	gtatggcatt	aatcgatctg	1500
gaccaatctc	tccacgttag	gtgtctgcat	tcaccgggtga	gatgatcagt	tgagaggtag	1560
ttgtgccatt	gatatcgcc	actgctttct	cgttatcctt	tctctgtagg	tgactctgaa	1620
tctagctctt	ttcacggatt	gcatcaactg	atatacgtcg	tattcttccc	aagttccctg	1680
cattcatgtc	ttatatatgc	taactggaca	attaatggct	catagcaggg	cagctgcctg	1740
tcagtgatca	ctccagcttg	gttgaatgg	tgagttagcc	tgtctgtttc	atgatatctg	1800
gtacagggtcc	cactcaatac	cgtttggtag	gaatttatac	ttgtgatgta	cttatgtcct	1860
caaataagtc	cagggagaat	atgaacgtaa	gcaaatttgt	gattgctaga	cagacatgaa	1920
aaaccaaagt	ctagccatgc	atgatctaca	gatttccttg	tttcttccct	ccaactttgg	1980
gaagtttgca	ctgaattaat	acggttttgc	ttcatggtag	gccattatat	ttgcaagtaa	2040
tggtcctact	agtataactg	gatcccaa	ttccgatgga	gcatcacaa	cacgatcata	2100
tgtcaagatc	gatattatgc	tgtttggggc	aaccgggcaa	acatagcacc	gggtaaactt	2160
tggttcaact	tacagttaat	cagtagtta	ccatgctgat	atgataata	gctgattagt	2220
ccgtgtaagt	taattaatta	acctaccggg	catgcatgca	agttttggat	ggattttacat	2280
aatggataag	ctgggtactct	gtttgtaaat	tctaactctac	tagcgtcctc	aattttttta	2340
cttctccatt	taaccagcga	cccatgcatt	attctgtatt	tccgttgcat	acttgcattt	2400
ccatgcacgt	gaactatgca	tgacatgatg	catcttgatc	aacgtggcac	gtgcacgc	2460
gcatgcagga	cgcgtagctc	gtgaccacct	cgcggaagca	ccggccggtg	ccggagagcc	2520
agctgctgag	aaccgtcgtc	ctccacgacg	gccgcctcgc	ccagcgccc	acggcgatca	2580
acacgctgct	cgtcttcctc	tggatgccgg	tcggcttcgc	gctcgcgctc	ctccgcgcgt	2640
gcctcagcct	gctcctcccc	gagcgcgtcc	tctcctacgc	ctacaagctc	acgggcgtcg	2700
ggctcgtggt	gcgcggccgc	ccgccgccgc	cggacgggag	ccccggcgtg	ctcttcgtgt	2760
gcaaccaccg	caccgtgctc	gaccgggtcg	ccgtcgccgc	cgcgctgggc	cgcaagggtga	2820
tctgcgtcac	gtacagcgtc	cctcgga	cgtacggcat	gagctccagg	ctcccggagg	2880
cgctcacggc	ctcgccggtc	aaggccgccg	tcgcgctgtg	cagggagcgc	gacagggacg	2940
cggaccgtgt	ccggcggctg	ctggaggagg	gcgtcgacat	cgtcgccttc	cccgagggca	3000
ccacctgccg	aggggcgttc	ctgctgcggt	tcagctcgct	gttcgccgag	ctcaccgacc	3060
gcatcgtgcc	ggtggccatc	gccaccagg	agaccatgtt	ccacggctcc	acggcgcgcg	3120
ggtttaaggg	catggacctt	tacttcttct	tcatgaaccc	gcggccggcg	tacgaggtca	3180
cgttcctgag	ccagctgccg	agtgagctca	cctccggcg	cggcggcaag	tcgccggctcg	3240
aggtggccaa	ctacgtgcag	aaggcgctcg	ccgggcagct	ggggtccgag	cacatcggca	3300

tcacgcgcaa ggagaagtag

3320

<210> 95
 <211> 2532
 <212> DNA
 <213> Oryza sativa

<400> 95
 tcaattggcc ttgggcttgt cggcggcggc ggcgacccgg ccgtcgttgc cggcgagctt 60
 catgtacttg tccttcctgg tgagcgtggt gcattggaac cctagctcct tggcgatcac 120
 cctctgcacg tggttcgcca cctccaccgc gtcctcccgc ccggcgacgc acgtctcctc 180
 cggcctcagc gccggcagga acgtcacctc gtaccccggc cgcgggttca tgtagaagaa 240
 gtaaggggtcc aggaacttcc accccctcgc cgtcgaccca tagtacgtcc cctgcctcgc 300
 ctccaccgcc accggcacga tcctcgccgt cagctccgcg aacagcgccg agaaccgcag 360
 caggtacggc tcccggcacg tcgtcccctc cgggcacacc accacgtccc cctcctccag 420
 cagcgccgcg atgcgcgccg cgtccgccgc gcggtcccgc gtcagcgccg ccgcgcggat 480
 cggcgagatc gccgtcgaca gcgggggctaa cgctgtacgt aacgcacgta acggggcggc 540
 caacgccgat ggacacgatg aaggggtccg aggcggttcc gtggttgac aacaggaggg 600
 agccccggcg tccccgggc gcagcgggcg ccgccgcgcc ccgcgcacgg cgaggcggat 660
 gccggtgagg cggtaggtgt ggcgacgag gcgcgccggg acggggaggt tgaggaacac 720
 gcggaggagc gccacggcga agcccacggg gaggtaggcc agcgcgaaca gcgcgctcgc 780
 tggctccggg cggcggacga ggcgccgctc gtggaagatg gcgcgcgaca gcagctcgtc 840
 ggcgggcggc gcgcgcgcgc gcttgttctt gggcaccatg taggcttcct gcgcattggc 900
 aatttggcgc caaaacagag ttagcacact aaggctgtgt ttagtttcac atcaaaattg 960
 aaagttcgga gaaattggaa cgggtgtaaca gaaaagttgg aagtttgtgt atgtagaaaa 1020
 attcgatgtg acagaaaagt taaaagttaa aagaaaaaag ttaaattctaa acaaggccta 1080
 atttttacia atatctggct ttggagtata ttcatatgta cttttgaac acttgtcttc 1140
 ttttttttgc gttaaataaa aacaaatatt gatgatgggc taggttgcat tctgaatgta 1200
 agcatgaaat tttagtttat gatagtctaa ccaattaaaa attgctagtg gtcaaagttg 1260
 aaacagtaga gtcaatattc agagtaaag caaatggaga gactactaaa tccacattaa 1320
 ttgatgcaaa gaaaacaata ctatatgagt aaagtgatca atgataatga tattactata 1380
 acaagtcagc atgagcaaac aatatttgaa tataggagta catactaatt tgggggtttg 1440
 gttttaaaaa ggacgaaata aataggaagc aactgctaac tgcctactaa aggaggtgat 1500
 ggtaccaaca gtattgagtt ttggccacaa gttttgcata accataatgt aaattctaga 1560
 gccttcttga tctttcctat gaaaagaggg gaaaaaaagg caattcaaca aaaaaaaagg 1620
 gcaaaaagct aagaagcgta catattcttg ctactcgtgc taccacaaat gagccaatcg 1680
 gatcgaagcc atgaacactc gaagaacaat cgagcgtgca attcctacca caattaattt 1740
 cacaatatatt tcttttttta tttttggcgt tctaaatgaa ttgttaaatt attagtggcc 1800

pctau2012001598-seql

acatgatgca aaacttcaat tttgttcatg tgtgattgat tcaattcaag aaatgatgaa	1860
aatttcgagc tttggatcgt aggccatgaa gtgtaccttg cagatggcca tgaagtcgtg	1920
gtcgtctctc cggtcgccga ggccgacgtc gggcatgtcg ccgccggcaa aaagcctctc	1980
gacgacctcg cgcttcttct cgccgacgag gacggccttg atgcggccgg taaaccgctt	2040
gccggaggca aacgtctcga gctcgggtccc ggccacctcg gcgccgagga actcgcgcac	2100
gaactcccc accatgacgg ccggcaacgc ggtgacaaga acgcggcggt tggcccagcc	2160
gtggaaaacg gcccaggttt cggcgcgaac cccggcgggc tagtggcgcg gcagcacggc	2220
gcgcgccacg gcctcgacgt cgccgacgcg cagcccgggc aacgtgatga acacgagcag	2280
ggcgatggcg gcggcctcgg agacgccgac gtagagcgcg aggagccacg gcgcggcgag	2340
cagcagcgcc agcgcgcgga ggtagctgcc ggcttcgaga gccacaagga agtagtacgg	2400
gaacgccgac gacgacacca gcagcgtccc gtcgaggtcc gccgccaccg tccgcctcgc	2460
gcccgccgcc gcgtcgtacg ccgtcaccgg cggaagaac ctctctctcc ccgccgccgc	2520
cgccgccgcc at	2532

<210> 96
 <211> 2941
 <212> DNA
 <213> *Oryza sativa*

<400> 96	
ctagttcttt tccttctcct tgttcttga agggacgcgg ccgtcgggtcc cggcgagtat	60
gctgtacttc tccttgccgc ttatggcggg gcactcgaag ccgagctggc cgctgagcgt	120
cttctggatg tagttggcca cctcgatggg cgacttgccg ccgctgcaag tgagctcctt	180
gggcagctgg ttcaggaagg tgatctcgta cgtcggccgc gggttcatga agaagaagta	240
gggggtccatg agcttgaagc cccgcacggg ggagccgtgg aacatgctct ccttgggtgtt	300
gatcgccacc ggcacgatgc ggtcgggtgag ctccggcgaag agcgcgctga accgcagcag	360
gaacggctcg cggcaggtgg tgccctcggg gcagatcacc aggtcgcctt cctccagcag	420
ccgccggatg ttctcggcgt ccttctcgcg ctcccgcgac agcgcgacgg ccttgatcgg	480
cgagatgagc tccgagaact tggagatgct gtatgtgacg cagctgacct tgcggcgag	540
cgccacggcg acctcaaccg ggtcgagcac ggtgcgggtg ttgcagacga agaggacgcc	600
aggatgtccc ttcttggggg gaggcggcgg gttgcccttc acgatgaggc ggatgcccat	660
gagctttag gtgtagaaga cgatccgctc cgggagtggc aggttgacgt acacgcggag	720
gagcgcgagc acgaagccga acggcatcca caggaagggt acgagcgcca ccagcggcgt	780
cggacgctgc acgaggcggc cgtcgtggag gatgagtggg ctacgcagct ggttcttggg	840
caccgcgctg tacttcttgc atgtcacgag gtaggcctcc tgcacagaca cgcagcatgc	900
atgttacaaa cacttcactg atgcaaaaat cacagcatca ccatcacttc ttattttatc	960
gaatatataa atgcttggtc atctcgccat gtacgccaaa cctacatcct accacgatta	1020

pctau2012001598-seq1

caattcagta	cagtccacta	ctgatggttc	cacagatgac	agatcttcac	atgcttaatc	1080
aaaattatgt	tttttcggag	ttaaactgtt	ggatttaagt	tgacgtgtat	cagccaatca	1140
ataggcgaga	gtagaccaga	gcagtggtaa	agaagtacag	cgattactgt	accgcaacaa	1200
ctgcaagact	ggcgcccata	aatgggtgat	ttcccctcca	cgagagaaat	tcgttatacc	1260
cccggggcag	tgacagaaaag	acgggtccca	tggagaaggg	aaagaagggg	caagaccatg	1320
aatggcagta	tttatttcag	tggttctctg	cattaccggc	tggaagaaaag	gaagaaagaa	1380
agggaaagca	agggcagccg	tggatctggc	aggcacagcc	agcaccagct	ctgatgtctc	1440
acgcgccagc	cggccgatct	ccttcaatgc	ggccgggaca	aggcttgaat	tctcaccaaa	1500
caaccgtaaa	taatggagcc	agccagccgg	ccggacggac	tgactcgacc	gtggcatact	1560
ggcatgagtg	aatccacgca	ttcaatccgg	cgggccacgg	gcgagtaaaa	accatcggct	1620
tccttccatg	ctctggctct	gcccgcctcc	agattccttt	tctgatcatt	tgctcgagct	1680
ttcgagttgt	taagcttcag	gtgaggatct	agtatgtgcg	acgtagacgt	tttgtgagtt	1740
ctgataatga	ccttgatcct	tattattagt	ttactgatta	aggatgttgg	tggaacagag	1800
cgaccgggtca	gtgcaggaag	cttaaactgg	gtttcgcata	atgggctgag	caacatctac	1860
agatcagtgt	cataaatggc	ctgacaatga	ctatacaccc	ctgctgttcg	cgtgctcgac	1920
cactgtggca	tactctactc	tgatctgaaa	ccactctact	ctgatctgaa	acctctgcat	1980
gtttgtacag	attgcagaat	ttatagtagg	cgggtgttcac	agcttcgtca	tagcaaaaat	2040
caattgtaac	tgatgtgcta	attccacgtt	atgaaagtaa	tatgtgtgat	ggagactata	2100
cttaattcca	attttcatgt	ttcttgtttg	gaatgcagta	attccgaaat	aaacagagca	2160
ttctatacgg	aatttatgta	caagatggtc	taaccagttc	gtagttgcca	gtaattggca	2220
gtgctactac	taattcgtcg	caggacttga	tcactgcaat	taacaatcaa	cacgacagta	2280
ctattactat	gagcaaaaaa	tgagtttgag	ggggcgtcaa	atttaccttg	catatggaca	2340
tgaagtcgaa	atccgtctcc	ctgtcgccca	agccgacgtc	gggcaccgcg	tcgcgtagct	2400
ccttgacgac	ggcctgcctc	ttgtggtcgc	ccacgagcac	tccgggcttc	accatgaacc	2460
ccgtggcctt	gccgttcttg	ccgacctcca	gctccgtccc	gaccaccttg	tcggcgccga	2520
ggaacgtctt	ggcgaagtgc	tccaccatga	tcctggggct	cgccgtgatg	atgtaccgct	2580
tgccgaagga	gttgaacacc	ctccagctct	ccgggtgcac	gtcctccgcg	tagaacttgg	2640
gcagcaccga	ccgcgccacc	atctcgatgt	tcctcacctt	gagccccgcc	acggagatgt	2700
acaccagcgt	gctgatggcc	agcgactcgg	agaagaaaat	gtaggtcacg	tagacgaacg	2760
gcacggacag	cagcagcacg	acggcgcgga	gcacgctgcc	ggcctcgagc	gcgacgagga	2820
ggtagtacgg	gaaggcgctc	cgcgacctga	ccagcgtgcc	gtcgaagtcg	gcggccaccg	2880
tctgtcccga	cctcccatcc	gagctgcact	cctccacggg	cttgaacctc	cgcgacacca	2940
t						2941

pctau2012001598-seq1

<211> 1930
<212> DNA
<213> Oryza sativa

<400> 97
ttaacatccc atggccttct tctcgtcgat gttgggcttg gacacgacgg tgccgtcggt 60
gccggcgagc gccctgtact tgtccttcct ggtgaagctg gtgcactcgt aggagagcgt 120
ggaggcgatg agcctctgga tgtagttagc cacctcgttg ctgctccggc cgccgccgcc 180
gttgcaggtg agctcgccgg ggagcttgct gaggaaggtg accacgtatc cggggctagg 240
gttcatgaag aagtagaagg ggtccagcgc cttccacccc ctgcccgtcg tgccgtggaa 300
catgctcatc tggttctcca tcgccaccgg cagcatctcg tcggtgagct ctgcaacag 360
cgccgagaac cggaggagga acggctcccg gcacgtcgtc ccctccgggc agatcaccag 420
gtcgccctcc gccagcagcc gccgatcat cgccgcgtcc gccgccctgt cccgcgtcag 480
ccgcaccgtc ctgatcggcg acaatatctc cgacagccgc gacacctgtt cagacacgca 540
ccaagatfff ttttttaaa attttgaaca acttaaatff taaaatttaa aaataaacct 600
ttctaagatt tattatcaac atctaaactt ttgacaatg ttagccgagt tggcatcata 660
caggtctggc tagcctaaca gcgttgctct gccacaccac tctaggtaac gtggcagaac 720
aaaactgtca cgcttgacac gacagtgtca ttttgccatg ttagacacag gctggcgtaa 780
ttaaagtgtt caactttcta aaataagcta tagaatagtt tgtttttaat ataaaaaag 840
aaaaattcaa aagttaatat aaattccatg caccggcgag ttagctggtg agctagttgc 900
tcgccggagt tgggtggtggt gtgatcatgt gtactaccga gtaggtgacg gcggtgatgg 960
ggcggccgag ggcggtggag aggaagatgg ggtcgaggag ggtgcggtgg gagcagatga 1020
agagcacgcc ggactggccg gtctcgccggc tggccggtgg cggcggggtg cccttgacgg 1080
tgacgcgcac gccgagggcg cggaaggcgt ggtacaccac gcgcatcggg aggagcgcgc 1140
cggcggcgat gcggaggcag gcgagcacga agccgatggg gatccagagc acggtgagca 1200
gcgcgagcgc cggcgacggc ttctggacca ggcggccgtc gtggaacacc accggcttcg 1260
ggagctcctc ccgcggcacg ggcctcggcg tcggcgacac cggcaccacg taccctcct 1320
tgcatagcct catgaacggg tagtccgtcc tccggtcgcc gatgccgatc tccggcgagg 1380
cgtcgccgaa cgctcccg agcgcggccg cttgttctc gccgacgagc acgccggggc 1440
tccggacgag ccccggtggc cggccacgcc agaccagag ctccgtgcc acgacgatgt 1500
cggcgccgag gtactccttc aagaacgcct ccaccatcat cctcgggttc gcggtgagca 1560
cgcaccggcg cccgcacgcc gagaacacgc gccacgactc cgggtggatg tcggagcagt 1620
agaacttggg cagcacggcg cgcgccaccg cctcaacgtc ggcgaccctc gcgccggcca 1680
tggaacgcaa gatgagcacc tggatgccgg ccgactcgga caccaggtag tagagcacgc 1740
cggcgagcgg cgccagcagg atcagcgcca gcagccggag gacgcccccc gtctcgaacg 1800
ccatgtgcgc gaagtacggg aacgagctcc ggccgcacag cagcgtgccg tcgaggtcgg 1860
cgacgatcgt gtcgccggcg cggtccttgg acgaacactt ctccaccgcc gggaacggcg 1920

pctau2012001598-seq1

acgcgatcat 1930

<210> 98
 <211> 1960
 <212> DNA
 <213> Oryza sativa

<400> 98
 ttaacatccc atggccttct tctcgtcgat gttgggcttg gacacgacgg tgccgtcggt 60
 gccggcgagc gccctgtact tgtccttcct ggtgaagctg gtgcactcgt aggagagcgt 120
 ggaggcgatg agcctctgga tgtagttagc cacctcgtgg ctgctccggc cgccgccgcc 180
 gccgccaccg ccgttgacgg tgagctcgcc ggggagcttg ctgaggaagg tgaccacgta 240
 tccggggcta gggttcatga agaagtagaa ggggtccagc gccttcacc ccctcgccgt 300
 cgtgccgtgg aacatgctca tctggttctc catcgccacc ggcacgatct cgtcggtagg 360
 ctccgcgaac agcgccgaga accggaggag gaacggctcc cggcacgtcg tcccctccgg 420
 gcagatcacc aggtcgccct ccgccagcag ccgccggatc atcgccgcgt ccgccgccct 480
 gtcccgctc agccgcaccg tcctgatcgg cgacaatatc tccgacagcc gcgacacctg 540
 ttcagacacg caccaagatt tttattttta aaattttgaa caacttaaatt tttaaaattt 600
 aaaaataaac ctttctaaga tttattatca acatctaaac ttttttgacc atgttagccg 660
 agttggcatc atacaggtct ggctagccta acagcgttgt cctgccacac caccctaggt 720
 aacgtagcag aacaaaactg tcacgcttga cagcagctg tcattttgcc atgttagaca 780
 caggctggcg tgattaaaat gttcaacttt ctaaaataag ctatagaata gtttgttttt 840
 aatataaaaa aagaaaaatt caaaagttaa tataaattcc atgcaccggc gagttagctg 900
 gtgagctagc tgctcgccgg agttggtggt ggtgtgatca tgtgtactac cgagtaggtg 960
 acggcggtga tggggcggcc gagggcggtg gagaggaaga tggggtcgag gagggcgcg 1020
 tgggagcaga tgaagagcac gccggactgg ccggtctcgc ggctggccgg tggcgggggg 1080
 gtgcccttga cggtgacgag cacgccgagg gcgcggaagg cgtggtacac cacgcgcatc 1140
 gggaggagcg cgccggcgcc gatgcggagg caggcgagca cgaagccgat ggggatccag 1200
 agcacggtga gcagcgcgag cgccggcgac ggcttctgga ccaggcgggc gtcgtggaac 1260
 accaccggct tcgggagctc ctcccgcgcc acgggcctcg gcgtcggcga caccggcacc 1320
 acgtaccctt ccttgcatag cctcatgaac gggtagtccg tcctccggtc gccgatgccg 1380
 atctccggcg aggcgtcgcc gaacgcctcc cggagcgcg cgccttggt ctcgccgacg 1440
 agcacgccgg ggctccggac gagccccgtg gcgcggccac gccagaccac gagctccgtg 1500
 cccacgacga tgtcggcgcc gaggtactcc ttcaagaacg cctccaccat catcctcggg 1560
 ttcccggtta acaaggcacc gggggccggc aagccgaaaa aaagggccaa gaatccgggt 1620
 ggatttcgga gcattagaac ttgggcagca cggcgggggg caccgcctta aagtcgggga 1680
 cccttgggccc ggccatggac gcgaagatga gcacctggat gccggccgaa tcggaaacca 1740

pctau2012001598-seq1

ggtagtagag caagccggca aacggcgcca gcaggatcag cgccagcagc cggaggacgc	1800
ccccgtttt gaacgccatg tgcccgaagt acgggaacga gctccggccg cacagcagcg	1860
tgccgtcgag gtcggggacg atcgtgtcgc cgccgcggtc cttggacgaa aacttttcca	1920
ccgccgggaa cggcgacgcg atcatcgggg ggtccccc	1960

<210> 99
 <211> 2311
 <212> DNA
 <213> Selaginella moellendorffii

<400> 99	
cttctgagcc agcagtacag gggctggacg gacttgaaag gcatttttct ccagttccat	60
ttcaagttag cacgttgga gcatttattc tcctctgtgg cgttctttct aactcggagt	120
tgtcccagga ttcgacatgg ccgacttaaa gtactcgcgg ctcgtccagt tcggcctcgt	180
ggagaattgc tcatctgaag gacgtggaat gcacacgggt gcggcggatt tggatggtac	240
actgctcaga cgcagaagct cgtttccata cttcctgctg gtggctcttg aaggaggtag	300
cctcttcaga gcagttgttc tcggactttt ggcccctatc gcctggctgc tttaccactt	360
cgtctcggag gcagcagcaa ttcaacttct catctatgtc tctttttcgg ggcagaaggt	420
gtccgagatt gagaaggtgg caacggcggg tctgccgaag ttttactccc aggacgtcca	480
cagtgatgct ttcagagttt tctcgtcctg cgggaagaag tatgtcatca ctgccaatcc	540
tcggatcatg gtggagtatt tctgcaaggc cttccttggc gcggataaag ttattgggac	600
tgaaatcgaa gtggacgatg atggacgtgc tacaggcttc gtcaagtacc cagggtgtgt	660
cgtgagtcaa aacaagagga ccgccttgaa gctcgagttc gtggacgggg agcttcccga	720
gatcggactc ggcgatcgag aaaccgactt cgccttcatg tcgctttgca aggtaagcgc	780
actctttcca tatcatatac atattttgca tttcgagcat ataatactta gctctgcctc	840
tttcgtgctg ctttaactgt tgaggggtgc caggatttgc gtgacactcc agaatcagtt	900
acgagaagcg agttttctaa ctctggttgt aaccgtgtga tacaggaagc ttacgttgtt	960
cccaagagga aggttgacgc agtggacggc agcgtcttg cgacgcggt gatcttccat	1020
gatggtaggc tgggtccagct cccaacgccg ttcgtggcgc tcgtcacttt cctctggctc	1080
cctatcggct tcgcgtcgc agtgctccg atcggcctca cctcccgttg ccccgaag	1140
tacatgggaa tgatctacgc catcctcggc gtccgccgca tcaccaaagg caaagttcct	1200
cgaaacaag acggcagggg gctcctcttc gtgtgcaacc acaggactct ggtggatccc	1260
atcttcctgg gattggccgc cgagaggac gtgaccgggc tacttacag catcagccgg	1320
gtgtccgagg tgctatgcc catcgtcacg gtccggctga cgagagacaa ggagcgcgac	1380
tttacgaaga tgaagtcgt gctggagagg gggagggacc tgtgcctgtg tccggagggg	1440
acgacttgcc gtgagccgtt cttgctgcga ttcagcgtc tcttcacca gctcacgcac	1500
aggatcgtcc cggcggccgt gaagacgagg cagagcatgt ttaacgggac ggcggtgcgg	1560
gggtggaagg ggatggatcc attcttcttc ttcatgaatc cattcccgcac gtacgaggtg	1620

pctau2012001598-seql

gagtttttgg	aggagctccc	cgcgagatg	agcgtccaaa	agggcggcaa	gagctcgttc	1680
gagacggcca	atttcatcca	gaagatgctg	ggcgacaagc	tgggatacga	gtgcacgacc	1740
ttcacccgca	aggacaagta	cctgatgctc	aatggcagcg	atggaacagt	ccccacgaaa	1800
taggaagggg	ttggaaaaga	aaacaagatg	gaaccgttgg	ctggtgttga	cttgaattga	1860
ctagagttag	cggttgaata	cttttttttt	agtgttggct	ggctatctag	cttgacttag	1920
taaatgctaa	aggttatagg	caggcgattt	taatccttcc	acatggctgg	gacaggaaat	1980
attgacttcg	ttcaattaaa	gggggaaagg	ttttagggca	ggggtgttct	agggcacggc	2040
agcgagctgg	catcctcaca	ccggcagcat	ggcgtccggc	cggcggtacg	gcgctcgtgt	2100
ctcgcattgc	cagagaaatg	ccactcccaa	cctccgccac	cttttctctc	tccgccacct	2160
ttgcgaacta	aacaaatgtg	ttagaggaat	caaagatcca	acacaccgaa	acttcgtttt	2220
tgttttcttc	ttgcctggtc	tttgtggagg	cgaagaaaaa	aaggaaacga	gagagtcttt	2280
gtacctttgg	gaatcaaact	gctattgcgt	a			2311

<210> 100

<211> 1989

<212> DNA

<213> Selaginella moellendorffii

<400> 100

atgtcgtctg	acggcggcaa	ggtcacggat	tctttcctgg	tagatgctcc	attcgagccg	60
gtggagcgat	gctcgtcggc	ggggcgtgag	aagcagacgt	gcgtggcgga	tttcgacggt	120
actctcatcc	gcggccggag	ctccttcccc	tatttcatgc	tgggtggcgct	agagggagga	180
ggcctcatcc	ggagcgtgat	tcttggcctc	ttcgcgccgc	tcgcctggat	cctctaccat	240
ttcgtctcgg	aatccgccgg	catccgcctc	ctcatcttca	tctccttcgc	cgggatgaag	300
gccaagcaga	tcgagtcggt	gtcgcgcgcc	gtgctgccca	aattctactc	ggaggacgtc	360
cacagcgagg	cctaccgagt	gttctcctcg	tgcggcaagc	gtgtggtggt	gaccgccaat	420
ccccggatca	tgggtggagca	cttcgccaa	acgtgcctgg	gagcggacaa	ggtgctgggg	480
accgagattg	aggtgacgag	gagtggctac	gcgacgggat	tcctcaagcg	gccgggagtt	540
ctcgtggggg	tcaacaagcg	caaggctgtg	aggcaggaat	tcggcgagca	tttgccccat	600
gtcggcattg	gcgatcgagc	cacggacttt	gctttcatgg	cgctctgcaa	ggtacttata	660
tcgaatccaa	tcccgtcttg	cttgggtggg	ggagtgggtg	cttagttctc	gatcgagctc	720
gatgcattag	atcaaacact	tgggcattgt	tttcgatcga	tgatccatcc	ctgcaattgc	780
tgcattgagt	gatgattctc	ttgtttcttt	tgtttatctt	ttttccttgt	ctcatgcaca	840
cattcttaca	agcgacttag	cctaaaagaa	tctcttccaa	gcagcagtag	aaacttcttt	900
ccagtcaact	aacgaacca	tttgctgacc	actatcagga	agcatacgtc	gtcccaccct	960
ccaaagtggg	ggcgtgagc	cgagacaagc	tcgtcaagcc	cgttgtcttc	cacgacggcc	1020
gcctcgtcca	gcggcccacg	cccttgaccg	ccctcgtcac	cttcctgtgg	ctgcccgtca	1080

pctau2012001598-seq1

gcttcgtgct ggcggtgttc cggatcatgg tcacgggtacc ctgcccgcgc gagtacgtga	1140
ccatcgtcta caagctcgtc ggcatccggc tcatcgtcaa ggggccgatc ccgccgccca	1200
agaagcgggg cgagtcgggc gtgctcttcg tctccagcca caggacactg tgcgatccgg	1260
tggtcgtggg ggtcggggcc cggcgggagg tgacggcgct cacctacagc atcagcaggg	1320
tgtcggagtt cctgtccccg atcaagacgg tgggcctgtc cagggaccgc gagagggacg	1380
cggccaagat caaggcgctg ctccagaagg gggacctgtg catctgtccc gagggaacca	1440
cctgccgcga gccattcctt ttgcgattca gcgcactgtt cgcgagactg actgacaaga	1500
tcgtgccggt ggcgctgtgc acgaagggga gcactttcca cgggacaacg gtgcgcgggt	1560
ggaaggggtt ggatccattc ttcttcttca tgaatccctt cccgacgtac gaggtgacgt	1620
ttctccagca gctgccgccg gagctgacgg tccagcaggg cggcaagagc gcgatcgagg	1680
tggcgaatca catccagcgg gtgatcgcg acacgctggg attcgagtgc acgaacttca	1740
cccgaagga caagtacggg atgctggcgg ggaacgacgg caccgttccg gagaggcgga	1800
ggagcggcgg cggaggtggc ggcgcccca aggatctccc gaaggacgag caagcggccc	1860
aggggcaaga ccaggtaaat ggctgtagct aaaattgatt tgcgaggatc gtggatccat	1920
gccatgggtc gggcttcacg cgctccttcg ttataagtga atgtaagcta agctactggc	1980
ttgctctct	1989

<210> 101
 <211> 1554
 <212> DNA
 <213> Selaginella moellendorffii

<400> 101	
atggcagagt ttccacccat cgaaaagtgt ggcaccacaa tcagccgcga gagcgaatcg	60
gtggcctccg atctggatgg aacgctgctg cgcgaggagga gttccttccc atactttttc	120
ctcaccactt tcgagacgtc ggggatcttg cgctcctggc tgctcctcct cgtctcgcct	180
ctcgtctggc ttctctatca cttcatctcc gaggccgctg gaatcaagct cctcatcttc	240
gcttccttg cgggtgttcc agtctctcag atcgaatgca tcgcccggag tgttcttccg	300
cggttttatc tccaggatct caaccccgag acgtggaagg tcttctcgtc ctttggaag	360
aagattgttg tcacagcgaa tcccaggatc ttggtggagt cgctctgcaa ggaattccta	420
ggtgctgacg aggtgattgg aaccgagctg gaggtggata accgaggacg agcaaccggt	480
ttcgtcaaga gccccggagt tctcgttga gagaacaaag ccaaggcctt gctccaggcc	540
tgtgatcccg agcagctccc agacgttga cttggcgatc ggcagaccga tttcccgttc	600
atgaagcttt gcaaggtacg taacaaaaga tcgtccctcc ttcccgcgac ttagcagttc	660
tcacgctggc ttctctcaat ccaggaagga tacgtcgtcc cccaagcaa caacgtggag	720
ccagttcaa agtcccagct tctgaagccc atcatctttc acgatggctg tctcgtccag	780
cttccgacgc cactcgtggc tctcgtcacc ctctctgga tccccgtcgg cttcgttcta	840
gcctgcctcc gaatcgccgc aggctcgtg cttcccatgt caatcgtgta ccatgccttc	900

pctau2012001598-seql

tggctgctgg	gagtccgagt	gaccgtgaaa	ggaacgccac	ctccgagagc	caccaagaac	960
tccccgaggcg	tcctcttcat	ctgctcgcac	cggacgctgc	tagatccaat	cttcctgtcg	1020
tgtgcgctgg	gccgaggcgt	ctccgccgtg	acttactcgg	tgtcccgact	gtccgagatc	1080
ctcagcccga	tcaaaacagt	ggctctctcc	aggaaccgct	acaaggacgc	ggctcagatc	1140
cgggagctcc	tgaaagaagg	cgacctggta	atctgccccg	agggaaacgac	gtgtcgggag	1200
ccattcctgc	tgcgtttcag	cgcactgttc	gcggagctgg	ccgacgatat	cgtccccgtc	1260
gccatggaca	accggatgag	catgttccac	ggcaccgctg	ctcgtgcctg	gaaggggatg	1320
gacccattct	acttcttcat	gaatcccagc	cccagctacg	aggtaacgtt	cctccagcag	1380
ctcccgaag	agctcacttg	caagggcggc	aagagctccc	acgaagtcgc	aaaccatatc	1440
cagcgagctc	tggcgggagt	gctagatttc	gagtgtacca	atctgacgag	gaaggacaag	1500
tacaaggcgc	tcgtcggcac	ggacggatcg	gtggagacca	agaagtcaac	gtga	1554

<210> 102
 <211> 1554
 <212> DNA
 <213> Selaginella moellendorffii

<400> 102	
tcacgttgac	ttcttgact ccaccgatcc gtccgtgccg acgagcgcct tgtacttgtc 60
cttcctcgtc	agattggtac actcgaaatc cagcactccc gccagagctc gctggatatg 120
gtttgcgact	tcgtgggagc tcttgccgcc cttgcaagtg agctcttccg ggagctgctg 180
gaggaacgtt	acctcgtagc tggggctggg attcatgaag aagtagaatg ggtccatccc 240
cttcaggca	cgagcagcgg tgccgtggaa catgctcatc cggttgtcca tggcgacggg 300
gacgatatcg	tcggccagct ccgcgaacag tgcgctgaaa cgcagcagga atggctcccg 360
gcacgtcggt	ccctcggggc agattaccag gtcgccttct ttcaggagct cccggatctg 420
agccgcgtcc	ttgtcgcggt tcctggagag agccactgtt ttgatcgggc tgaggatctc 480
ggacagtcgg	gacaccgagt aagtcacggc ggagacgcct cggcccagcg cacacgacag 540
gaagattgga	tctagcagcg tccggtgcga gcagatgaag aggacgcctc gggagttctt 600
ggtggctctc	ggagggtggc ttcctttcac ggtcactcgg actcccagca gccagaaggc 660
atggtacacg	attgacatgg gaagcagcga gcctgcggcg attcggaggc aggctagaac 720
gaagccgacg	gggatccaga ggagggtgac gagagccacg agtggcgctc gaagctggac 780
gagacgacca	tcgtgaaaga tgatgggctt cagaagctgg gactttggaa ctggctccac 840
gttggtgctt	ggggggacga cgtatccttc ctggattgag agaagccagc gtgagaattg 900
ctaagtcgcg	ggaaggaggg acgatctttt gttacgtacc ttgcaaagct tcatgaacgg 960
gaaatcggtc	tgccgatcgc caagtccaac gtctgggagc tgctcgggat cacaggcctg 1020
gagcaaggcc	ttggctttgt tctctccaac tagaactccg gggctcttga cgaaaccggt 1080
tgctcgtcct	cggctatcca tctccagctc ggttccaatc acctcgtcag cacctaggaa 1140

pctau2012001598-seq1

ttccttgacag agcgactcca ccaagatcct gggattcgct gtgacaacaa tcttcttccc	1200
aaaggacgag aagaccttcc acgtctcggg gttgagatcc tggagataaa accgcggaag	1260
aacactccgg gcgatgcatt cgatctgaga gactggaaca cccgccaggg aagcgaagat	1320
gaggagcttg attccagcgg cctcggagat gaagtgatag agaagccaga cgagaggcga	1380
gacgaggagg agcagccagg agcgcaagat ccccgacgtc tcgaaagtgg tgaggaaaaa	1440
gtatgggaag gaactcctcc cgcgacgag cgttccatcc agatcggagg ccaccgattc	1500
gctctcgagg ctgattgtgc tgctacactt ttcgatgggt ggaaactctg ccat	1554

<210> 103
 <211> 1757
 <212> DNA
 <213> Selaginella moellendorffii

<400> 103	
atggagcacc cgtatgacgt gatcgacaat tgctcggccg acaaccgcga gaagcagagc	60
gtggtagccg acttgatgg ctcctgcgtt cgcagccgga gctccttccc ttacttcatg	120
ctggtggctc tggaaggagg cagcctcctg agaagcgtgc tcctcatggt catggcccc	180
gtcgcgtggc tgctctacca cttcgtctcg gaggcagcag gcatcaaagt cctcatcttc	240
ttctcctgcg ccggccagag agctcgggac atcgaagcgg tggctcgcgc agttctcccc	300
aagttctact ccgaggacgt ccacagctcc acctggagag tcttctcgtc gtgcggcaag	360
cgctacattc tcacggccag cccgagactg atggtggagc actttgccac gacatacctg	420
ggcgtgaca aggtgatagg gacggagctg caagtgacaa ggagcgggta cgctacgggg	480
ttcgtcaaga gccccggcgt tttggttggc gtcaacaaga ggcgggcatt gaaggccgtg	540
tttggtagagg aggtgccaga cctggggatt ggcgatcgca gctctgatta tccgttcatg	600
gctttttgca aggttggtag ttctcgttct ctctgagctc tcttcttttt cgctgttagt	660
cagcaggagc atatgcatat attgcagtgc atgctatgct agtctgtgat ctttggagca	720
actcgtcgcc tggctttctt tttctttgat tttgttgtat ggttggttgt gtcaagtccg	780
cgttgtcttc gttgtctccg tcgtcttctg gctttatttc atccacggat gtggaacttc	840
cgagtctcac cagaactaat aaagtgcct ctcttttttc cttttcggct ttgcaggaag	900
cttacattgt tcctagccag cctgccgacg ccgttccagc tcacaaactc gcgcgcaaga	960
ttatctttca cgacggcagg ctggtccaga ggccgacgcc gttcgtggcg ctggtcacct	1020
ttctgtggct acccgccggc ttctccttg cttcttccg catcctgggt gcgcttccat	1080
gcccggtcga gtacgtcccg ctgctctaca aggtgatcgg cgtccggctc atcgtcaagg	1140
gcaaagtccc gccaccttcc aagaacggga tcctcttcgt ctgcacgcac cggacgtca	1200
tcgacccgat aatggcctcg gcggcctcgg ggagcgacgt gaccgcgctc acctacagca	1260
ttagccgggt gtcggagggt ctctcgccca tcaagacggt gccgctgtcg cgggaccggg	1320
agagggacgc ctcgaacatc aagcggctgc tgctgaagga gaacttggcg atctgtccc	1380
agggaaacgac gtgccgggag ctttctctgc tgcggttcag cgccctgttc gcggagctgt	1440

pctau2012001598-seq1

ccgagaacat	cgtccccggtt	gcgacgtcga	ccaaaacgag	catcttccac	gggacgacgg	1500
tgcgggggctg	gaaggggatg	gacccgttct	tcttcttcat	gaacccgttc	ccggtgtacg	1560
aggtagacgtt	cctcgagcgg	ctgccgcacg	atctgagcgt	ccaggggggc	aagtcggcga	1620
tcgaggtggc	caactacatc	cagaggggtgc	tggcggacac	gctgggattc	gagtgcacga	1680
actttacgcg	caaggataag	tatgggatgc	tcacgggcag	cgatggaacg	gtgcctgcc	1740
ggaagaagtc	cttttga					1757

<210> 104
 <211> 2256
 <212> DNA
 <213> Selaginella moellendorffii

<400> 104	
agcttctgag	ccagcagtac aggggctgga cggacttgaa aggcattttt ctccagttcc 60
atttcaagtg	agcacgttgg aagcatttat tctcctctgt ggcgttcttt ctaactcgga 120
gttgtcccag	gattcgacat ggccgactta aagtactcgc ggctcttccg gttcggcctc 180
gtggagaatt	gctcatctga aggacgtgga atgcacacag ttgcggcgga tttggatggt 240
acactgctca	gaggcagaag ctcgtttcca tacttctctgc tgggtggctct tgaaggaggt 300
agcctcttca	gagcagttgt tctcggactt ttggccccta tcgcctggct gctttaccac 360
ttcgtctcgg	aggcagcagc aattcaactt ctcatctatg tctctttttc ggggcagaag 420
gtgtccgaga	ttgagaaggt ggcaacggcg gttctgccga agttttactc ccaggacgtc 480
cacagtgatg	ctttcagagt tttctcgtcc tgcgggaaga agtatgtcat cactgccaat 540
cctcggatca	tgggtggagta tttctgcaag gccttccttg gcgcggacaa agttattggg 600
acggaaatcg	aagtggacga tgatggacgt gctacaggct tcgtcaagtt cccaggtgtg 660
ctcgtgagtc	aaaacaagag gaccgccttg aagctcgagt tcgtggacgg ggagcttccc 720
gagatcggac	tcggcgatcg agaaaccgac ttcgccttca tgtcgttttg caaggtaagc 780
gcactctttc	catatcatac acatattttg catttcgagc atataatact tagctctgcc 840
tctttcgtgc	tgctttaact gttgaggggt gccaggattt gcgtgacact ccagaatcag 900
ttacgagaag	cgagttttct aactctggtt gtaaccgtgt gatacaggaa gcttacgttg 960
ttcccaagag	gaaggttgac gcagtggacg gcagcgtctt ggcgagcgc gtgatcttcc 1020
atgatggtag	gctggtccag ctcccaacgc cgttcgtggc gctcgtcact ttcctctggc 1080
tccctatcgg	cttcgcgctc gcagtgtctc ggatcggcct cacctcccgt tgccccgaa 1140
agtacatggg	aatgatctac gccatcctcg gcgtccgccg catcaccaaa ggcaaagttc 1200
ctcgcaaaca	agacggcagg gggctcctct tcgtgtgcaa ccacaggact ctggtggatc 1260
ccatcttcgt	gggattggcc gccgagaggg acgtgaccgg gctcacttac agcatcagcc 1320
gggtgtccga	ggtgctatcg cccatcgtca cggtcgggct gacgagagac aagcagcgcg 1380
actttacgaa	gatgaagtcg ctgctggaga gggggagggg cctgtgcctg tgtccggagg 1440

pctau2012001598-seq1

ggacgacttg	ccgggagcct	ttcttgctgc	gattcagcgc	tctcttcacc	gagctcacgc	1500
acaggatcgt	cccggcggcc	gtgaagacga	ggcagagcat	gtttaacggg	acggcgggtgc	1560
gggggtggaa	ggggatggat	ccattcttct	tcttcatgaa	tccattcccg	acgtacgagg	1620
tggagttttt	ggaggagctc	cccgtggaga	tgagcgtcca	aaagggcggc	aagagctcgt	1680
tcgagacggc	caatttcac	cagaagatgc	tgggcgacaa	gctgggatac	gagtgcacga	1740
ccttcacccg	caaggacaag	tacctgatgc	tcaatggcag	cgatggaaca	gtccccacga	1800
aataggaagg	ggttgaaaa	gaaaacaaga	tggaaaccgtt	gactggtggt	gacttgaatt	1860
gactagagtt	gacggttgaa	tttttttttt	ttggtgttgg	ctggctatct	agcttgcaact	1920
agtaaagtct	aaaggttttt	caagcaggca	ggcgatttta	atccttccac	atgggtggga	1980
cagaaaatat	tgacttcgtt	catttaaagg	gggcaaaggt	tttagggcac	ggaggggtgg	2040
tctagggcac	ggaagcgagc	tggcatcctc	acacaggcag	catggcgtcc	ggccggcggt	2100
acggcgctcg	tgtctcgcat	tgccagcgaa	atgccactcc	caacctccgc	caccttttct	2160
ctctcctgga	gttcttctgc	cacctttccg	aactgaacaa	atgtgttaga	ggaatcaaag	2220
atccaacaca	ccgaaacttc	gtttttgttt	tcttct			2256

<210> 105
 <211> 2044
 <212> DNA
 <213> Selaginella moellendorffii

<400> 105	
atgtcgtctg	acggcggcaa ggtcacggat tctttcctgg tagatgctcc attcgagccg 60
gtggagcgat	gctcgtcggc ggggcgtgag aagcagacgt gcgtggcgga tttcgacggt 120
actctcatcc	gcggccggag ctccttcccg tatttcatgc tgggtggcgct agagggagga 180
ggcctcatcc	ggagcgtgat tcttggcctc ttcgcgccgc tcgcctggat cctctaccat 240
ttcgtctcgg	aatccgccgg catccgcctc ctcatcttca tctccttcgc cgggatgaag 300
gccaagcaga	tcgagtcggt gtcgcgcgcc gtgctgccca aattctactc ggaggacgtc 360
cacagcgagg	cctaccgagt gttctcctcg tgcggcaagc gtgtggtggt gaccgccaat 420
ccccggatca	tgggtggagca cttcgccaag acgtgcctgg gggcggacaa ggtgctgggg 480
accgagattg	aggtgacgag gagtggctac gcgacgggat tcctcaagcg gccgggagtt 540
ctcgtggggg	tcaacaagcg caaggctgtg aggcaggaat tcggcgagca tttgccccat 600
gtcggcattg	gcgatcgagc cacggacttt gctttcatgg cgctctgcaa ggtacttata 660
tcgaatccaa	tcccgctctg cttggtgggt ggagtgggtt cttagtcttc gatcgagctc 720
gatgcattag	atcaacactc tgggcattgt tttcgatcga tgatccatcc ctgcaattgc 780
tgcattgagt	gatgattctc ttgtttcttt tgtttatctt ttttccttgt ctcatgcaca 840
cattcttaca	agcgacttag cactagccta aaagaatctc ttccaagcag cagtagaaac 900
ttctttccag	tcaactaacg aaccatttg ctgaccacta tcaggaagca tacgtcgtcc 960
cacctccaa	agtggaggcg ctgagccgag acaagctcgt caagcccgtc gtcttccacg 1020

pctau2012001598-seql

acggccgcct	cgtccagcgg	cccacgccct	tgaccgccct	cgtcaccttc	ctgtggctgc	1080
ccgtcagctt	cgtgctggcg	gtcttcggga	tcatggtcac	ggtaccctgc	ccgcgcgagt	1140
acgtgaccat	cgtctacaag	ctcgtcggca	tccggctcat	cgtcaagggg	ccgatcccgc	1200
cgccaagaa	gcggggcgag	tccggcgtgc	tcttcgtctc	cagccacagg	acactgtgcg	1260
atccggtgtt	cgtgggggtc	ggggcccggc	gggaggtgac	ggcgctcacc	tacagcatca	1320
gcagggtgtc	ggagttcctg	tccccgatta	agacgggtggg	cctgtccagg	gaccgcgaga	1380
gggacgcggc	caagatcaag	gcgctgctcc	agaaggggga	cctgtgcatc	tgtcccgagg	1440
gaaccacctg	ccgcgagcca	ttccttttgc	gattcagcgc	actgttcgcg	gagctgactg	1500
acaagatcgt	gccggtggcg	ctgtgcacga	aggggagcac	tttcacggg	acaacggtgc	1560
gcgggtggaa	ggggttgat	ccattcttct	tcttcatgaa	tcccttccc	acgtacgagg	1620
tgacgttcct	ccagcagctc	ccgccggagt	tgacgggtcca	gcagggcggc	aagagcgaga	1680
tcgaggtggc	gaatcacatc	cagcgggtga	tcgcggacac	gctgggattc	gagtgcacga	1740
acttcacccg	caaggacaag	tacgggatgc	tggcggggaa	cgacggcacc	gtgccggaga	1800
ggcggaggag	cggcggcgga	ggtggcggcg	ccgccaaagga	tctcccggag	gacgagcaag	1860
cggcccaggg	gcaagaccag	gtaaatggct	gtagctaaga	ttgatttgcg	aggatcgtgg	1920
atccatgcca	tgggtcgggt	ctcacgcgct	ccttcgttat	aagtgaatgt	aagctaagct	1980
actggcttgc	tctctcgatt	gactcatcag	gaatagatcg	atcgggtgga	gagcatcgtg	2040
atcg						2044

<210> 106
 <211> 1715
 <212> DNA
 <213> Oryza sativa

<400> 106	
actcccacac	catcagcgag cacctagcgg tcgcatcgct agctgccatc tcttggtgag 60
atccatcgag	acgcccggca cgacgacgat ggtgtcgcgg aggttcaagc cctgggagga 120
gtgcagctcg	gatgggaggt cggagcagac ggtggccgcc gacttcgacg gcacgctggt 180
caggtcgcgg	agcgccttcc cgtactacct cctcgtcgcg ctcgaggccg gcagcgtgct 240
ccgcgccgtc	gtgctgctgc tgtccgtgcc gttcgtctac gtgacctaca ttttcttctc 300
cgagtcgctg	gccatcagca cgctggtgta catctccgtg gcggggctca aggtgaggaa 360
catcgagatg	gtggcgcggt cggtgctgcc caagttctac gcggaggacg tgcacccgga 420
gagctggagg	gtgttcaact ccttcggcaa gcggtacatc atcacggcga gccccaggat 480
catggtggag	cacttcgcca agacgttcct cggcgccgac aaggtggctg ggacggagct 540
ggaggtcggc	aagaacggca aggccacggg gttcatggtg aagcccggag tgctcgtggg 600
cgaccacaag	aggcaggccg tcgtcaagga gctacgcgac gcggtgcccg acgtcggctt 660
gggcgacagg	gagacggatt tcgacttcat gtccatatgc aaggaggcct acctcgtgac 720

pctau2012001598-seq1

atcaaggaag	tacagcgcg	tgcccaagaa	ccagctgctg	agcccactca	tcctccacga	780
cggccgcctc	gtgcagcgtc	cgacgccgct	ggtggcgctc	gtcaccttcc	tgtggatgcc	840
gttcggcttc	gcgctcgcgc	tcctccgcgt	gtacgtcaac	ctgccactcc	cggagcggat	900
cgtctttctac	acctacaagc	tcatgggcat	ccgcctcatc	gtgaagggca	acccgccgcc	960
tccccccaag	aagggacatc	ctggcgctcct	cttcgtctgc	aaccaccgca	ccgtgctcga	1020
cccggttgag	gtcgccgtgg	cgctgcgccg	caaggtcagc	tgcgtcacgt	acagcatctc	1080
caagttctcg	gagctcatct	cgccgatcaa	ggccgtcgcg	ctgtcgcggg	agcgcgagaa	1140
ggacgccgag	aacatccggc	ggctgctgga	ggagggcgac	ctggtgatct	gccccgaggg	1200
caccacctgc	cgcgagccgt	tcctgctgcg	gttcagcgcg	ctcttcgccg	agctcaccga	1260
ccgcatcgtg	ccggtggcga	tcaacaccaa	ggagagcatg	ttccacggct	ccaccgtgcg	1320
gggcttcaag	ctcatggacc	cctacttctt	cttcataaac	ccgcggccga	cgtacgagat	1380
caccttcctg	aaccagctgc	ccaaggagct	cacttgcagc	ggcggcaagt	cgcccatcga	1440
ggtggccaac	tacatccaga	agacgctcag	cggccagctc	ggcttcgagt	gcaccgccat	1500
aacgcgcaag	gagaagtaca	gcatactcgc	cgggaccgac	ggccgcgtcc	cttccaagaa	1560
caaggagaag	gaaaagaact	agcgagtcac	agcctcctat	cgatcggagt	actccgtatt	1620
cgcccagccg	actagtgtca	ctcatggctt	ctatctattg	tttacgattt	attgttgttt	1680
ttcaagaagt	tggtaaatta	catattgttt	ccgag			1715

<210> 107
 <211> 1521
 <212> DNA
 <213> *Oryza sativa*

<400> 107	
atggcggcgg	cgccggcgcc
gcgccggcgg	gacgggtggc
tcgtcggcgt	ttccgtacta
ctggcgctgc	tgctcgccgc
gccgccatcg	ccctgctcgt
gccgtggcgc	gcgccgtgct
gcggtgttcc	acggctgcgc
gtgggggagt	tcgtgcgcga
ttgcctccg	gcaagcggtt
cgcgaggtcg	tcgagaggct
cgcgagagcg	accacgactt
aagcgcgcg	cgcgccgcgc
cgctcgtcc	gcccccga
ggcttcgccg	tgccgctcct
	ccgcgtgttc
	ctcaacctcc
	ccgtcccggc
	gcgcctcgtc
	ggcggtgac
	ggtgtgacg
	gctggtgtcg
	ccggcagcta
	cctccgcgcg
	acgtcggcgt
	ctccgaggcc
	cgacgtcgag
	cgacacgtgg
	ggccgtcatg
	gctcgagacg
	cgagaagaag
	cctcggcgac
	gcccagaac
	ccacgacggc
	cctccccgtg
	gcgctcgtc
	ggttcgttc
	ccgtcccggc
	gcgcctcgtc

pctau2012001598-seq1

cgccacacct	accgcctcac	cggcatccgc	ctcgccgtgc	gcggggcgcc	gccgccgccc	900
ccgcgccccg	ggacgccggg	ctccctcctg	gtgtgcaacc	accgcaccgc	gctggacccc	960
atcatcgtgt	ccatcgcgct	gggccgcccc	gtgacgtgcg	tgacgtacag	cgttagccgg	1020
ctgtcgacgg	cgatctcgcc	gatccgcgcg	gcggcgctga	cgcgggaccg	cgcggcggac	1080
gcggcgcgca	tcgcggcgct	gctggaggag	ggggacgtgg	tgggtgtgcc	ggaggggacg	1140
acgtgccggg	agccgtacct	gctgcggttc	tcggcgctgt	tcgcggagct	gacggcgagg	1200
atcgtgccgg	tggcgggtga	ggcgaggcag	gggacgtact	atgggtcgac	ggcgaggggg	1260
tggaagtcc	tggaccctta	cttctttctac	atgaaccgcg	ggccggggta	cgaggtgacg	1320
ttcctgccgg	cgctgaggcc	ggaggagacg	tgcgtcgccg	gcgggaggag	cgcggtggag	1380
gtggcgaacc	acgtgcagag	ggtgatcgcc	aaggagctag	ggttccaatg	caccacgctc	1440
accaggaagg	acaagtacat	gaagctcgcc	ggcaacgacg	gccgggtcgc	cgccgccgcc	1500
gacaagccca	aggccaattg	a				1521

<210> 108
 <211> 1755
 <212> DNA
 <213> Zea mays

<400> 108	
aaccagaca	ggcaccgtac caccgtgttc cccgtcgcca tggccgcgag gttcccccg 60
gtgtcgtcct	acgacgttc ggcgcggttc cgccgcacgg ccgcggcgga cctggacggc 120
acgtgctgg	cgtcgtcgtc ggcgttcccc tactacttcc tgggtggcgct ggaggccggc 180
tcctgcgcgc	gcgcgctgct gtcctccttg gccgcgccgc tgctgctggc gctctacacg 240
ctggcttccg	aggcgggcggc catcgcgctg ctggcgttcg tcacgttcgc ggggctccgg 300
gtgcgcgacg	tcgaggccgt cgcccggggc gtccctccgc gccactacgc cgcaggcgctg 360
cgcgccgaca	cctggggccgt gttccgcgac tgcggcgccg cgcgaggagg cgtcgtcacc 420
gcgtcccccg	ccgtcatggt cgccgacttc gtccgcgagt tcctcgggcg cgagctcgcc 480
gccaccgagc	tcgagacctg ccgcgcgctc ggcgacgcct ggttcacggg caggatcaag 540
gccgtgctcg	tcgggggagag gaaggccgag gtcgtgcgcc ggctcttcgc cgccggcgac 600
ctgcccgcag	tcggggtcgg ggaccgcgag agcgaccacg acttcatggc catctgcaag 660
gaggcttaca	tgggtgcctc ggaccggcgc gcgccgcgag cgcccgcgga cgcgctgctg 720
tcccgcgccg	tcttcacga cggccgcctc gtgcggcgcc cggaaccggc gcacgcgctc 780
ttcgcgctgg	cctacctccc tctcggcttc ctgctggcgc tgttcgcgct gctcttcaac 840
ctcgccatgc	cggcgcgccct ggtccgccac acgtaccgcc tgacggggat ccggctccgc 900
gtgcgcggga	ctccgccgcc ggccccggcg ccgggcgcgc cggggtcgct gctggtgtgc 960
aaccaccgca	cggcgctgga ccccatcatc gtgtcggtgg cgctggggcg gcccggtggc 1020
tgcgtgacct	acagcgtgag ccggctgtcg acggcggtgt cgccggtccg cgcggtggcg 1080

pctau2012001598-seq1

ctgtcgcggg	accgggaggc	cgacgcggcg	cgtatagcgg	cgctgctggc	ggagggcgac	1140
gtggtggtgt	gccccgaggg	gaccacgtgc	cgggagccct	gcctgctgcg	cttctcggcg	1200
ctgttcgcgg	agctcacgga	ccggatcgtg	cccgtggcgc	tggacgcgcg	gcaggggacc	1260
tactacgggt	ccacggcgag	ggggtggaag	tggctcgacc	cttacttctt	ctacatgaac	1320
ccgcgccccg	ggtacgacgt	caccttcctg	ccgccgctgc	ggccggagga	gacgtgcggc	1380
gccggcggga	ggagcgccgt	cgacgtggcc	aaccacgtgc	agacgctcat	cgccaaggag	1440
ctcgggttcc	ggtgcaccaa	gctcaccagg	aaggacaagt	acatgaagct	cgccggcaac	1500
gacggcacgg	tcggggggcag	gcctgagaat	gagaagggcg	ccgccgatga	tagtgctgcc	1560
acggccacca	ccaagaagct	tgtataaaga	gactgtattt	atagctagct	ggtcgactgt	1620
atcggagcag	catccattat	gtatttatgt	acgtatgcta	atgacgacaa	cattagtgtg	1680
atgtgatgga	ttcgtgaaca	tagaagatta	caagcttaac	ttcatgttta	ccatgaccaa	1740
aaaaaaaaaa	aaaaa					1755

<210> 109
 <211> 1852
 <212> DNA
 <213> Zea mays

<400> 109						
gcgcgcgtat	attaactcca	cctcgccgcc	ccactccact	gccaagcaac	tcgcacgcag	60
cagtacgacg	tctctcctct	gggtctgggg	ccgagacacc	gagcacgtac	taccagcaag	120
atggtggcgt	ctcccagatt	caagcccatc	gaggagtgtc	gctcggaggg	gcggtcggag	180
cagacggtgg	ccgccgacct	ggacggcacg	ctgctcatct	ccaggagcgc	gttcccctac	240
tacctcctcg	tggctctcga	ggccggcagc	gtcctccgcg	ccgcgctgct	gctcctgtcc	300
gtgccgttcg	tctacgtcac	ctacgccttc	ttctccgagt	cgctggccat	cagcacgctg	360
gtgtacatct	ccgtggcggg	gctcaagggt	cgcarcatcg	agatggtggc	gcgctccgtg	420
ctcccaggt	tctacgcggg	cgacgtgcac	ccggagagct	ggaggggtgtt	cagctccttc	480
ggcaggcgct	acgtcgtcac	ggccagcccc	agggtcattg	tcgagccctt	cgccagggcc	540
ttcctcggcg	ccgacaaggt	cgtcgggacc	gagctggagg	tcggcagggg	cggcaaggcc	600
acgggcttca	tggccaggcc	aggcgtgtc	gtcggcgacc	acaagaagaa	ggccgtcgtc	660
aaggagctcg	gcgacgcgct	gcccgcgtc	ggcatggggg	ataggagagc	cgacttcgac	720
ttcatgtcca	tctgcaagga	ggcctacctg	gtgacgtcaa	ggaagtacag	cccggtgccc	780
aggaaccagc	tgctgagccc	gctgatcgtg	cacgacggcc	gcctcgtgca	gcgcccgcgc	840
ccgctcgtcg	cgctcgtcac	cttcctctgg	atgccgttcg	gcttcgcgct	ggcgtcatg	900
cgcgtgtaca	tcaacctgcc	gctgcccgcg	cgcctcgtct	actacacctc	caagctcatg	960
ggcatcaggc	tcgtcgtcaa	gggcaccccg	ccgccgccgc	ccaagaaggg	ccacccgggc	1020
gtcctcttcg	tctgcaacca	ccgcaccgtg	ctcgaccccg	tcgaggtggc	cgtggcgctg	1080
cgccgcaagg	tcagctgcgt	cacctacagc	atctccaagt	tctccgagct	catctcgccc	1140

pctau2012001598-seq1

atcaaggccg	tcgcgctgtc	gcgggagcgc	gacaaggacg	ccgagaacat	ccgccgcctg	1200
ctggaggagg	gcgacctggt	catctgcccc	gagggcacca	cctgccgcga	gcccttcctg	1260
ctgcgcttca	gcgcgctctt	cgccgagctc	accgaccgca	tcgtgcccgt	ggccatcaac	1320
accaaggaga	ccatgttcca	cggctccacc	gtgcgcgggt	tcaagctcat	ggatccctac	1380
ttcttcttca	tgaacccgcg	cccgcgtac	gagatcacgt	tcctcaccca	gctcccaaaa	1440
gacctcacgt	gcagcggcgg	caagtcgccc	atcgaggtgg	ccaactatat	ccagaagacg	1500
ctcagcggac	agcttggctt	cgagtgcacg	tccatcacgc	gcaaggagaa	gtacggcatg	1560
ctcgctggca	ccgacggccg	tgtcccgtcc	aagaacaagg	agaaggagaa	ggataagaac	1620
taacggccaa	cgtcgctctt	cctgcttgcc	cagaagactg	ccgtttaatt	actgctcaaa	1680
tttattccaa	tttattgttc	cttgtcgaag	agttggccag	tgtgatgtgg	gccgtgtact	1740
agtacaagct	tgatgatcat	tagtatgtct	acaattcttc	ctttccggaa	gaaggaggga	1800
agagttgata	gagccgtctt	aatacaaagt	ttcgccaaaa	aaaaaaaaaa	aa	1852

<210> 110
 <211> 1704
 <212> DNA
 <213> Zea mays

<400> 110	
caccaatata	atcaccatth ctgttcacgc gatcgagagg caaaattaag catgggtgtcc 60
cgaagggtca	agccgatcga gctgtgcgac tccgaggggc ggtcgcgcca gacggtcgcc 120
gccgacctcg	acggcacgct cctcctgtcc cgcagcgcgt tcccgtacta cctcctcatc 180
gcgctggagg	ccggtagcct ccttcgcgcg gtggcgcttc tcttgtccgt gccctcgtg 240
tacctcacct	acgtgaccgt gtccgagacc ctggccgtgc gcgcgttcat ctacgtcgcc 300
atggccgggc	tggaggcgag ggacatcgag gccgtcgccc ggtccgtgct cccagcttc 360
tacgccggcg	acgtccaccc ggagggctgg cgcgtgttcc gctcgttcgg caggaggtgc 420
gtcgtcaccg	ccagcccgcg cctcatggtg gagccgttcg ccaaggcggt tcttggcgcg 480
gacgtggtga	tcgggacgga ggtggaggtg gccagtcgg ggaaggccac ggggttcgtt 540
gccggggccc	gggtgctcgt cggcgagcac aagaggcggg cgggtggcgag ggagttcggc 600
gactcgctgc	cggacgtcgg catgggggat cgcgagagcg acttcgactt catgtccatc 660
tgcaaggaag	cgtacaccgt aacacgacag aagtaccgcg cgctgccccg ggatcagctg 720
cagggccggg	ccatcctcca cgacggccgc ctggcccgcg gcccgacggc gaccaacgcg 780
cttctcactt	tcctgtggat gccgctcggc ttgcgcctcg cgctgggtgc cgcgcatctg 840
cacctgtctc	tcccgggtgc cgccctcgcc tacgcctaca agcttacggg cgtcaagctg 900
gtgggtgcgc	gcaacaggcc gccgcccgtc aagaagaaga agggcgacca gctaggggtt 960
ctcttcgtgt	gcaaccaccg tagcacgtc gacgccgtcg cgggtggcggt cgcgctgggc 1020
cgcaagggtg	gatgggtcgt cacagacggc ggtgcgggtg cctccaggtt ctcggagccc 1080

pctau2012001598-seq1

gtagtgtcgc	ctatcatgac	aggcgtgccg	ctgcccgtgc	ccaccagcct	cgagggcgac	1140
gccgacgcgg	acgccgcccc	ccgcatccgg	cggctgctcg	aggagggaga	cgacgtcgtc	1200
gtcttccccg	agggcgccat	ctgccgcgag	ccgttcctgc	tgcggttcgg	cgcgctcttc	1260
gctgaggtca	ccgaccgcat	cgtgcccgtg	gccatcgaac	ctagggagag	catgtttcac	1320
gggtccacgg	cgcgcgggct	cagagggatg	gatccctact	tctttttcat	gaaccacacg	1380
ccgacgtacg	aggtcacgtt	cctgaaccag	cttcacgggg	agctcacctg	cggcggcggg	1440
aggtcgccgg	tcgaggtggc	cagctacgtc	caggaggtcc	tggccgcgca	gcttggattc	1500
gagtgcacct	gatgaagcag	caagtgcaaa	agtggcatgg	tgacggttgc	gtggcgttca	1560
aaggtggact	aggaacaca	gagcatgcaa	ccaactgtca	attgccgtgt	gttgagagtg	1620
gtgatcagtt	catggtttca	cgtttgctgt	tgatgtctgt	gataagaacc	atatgtatct	1680
ctacaactat	taagagcggt	gtgt				1704

<210> 111
 <211> 3202
 <212> DNA
 <213> Brassica napus

<400> 111

ccgcaaagat	ccgaaccgaa	cctgaaccga	aatttagaaa	taccggaatg	gggctgaaat	60
ctttaaaccc	gaaaaccgga	aacccgatta	gatccgaacc	gaaccggaat	gggtaccgga	120
atgccacccc	ctaactcat	ctataaaata	aatatatcca	tatataacat	ttcaaatttc	180
gaaatattat	tatttttgta	taatttgtat	tactaaagtt	ttcaaaaatt	tctacaattt	240
ttttaaaaaa	taaatatcta	atcgtaaaat	cattagtttc	ttatatatct	acaaatttta	300
taaatattgt	ttaggctaaa	tttttgataa	ttatatagtt	tttattaatt	ttatacaaat	360
tgatttaata	tatattaaat	ctatattaaa	tattagtaaa	aaaatagtaa	aatatataat	420
atttaataaa	tttattttta	aatataaatt	agattttaaaa	aattccttgc	tgaaaaacac	480
ctagtagtat	atcataaagg	atgtttattg	cgatgtgtaa	ttttaaatct	aatgggggac	540
aaagtgagaa	tcagggttat	aattgttaga	aaaggatagt	taaaggataa	aatcgtaaag	600
aggtcatatt	gattggagtg	ataactaata	gttgagcggg	taaccacttt	gaatgagcgc	660
atttaaaaca	gagaccctt	ccctctcttt	ctttcccaaa	ttgacaaaaa	gactcaagag	720
agaaagaaag	aaaatgtctc	cggcgaagaa	gaagagcaca	agctttcctc	cgataagcga	780
gtgtaaaagc	ggagagtatg	attcgattgc	tgcggatctt	gacgggactc	ttctcctctc	840
gcgaagctcc	ttcccttatt	tcatgctcgt	cgccatagaa	gcaggcagtc	tcttccgtgg	900
attgatcctc	cttctatctc	tcccgatcgt	catcatcgtc	tacctcttcg	tctccgaagc	960
tctcggcatc	caaactctca	tctacatctc	cttcgccggc	atcaaaaatcc	gcgacatcga	1020
actcgtctct	cgcgccgttc	ttccacgggt	ctacgcggcg	gatgtgagga	aggacagttt	1080
cgaggtgttt	gataaatgca	agaggaaagt	ggtggtcact	gcgaacccca	tagtgatggt	1140
tgagcccttc	gtgaaggatt	acttgggagg	agacaaagtt	ttaggaacag	agatcgaagt	1200

pctau2012001598-seql

gaatcctaaa	acgatgaaag	ccactggggtt	tgtgaagaag	cctgggtgttc	ttgttggcga	1260
tttgaaaaga	ttagccatct	taaaggagtt	tgagagacgaa	tcaccagatc	ttggccttgg	1320
tgaccgcacc	tccgatcacg	acttcatgtc	catttgcaag	gtgCGtatta	tctccttttc	1380
caagcttatt	ttacgtttct	tgaatcgttt	aaacacccga	tttaggtggc	ctagttaatt	1440
aagttgagtg	tctgatactt	caaaaacctt	tcgtaatcaa	agaatttaag	tactttaaat	1500
aaatgcagta	ggacctatat	taatgccgCG	ttaaggggttc	acatgatcta	gctagtggga	1560
tagattatac	aataatttgt	ccatccaccc	cccttttttt	taacattaac	agccttcaca	1620
tgatccattt	acgtccagat	gatttattta	ctttgtatgt	gtatgtctgc	atggaaacaa	1680
ttagattaag	tattcttcta	atgtttttga	aaaactgatt	tagcatacaa	aaaaaacttg	1740
taatcatacg	tatatgtgaa	ttcaataaat	gaatgtaata	aaattcaact	tacaaaaagt	1800
aggtagatag	aatatgatag	aagtagtagt	tttcataaat	gctttaaacc	tatcaccaac	1860
taccatcccc	acgtaagttg	gatgggtccat	ttaatgaccg	atcccatatg	tcaacatttt	1920
ttttcttttt	ttttctcata	taccctgaaa	ctttcagtat	tttaccatat	ttcttactat	1980
aatagtgaag	cttttgaatc	ttttaaatga	ttggtaacga	gttgcttatt	ttaatgttaa	2040
caggaagggt	acatgggtgca	tgagaccaa	tcagccacaa	cagtcccat	agagcgtctc	2100
aagaaccgca	taatcttcca	cgacggccg	ctagtccaac	gtccgacgcc	tttaaacgcc	2160
ttaatcattt	acctttggct	tcctttcggc	ttcatgcttt	ccatcttccg	cgtctacttc	2220
aaccttcctc	tacctgaacg	cttcgttcgt	tacacttacg	agatcctcgg	tattcacctc	2280
actatccgCG	gccaccgtcc	ccctcctcct	tcccccgga	ctcccgaaa	cctctacgtc	2340
cttaaccacc	gtactgctct	tgaccccatc	atcatcgcca	ttgctctcgg	ccgtaagatc	2400
tcatgcgtca	cttatagtgt	ctctcgctc	tcccgcatgc	tctctcccat	ccctgccgtc	2460
gctctgaccc	gagaccgtgc	cgccgatgct	gcccgcata	gaaaactcct	agagaaagggt	2520
atttgaaca	gcagtcctta	ttactatagg	cagagtgact	ctgaatcttt	actagaaaat	2580
tggtttgaaa	aagataatat	aacaactact	tcataaatat	cagaaacaga	gtatctctca	2640
ctttttggat	acataactgg	tctcattttt	atcacatttg	gtctagtctg	aatatttaat	2700
catattcatg	cgtagaataa	gataactttc	gtctatataa	aagtagaaaa	gagagcatct	2760
ctataagtgt	tttactatat	ttgcgacagg	tgatttggtg	atctgccctg	aaggcacgac	2820
tttagagaa	ccgtatctgc	tgcgattcag	tgctctattt	gcagagctaa	gcgaccggat	2880
cgtacctgtg	gccatgaact	gtaagcaagg	gatgttcaat	ggaacgacag	tcagaggtgt	2940
aaaattctgg	gatccttact	tcttcttcat	gaaccaaga	cctagctacg	aggccacttt	3000
cttgaccgt	ttgcctgaag	aaatgacgg	aaacggcggc	ggtaaaactc	caatcgagggt	3060
ggctaattac	gttcagaagg	tgatttggtg	agtattgggg	ttcgagtgca	ctgaacttac	3120
aaggaaggat	aagtatcttt	tgcttgagg	caacgacggg	aaggttgagt	ctatcaagaa	3180
gacaaaggat	gataaatcca	gc				3202

pctau2012001598-seq1

<210> 112
 <211> 1872
 <212> DNA
 <213> Arabidopsis thaliana

```

<400> 112
caatataggT gcatttaaaa cataacgttc cctctcttTgt cttcttattt ttgctctctg      60
ttcttttagaa ctgccattaa cgagagtTgag agagagagac acacacacag cgagtTgaaaa      120
tgtctccgga gaagaagagt caaaacttcc ctccaataac ggaatgcaga gacggagagt      180
acgattcgat agccgccgat cttgacggga ctctgcttct ctcaagaagc tccttccctt      240
acttcatgct cgtcgtctgtt gaagctggaa gcctttttacg tggactaatc ctccttctct      300
cgttaccatt cgtcattatc tcttacctat tcgtatccga atctcttTgt atccagatcc      360
tcattcttcat ctattcgtc ggtctcaaaa tccgcgatat cgaacttTgc tctcgcgcag      420
ttcttccacg gttttacgcg gcggatgtga ggaaagacag ttttgaggTg tttgataagt      480
gtaagagaaa agtggtagtT acggcgaatc cgattgtgat ggtggaggcg tttgtgaagg      540
attatcttgg aggtgataaa gttttgggaa cagagattga agttaaccct aaaaccaata      600
gagccactgg atttTgtaag aagcctgTgt ttcttTgtTg tgatcttaag aggttagcca      660
ttttaaaga gtttgTtaac gaatcacctg atctcggcct cggTgatcga acctctgatc      720
atgatttcat gtctctctgc aagaaaggTt acatgTttca tgcgaccaag tcagccacaa      780
cgattccaaa agaacgctta aagaaccgca tagtcttcca tgatgggcgt ttagcgcaac      840
gtccaactcc gttaaacgcc attatcacat acctatggct tccttttTgt ttcatcctct      900
ccatcattcg cgtctacttc aacctccctt tacctgaaag atttTgtccgt tacacttacg      960
agatgctcgg gatccactta accattcgtg gtcacgtcc tccacctcct tcccctggaa     1020
ctcttggcaa cctctatgtc cttaccacc gtaccgcgt tgatcccatc atcgttTgcta     1080
ttgctcttgg acgtaagatc tgttgcgtca cttacagtgt ctctcgtctc tcccttatgc     1140
tttctcctat tcctgctgtt gccctcacc gtgaccgtgc caccgatgct gccaatga      1200
gaaaacttct cgagaaaggc gacttggTga tatgtccgga aggcacgacg tgtagagaag      1260
agtatctact gagatttagc gctctattcg cagagctaag cgaccggatt gtgccagtag      1320
cgatgaactg taaacaagga atgttcaacg ggaccacagt taggggtTgt aagttttggg      1380
acccttactt cttcttcatg aaccaagac caagctatga agccactttc ttggatcgtt      1440
tgctgaaga aatgactgtc aacggtgTgt gcaagactcc tatagaggTg gctaattacg      1500
tccagaaagt tatcggcgcg gttttgggct tcgaatgcac cgaacttact cgcaaggata      1560
aatatctttt gcttggaggT aatgacggca aggtggagTc tatcaacaac accaagaagt      1620
gaatatctta ttcttTttat aacttggcat tgacttattg caagtaaggT tggatcatat      1680
acgaattttc atgtgtgtaa ttgcttgagc tcttTgtcatc attatggTac gttattTtta      1740
tgtttattct attttTttct taatgtcgaa cgggtaacta tatccttaaa aactattttg      1800
    
```

pctau2012001598-seq1

tgatcccggtt tctatacggg tgtgtataga gacttttctat ttaaagtgcg aaatttttatt	1860
ggaggaagta aa	1872

<210> 113
 <211> 1545
 <212> DNA
 <213> Physcomitrella patens

<400> 113 atgcagcaac tcgacccatt gactagtttt cactgcagtg aaagcttcga aggagtggag	60
acctgcagag tggaagggcg tgcgaaccag agcattgttt cagatttgga cggaacgctt	120
ctgcggtcaa gaagctcttt tccctacttt atgttgctcg ctttcgaagc agggagccca	180
ctccggtctg tcgttcttct gcttatttct cccgtagtat ggcttgctta caacttcgtc	240
tccgaggcag ctgggattaa gatgctcatt ttcgtctcgc tggcgggact caaagtgtcc	300
gccgtcgagt cagtggcgcg ggggtgttctg cccaaattct acttggaaga catgcactcc	360
atcagctacc gagtcttcac ctccctgcggg aagcgctacg tcgtgacggc gaatccaagg	420
atcatagtgg agccgtttct gaaagagtat cttgacgtag aagctgtaat gggcactgaa	480
ctacaaatct cgagcggggg atatgtcacc ggttttgtaa cgggccccgg tgtactcgtt	540
ggtgctgtca aagagcgagc tgtgaagaaa tactttggac ccgaccctcc ggacgtcggc	600
ctaggagaca ggagagcga cttcttattc atggatctct gcaaggaatc ttacattgta	660
cacagcgaca aggatgtgcc cgccgtgtcg aaggaagact tcttaaaacc tcttatcttc	720
cacgacggta gactggtgtg ccgtcctacg ccgctgatgt cgcttatcgt catgctatgg	780
tgtcctattg gctttgtcct tgccgttggt cggtatgtca tcggcattgc gttgccaatg	840
tggtggggcg ttcctcttga ggcaatgtc ggcgtcaccg tccgtgtgaa aggaattccc	900
ccgacgtgcc ccctccgaaa caagaagcgg ggtgtgtctt tcgtgtgtc tcatcgcact	960
cttcttgacc ctattttctt gtccattgcc gttcgaagga aagtcactgc agtaacatac	1020
agcatcagtc ggctctccga agtgctgtcc ccaatccgca ctgtgcgct cagcagggac	1080
aggaacacgg acgccaacac catggcaagt ctattggagg aaggtgacct tgtggtgtgc	1140
cccgagggca ccacttgccg ggagccttat ctgttgcggt tcagttcctt gttcgctgag	1200
ctgacggatc aaattgtccc cgtgacgatg aatatcaaaa ttaccatgtt tcatgggacg	1260
agcgctcgag gctggaaagg catggacccc ttcttcttct tcatgaacct ttgccctaag	1320
tatgaggtaa catttctgga ccaactcccg catcagctga catgcaatgg cggaagacc	1380
agccacgagg tggcgaacta cattcagcgc gtgttagcga actctcttgg atttcagtgc	1440
accaatctca cgcgcaagga caagtatact gtattggccg ggaacgatgg gattgtgccg	1500
gatcacactc tgcggcagcg agctggcaga atgttcggct tgtag	1545

<210> 114
 <211> 1506
 <212> DNA
 <213> Physcomitrella patens

pctau2012001598-seq1

```

<400> 114
atgaagagag atcccttcga taccatcgac aagtgcagtg acaaaggcag aagcaagcag      60
acggttgctc cggatctcga cggaactctt cttcgagcc gaagctcggt cccctacttc      120
atgctggtgg cattcgaagc tggaggagtg gcccgggcct ttgtgctcct cttgtgctct      180
ccgctgtgtt ggcttctgta ccattgcgtc tcagaatcca tcggcattcg cctcctgac      240
tttgtcacct tcgcgggcct aaaaatcggc ggtatcgaag ccacgcacg cggaacgttg      300
ccaaagtttt acgcggaaga cgtgcatccg gacacatggc gaatcttttc ttcgtttggg      360
gagcggtaga tacttacggc gacgccgagg atcatggcgg agaccttcgc caagacgtat      420
ctgggagtg acggtgtgtt ggggacggag ctgcatttca cttctggcgg catcgcgact      480
ggtttgttga tgaaccagg ggttctaacc gggagaaaca aagaaattgt gctgaggcaa      540
gaattccaag gcttgaacct tcctgacgtg ggtctcggag accgtccgtc cgaccacaac      600
ttcatgtcca ttgcaaaga aggatacatc gttccgccga gcaacacaat actggcagcg      660
tcgaaagaaa gtctaataa cctgctggtg ttccacgacg gccgcctgat ccaaaggccc      720
acagcaggca tagcgctcat catactgctg tggtagccga ttggcgctgt gttggccgtg      780
cttcggatcc tggctggcat ctttcttccc ttccacttgt tgaaattggt atacaagttt      840
cttgggggtg ggggtggtgt tcgtgggacc ccacccccag agcccaaaga cggccccgga      900
cgaggaggct acctgtacgt gtgctcgcat aggacgtac tagaccccgat gatggtgggc      960
gtggcggtga agcggcgctg gacggccgtc acgtacagca taagccggct gtcggagggtg     1020
ttgtcccaa tcaaacagt ggactgaag aggaatcggg agaaagacgc ggcgaagatc     1080
aggtctctgc tgcgagaagg agaccttgcg atctgccccg agggcaccac ctgtcgggag     1140
ccgtatctgc tgcgattcag cgcgctgttc gcggagctgt cgaaccagtt ggtacctgtg     1200
gcgatgaaca cgcgcatgag catgttcac ggcaccaccg cgcaggggtg gaaatgcttg     1260
gaccttttct acttcttcat gaatccaac cctatctatg aggttacgtt cctcaacgaa     1320
ttaccggttg agcttacatg tgccgggtgg aaatcttcgt atgaagttgc gaaccacatc     1380
caacaacttt tgtcgaaac gttgggcttt gaatgcacaa attatacacg aaaagacaag     1440
tatggggtgc tctgtgaaa cgacggttca attcctttaa agtctcaaga ctcctttggt     1500
tcgtag                                           1506

```

```

<210> 115
<211> 2100
<212> DNA
<213> Physcomitrella patens

```

```

<400> 115
atggagacgc aagaaccga ccctgtcaca agctttcact gcagcgaaaa ctatgacgag      60
gtggagaact gcagggtgga gggtcgggca aatcagagca ttgtatcaga tttggacgga     120
acgcttcttc ggtcgcgaag ctcttcccc tatttcatgc tgctagcctt cgaagccggt     180
agtccattac ggttcatcat tcttcttctt gtctctccca tcgtttggtt cgtctacaac     240

```

pctau2012001598-seq1

ttcattctctg	aggtagtcgg	tattaagatg	ctaattattca	tatcgcttgc	ggggctcaaa	300
gtgtcagcca	ttgagtcagt	ggcgcgtggt	gtactgccga	aatttttctt	ggaggacatg	360
cactccatca	gctacagagt	gtttatttcc	tgcgagaaac	gctacgtcgt	cacagccaat	420
ccaaggataa	tagtggagcc	attccttaag	gagtatttag	atgtagacgc	tgtgatgggc	480
accgagcttc	aaatttcgag	tgggggatat	gctaccggct	tcgtcaccgg	ccctgggtgc	540
ttagtgggga	cagcgaagca	acgagctgtg	aagaaatatt	ttgggtccga	ccagccggac	600
cttggcctag	gggaccgaac	gagcgacttc	gcattcatgg	acctctgcaa	agaggcttac	660
atagtgccta	gctataagga	ggtgccgtcc	gtgacgaaga	aagactactt	gaagccgttg	720
attttccatg	acggaaggct	ggtgtgtcgt	cctacgccgc	tgatggcgct	ggcggtgaca	780
ctttggtgcc	cgattggctt	cgtcttggtc	atcatccgaa	tgttcattgg	cataatgtta	840
ccaatgtggt	gggctctccc	tctagaagcc	atgcttggcg	tcacgggtccg	tataaagggg	900
attcccccca	catgtcccct	tcggagcaag	aaacgcggcg	tgctgttcgt	ttgttcacat	960
cgactctcc	tggaccccat	tttcttgtct	atagccgtgc	gaaggaaagt	cacggcagtg	1020
acgtacagca	tcagtcggtt	gtcgggaagtc	ctggctccga	tccgaactgt	acgtctcacg	1080
agagacagga	acacagatgc	gagcactatg	gctagtctct	tggacgaggg	tgacctcgta	1140
gtgtgtcctg	agggcaccac	ttgccgcgag	ccttatcttc	tgcggttcag	ttccttgttt	1200
gcagaattga	ctgaccaa	cgtccccgta	accatgaaca	tcaagatttc	aatgtttcat	1260
gggacgagtg	ctcgaggctg	gaagggcatg	gatccgttct	ttttctttat	gaacccccgc	1320
cctaagtacg	aggtaacatt	cttagatcaa	ctccgcac	agttgacttg	caatggcgga	1380
aaaactagcc	acgaggttgc	gaattacatc	cagcgtgtgt	tggcgaactc	tctcggattt	1440
caatgcacca	atcttacgcg	caaagataag	tacaggggtg	tggcggggaa	cgatggaatt	1500
gtaccagatc	gaagtctg	gcagagggcc	agcagaatat	ttggtttgca	gtgacattag	1560
ggtttacttg	taacgccaca	tgtcaa	cagcatcgga	tggtcctgta	cccaaataat	1620
ttcctaggtg	gcggaggacc	gatttttttc	ctcaaattgt	gcggagttag	ggaaaacctt	1680
cgggtccccag	ctatgaccca	catgtgctgc	gttgtctgca	tccagcccga	cctcaactgt	1740
gtcgagcttc	tgctgttgat	tgattgacac	cgactagaaa	ccaaatagta	caattttaag	1800
aggataaaca	tcgtgttata	atttggaag	ttgaaccata	caattccttt	ctatcagtta	1860
cagcgtttga	ctaactacag	aaacaggaag	acactgttgg	acattgctct	atctgcggcg	1920
gtaaaaagga	cactggctgc	ttttcagaga	cgacatataa	cttcggctta	tgactgagtt	1980
cttggtgctc	ctcaa	gaactgtagg	actgtagcta	gtgtctctgc	tgtctcgctt	2040
gtaagcagaa	gcattaggcg	tttgtattgt	gtattattcc	taccattcat	tgaaggacat	2100

<210> 116

<211> 1676

<212> DNA

<213> Physcomitrella patens

pctau2012001598-seq1

```

<400> 116
atggagggtga cgaagttgaa agaggccggg aactacagct tcgcgagggt gaaatcgtgc      60
agctcaggga atgatcgag gagccagacc gtggtggcgg acttggaagg caccctcatt      120
cgggggcggt ccgctttccc gtatttcttc ctggtggcgt tcgaggccgg gagttacttc      180
agggcgctgc tgctgtcgct catggcgccg ctcatgggtt ttctctacta ctttgtggac      240
gaggcgctcc ggattcacct catgatcttt ctctcattcg cggggctgaa gattaaggac      300
attgaggggtg tggcgcgcg cgtgtgcttacc aagttctacg ccgatgacct tcaccctgag      360
acttggcggg tgttctcgag ctttggcaag cgtgtggtgc tgacggcaaa cccgcgcgtg      420
atggtggagc cgtttctgaa ggaggtcctc ggcgcggacg aggttatggg gacggagatt      480
gaggtgaaca aaaagggccg cgctacaggc aggccttctga agccgggtgt tctggttggg      540
acacacaagg aggaagctct aaagcgggtgt aacattggag gtgagcgccc gcacgtcggg      600
ctaggcgacc gcgttacgga tttccctttc atggcgctact gcaaggaagg ctatgtagta      660
cccaaaacga aagtgccagc cgtaagaag caagacatgc cgaagcagct aatctttcat      720
gacggtcggc ttgtgcaact cccaacgccc cgtaatgctt tcaaagtgtt ggcattggatg      780
cctatcggcc tgcttttgtc catcgtccgc gtcacgatgg gtgtatgggt ccccatacgc      840
ttgatgccac tgtgttacia actcacaggc atcaacctcg ttgttaaggg aaacatccct      900
gagaagccca agaattggaga accggggaga ctcttcgtct gcaaccaccg gactcttctg      960
gaccccgctc tcatcgctct ggcgcttggg cgccccgtcc ctgccgtgac ttactccata     1020
tccaaagtct cagagttcct gtcccccatt ccaacaatcg gcctgtgccg agataggagg     1080
aaagactcgg ccaatatgag aagggtgctg aaagaggggtg agctcactct ttgtcctgag     1140
ggcactactt gcaggagacc tttcttgctt cgattctcag ctctcttcgc cgagttgagt     1200
gacaggattg ttcctgtggc cgtaaaaatt aacatgaata tgtttcacgg cacaactgct     1260
cgcggttaaca aagccatgga ccccttcttt gcttacatga accctcggcc gaccatcgaa     1320
ctcaaattcc tagacgagat tccaaagaac atgacatgtg gtagcggaaa atcgtccatt     1380
gaagtggcga actacataca acatctcctt gccggagaac tcggatatga ttgcacagat     1440
ttcactcgca aggataagta ccggttgctc gcaggaaacg atggcattgt acctgttaag     1500
gagaaaacga ggtaatgaat cgacctgcaa gcacattacc acagaacaag acattgaaag     1560
tgcttacgat tcagtatcac tagatggagg ggttaacccc gactcagtac tataagtgtg     1620
gaattcagtg ccaaaggcaa taaactgtga ttgattgtta cataaactcg atgcaa         1676

```

```

<210> 117
<211> 1590
<212> DNA
<213> Physcomitrella patens

```

```

<400> 117
atggaaaattc caggagtgtt tgatgctaatt ctcttctgcg ctgagtcatt tcctgaagtc      60
gaaacttgta aagtggaagg tcgtgaaaaa cagacgatta tttcagactt agatgggacg      120

```

pctau2012001598-seq1

ctcctcagat	caagaagctc	gttcccttac	ttcatgctca	ttgcgttcga	cgctggaagc	180
cctctccgct	tcatcgtcct	cttgctggcc	tctccaataa	tttggttggt	ctacaacttc	240
ttctccgagg	ctgccggtac	caaactgctc	atthttcattg	catgtgctgg	cttgaaagcg	300
tccgacatcg	agtctgtagc	gcgcggcggtg	ttgcccaa	tttacttgga	agacatgcac	360
tccgtcagct	acagcatctt	cgtctcctgc	ggtaagcggg	atgtggtgac	tgccaatccg	420
cgcacatcatg	tagagtcatt	tctgaaagag	tacatggg	tcgaggcagt	cataggtact	480
gagcttcata	tcacgaagag	cggttacgcc	acgggtttgg	taatgggtgc	cgggtgtgatt	540
gtgggagcta	acaaaacgcg	cgctgtgaag	aaatactttg	gcgatgacct	tccagacatc	600
gctctcggag	accgagccag	cgactttcct	tttatggctc	tttgcaagga	ggcatactta	660
gtaccgagct	tcaagcctgt	cgagccagtt	gccaaagctg	actacctgaa	gccgctcatt	720
ttccatgacg	gcagattagt	atgtcgacca	acacctctca	tgtcactcgt	aattattctc	780
tggtcgccga	tcggtcttat	tctagcactc	attcggatgg	tgtttggcac	cgtgttacct	840
atgtggtatg	cgctgccgct	ggaggccatg	cttggcggtta	gcacccgcgt	gaaaggaatc	900
ccccccagct	gcccccaa	cagcggcaga	cgcggtgtgc	tcttcgtctg	ctctcaccgc	960
actttgctcg	accccatctt	cctgtcaata	gcctgtcggc	gacaagttac	agcagtaacc	1020
tacagcatca	gcaaggtctc	ggagatactg	tcgccgatcc	ccacagtccg	actcacaagg	1080
tgtcggaaaa	ccgatgcgga	gacatgaca	aagcttctca	acgagggaga	tcttgcatgtg	1140
tgcccagaag	gcacgacgtg	ccgcgagcca	taccttcttc	ggttcagctc	cttgttcgcg	1200
gagttagccg	accaa	tcccgtcacg	atgaacgtca	aaacgacct	gttcacggc	1260
acgacagccc	ggggctggaa	ggcactggac	ccctttatct	ttctgatgaa	tcccagccct	1320
aggtatgaat	tggagtttct	agatcagctt	ccgcacgaga	tgacctgcag	cggcggaag	1380
agtagtcacg	aggttgcgaa	ctacattcaa	cgagtcctcg	cggcatctct	gggtttcaag	1440
tgcacgaacc	tgacgcggcg	agacaagtat	cgaatcttag	caggaacga	tggcgtggtt	1500
cctgcaagcc	atthtttgaa	acaaagaagc	aggatcttcg	catgcttgcc	ttccaggaat	1560
agtgaagagag	gccagcgatc	agacccgtaa				1590

<210> 118
 <211> 1512
 <212> DNA
 <213> Vitis vinifera

<400> 118	
atgggtgtaa	ctacctttcc aactgtggac cagtgtcaat ccattggccg agaaaagcac 60
actgtggttg	cggacatgga tggaaccttg cttattggca ggagctcctt tccttacttt 120
gctttggttg	cttttgaggg gggaggtgtt ctaaggctgc tcttcctgct cttggcttca 180
ccactcgctg	gacttcttta ctatthttgtg tcggagtctg caggaatcca gattctcatc 240
tttgtaacct	tcgctgggat gaaggtgtca gacattgagt cagttgcacg tgctgtgctg 300

pctau2012001598-seq1

cccaaatttt	atttgagcga	tctccatccc	gagtcattggc	gggtgttctc	atcatgtgga	360
aagcgttgtg	ttcttaccgc	aaaccccagg	ataatgggtg	aggcatttct	gaaagatttt	420
ttgggagctg	atttagtttt	ggggactgag	atagccactt	ataagggtag	ggcaactggg	480
tttgtttgtg	atccaggtgt	acttgtgggc	aagaacaagg	catatgtctc	caacaagact	540
tttggagaga	cccagcctga	gattgggctg	ggtgatcgtc	acactgattt	tcccttcattg	600
gcttcattga	aggagggcta	tattgtggca	gccaaaccag	aggttaaggc	ggtaacaagt	660
gacaagcttc	ccaagcccat	catcttccac	gacagccgcc	tcgtccaaaa	gccaaactcct	720
ctcacggcac	tcctcaccct	tctctggatc	cccataggct	tcgcactcgc	atgcttgcca	780
atcgccgctg	gctcactcct	ccccatgcgc	attgtctacc	acgctttctg	ggcactgggt	840
gtccgagtct	ccgtcaaggg	ttccccacct	cctgttgcca	aaaagtcaat	aggccagtca	900
ggcgtgctat	tcattctgctc	ccaccgaacc	ctccttgacc	caatattcct	ctccgcagcg	960
ctcggccgctc	ccattcctgc	agtcacgtac	tccgtctccc	gcctctccga	gttcattctcc	1020
cctatcaaaa	ccattcgcct	aagccgagat	cgagtcaaag	acgcagccat	gataaagaaa	1080
ctgcttgaag	aaggtgacct	agctatttgc	cccgaaggaa	ccacttgccg	tgagcctttc	1140
cttctcaggt	tctcagcatt	gtttgctgaa	ctgactgacc	agcttggtcc	agtggctatg	1200
atgaaccgca	tgagcatgtt	ccatggaaca	acagcaagag	ggtggaaagg	gatggaccca	1260
tttttcttct	tcattgaatcc	aagcccagcg	tatgaggtaa	cgttcttgaa	caagctgccc	1320
ctggaactta	cttgacgttc	ggggaaatct	agccatgagg	tggccaacta	catgcagagg	1380
gtgattgctg	ccactctttc	ctatgaatgc	actagcttta	ccaggaaaga	taagtaccga	1440
gcactcgtcg	gaaatgatgg	gaccgtgccc	aaaaaacctc	tgctatatcc	cgacaaagtc	1500
atgggatgct	ga					1512

<210> 119
 <211> 1784
 <212> DNA
 <213> Vitis vinifera

<400> 119						
gtttcttttc	gttggtccact	gtgctgcctg	aaaggctgaa	tccggttcta	ccggccgaga	60
catgtctccg	ccaccgaaac	gcgcccgaac	atttcctcc	atcactacct	acgccggcgg	120
cgaccatcgc	tctattgcgg	ccgacctcga	tggcactctc	ctcgtctctc	gctcctcctt	180
cccttacttc	atgcttgtcg	ccgtcgaagc	cggcagcctc	ctacgcgggc	tcttccttct	240
gctctctctt	cctatagtga	tcgttgccca	cctcttcctc	tcagaagaaa	tcggaattca	300
gattctcatc	ttcatttcct	tttccggcct	caagatccga	gacatcgagc	tcgcgtcgca	360
ggcgggtgctg	ccgcgggttc	acgccaacga	cgtgaggcga	gagagctggg	agggtgttcga	420
gaagtgcgaa	cggaagggtg	tgggtgacgg	caacccgacg	ttgatgggtg	agccgttcgt	480
gagagatttt	cttggcgga	cgaagggtgt	agggacggag	attgaggtga	atccgaagac	540
gaagaaggcc	actgggtttc	tgaagaagcc	aggggtgttg	gttgagagaca	gaaagagatt	600

pctau2012001598-seq1

```

ggcgctgttg aaggagtttg gtgatgaatt gcctgatatc gggattgggtg accgcgagtc      660
tgatcatgat ttcatgtcca ttgcaagga gggatacatg gtgcttccga gtaaatacagc      720
tacgccggtg ccaccaaac gcctgaaaac cccaataata ttccacgatg gtcggttcgt      780
ccaaccccc acgccactca cggccctcat catctacctc tggcttccat ttgggtttgc      840
actctctatc ttccgagtct acttcaacct cccctccca gagcgcacgc tccgctacac      900
ctacccgatg ctgggcatca accttggtat ccgaggcaac cctcctccgc cccctcccc      960
aggttcccc ggcaacctct acgtctgcaa ccaccgacc gccttgacc ccatcgatcat     1020
cgccatcgcc ctccgccga aggtctcctg cgtcacctac agcgtcagcc gcctctctcg     1080
cttctctcc cccatccccg ccgtcgccct caccgcgac cgccgcccg acgtgccccg     1140
catctctcc attctccaga agggcgacct cgtggtctgc ccagagggga ccacctgccg     1200
cgagccctac ctctccggt ttagcgact gttcgggaa ctcagcgacc gcatcgatgcc     1260
ggtggcggtg aacgtgaagc agaacatgtt ccacggcacc accgtccgag ggggtgaagtt     1320
ctgggacgcc tatttttact tcatgaacct tcggcccacc tatgagataa cgtttctgga     1380
tcggttgccg gaggagatga cgtgtaaggc cggagggaa tcggccatag aagtggcgaa     1440
tcacgtgcaa aagggtgttg gtggagtgtg ggggtttgag tgactggac tgactagaaa     1500
ggataagtac atgttgctgg gaggaaatga tggcaagggt gaatccatgt acaatgcaa     1560
gaagtgaag gctgggatcg gagaagaagg ccacattgga attgtatggg ttttagaaaa     1620
attacgaggg aggtatatcg gttaatgtag aattacaact tctatgaaa ttcccttagt     1680
atgattgata attgttgatc atttttcatc tggattattc atattatttg gacattcaag     1740
tttagagtta tgaagagtgt cttctgattt ctgtggttga tttt                        1784

```

<210> 120
 <211> 1533
 <212> DNA
 <213> Vitis vinifera

```

<400> 120
atgaccaa at caaacctcca ctctcctccc acctccctaa ccatcgatca atgcacctcc      60
atcgccggg aaaaggagac ggttggtggc gacatggacg gaactcttct cagagggcgc      120
agttcgttcc cctactttgc acttggtgcc ttcgagggtg gtgggatttt taggcttctt      180
ctgttggtgc tatgttctcc attggctgga gttctctact acttcatatc ggagccagct      240
ggcattcagg tgctcatatt cgcgaccttt gtggggatga aggtttcgga tatcgagtcg      300
gtggcgcggg cagttttgcc caagttttac gcgagcgacc tgcaccgga ggcgtggagg      360
gtgttctcgt cgtgtgggaa gaggtgtgtg ctcaccgcga atccgaggat tatggtggag      420
gccttcctga aggatttttt ggggtgcagac atggtgttgg ggacggaggt ggacacgtat      480
aaagggagag ctactgggct tgtccgtggt ggtggggttc ttgttgggaa gagaaaggca      540
gaggccctcc gaaatgcctt cagcgagacg tctttaccgg agattggact cggatgatcg      600

```

pctau2012001598-seq1

gagacggatt atccatttat	gagtctctgc aaggagagtt	acatagttcc agcgaagcct	660
gaggtggagg ctgttagcca	tgacaaattg ccaaagccca	ttgtattcca cgacggccgc	720
ctcgtccaaa aaccaacccc	tttcatggcg cttctcacca	tcctctggat cccagtaggc	780
ttcctcttgg cctgcttgcg	catcgccgct ggtgcgctcc	ttccaatgcc gcttgtctac	840
tacgcctttt gggcactcgg	tgtcagagtt tatatcaagg	gcaacccccc tccaccagcc	900
aaaaaatcaa taggccaatc	tggcgtcctc ttcatttgct	ctcacagaac cctcctcgac	960
cctatcttcc tctccgccgc	attaggccgc cctatccctg	cagtcacgta ctccctttct	1020
cgtctctccg aattcatctc	ccccatcaag acagtcgggc	tcagtcgcca tcgagccacc	1080
gacgcagcaa tgataaagaa	gctcttagag gcaggtgacc	tagccatatg tcccgaagga	1140
accacttgca gggagccatt	tctgctaaga ttttctgcct	tgttttctga gctaaccgac	1200
gagctgggtc cggtaggtat	ggcaaaccgc atgaccatgt	ttcatgggac aacagctaga	1260
ggttgaagg ggatggacc	cttttactac ttcatgaacc	ctagccctgc ctatgaggt	1320
aacttcttga ataagcttcc	acatgagctc acttgcggtg	ctggaaaatc gagccacgac	1380
gtcgcgaact acatacagag	gatgattgcc gcaagtctat	catatgagtg caccaacttc	1440
actaggaagg ataagtatcg	ggcactggcc gggaacgatg	gaactgtagt ggaaaagccc	1500
aaggttgatg ctaataaagt	catgggctgc tag		1533

<210> 121
 <211> 1726
 <212> DNA
 <213> Vitis vinifera

<400> 121			
atgggggctc aacgccattt	ccaaccaata tcgaaatgca	gctccgacgg tcgagccaat	60
cagacagtgg cggcagactt	agatggcacg cttcttgtgt	cccgaagcgc atttccttac	120
ttcatgcttg tggcacttga	agctggtagc cttgtgaggg	caatcctcct cctagcatcg	180
gttcactag tgtatttcat	ttacttgttc atttcggagt	cgctagccat taaaatcttt	240
atctatatgg cttttgcggg	actcaaagtt agagagattg	agcttgtttc gaggtcagtc	300
ctgccaagt tctatgctga	ggatgtgcac cccgagagtt	ggagagtgtt taattccttt	360
gggaagaggt atatcattac	tgccaatcct agaattatgg	tggaaccctt tgtgaagacc	420
tatatgggtg ccgataaggt	tcttgggacc gagttggaag	tatcgaactc cggaagggcg	480
actgggtttg tgaagaaacc	aggagttttg gttggggaac	ataagaaagc ggcagtgcag	540
caagaatttg gcatgaattt	gccagatgtg ggcctgggag	atagagatac ggatcatgat	600
ttcatgtcct tatgcaagga	aggatacatg gtgccaagga	caaatgtga gccactacca	660
aggaacaagc ttctgagtcc	agtcattttt cagcagggac	gtttggttca gagaccacc	720
cctgtggctg ccctgttgac	tttcttatgg atgccaatg	ggatcattct ctccatcctt	780
agggcttacc tcaacattcc	cttgccagag cgtattgtga	ggtacaatta caagatcctg	840
ggaatcaagc ttattgttaa	gggtacacct cctccccgc	caaagggagg ccaaagtga	900

pctau2012001598-seq1

gtcctcttcg	tttgaacca	taggacagtt	ttggaccccg	tggtcacggc	tgtggcttta	960
ggcagaaaaa	tcagctgtgt	cacttacagt	attagcaaat	tctcggagct	aatttctcca	1020
atcaaagccg	ttgcactatc	cagggagaga	gaaaaggatg	cagcaaatat	caagcgctta	1080
cttgaagaag	gtgatttagt	aatttgtccg	gagggaaacta	catgtagaga	gccatttctg	1140
ttgagattta	gcgccctttt	tgctgagctc	actgatagaa	ttgtgcctgt	tgccattaat	1200
accaagcaga	gcgtgtttta	tggcacgtca	acccgggggtt	ataagttggt	ggatccttac	1260
ttcgtgttca	tgaatccaat	gcccacctac	gagatcacct	tcttgaatca	gctgcccgcc	1320
gagctcactt	gcaaagggtg	gaaatcccca	atcgaagttg	caaattacat	ccagaggggtg	1380
ctggctggaa	cccttggctt	tgagtgcact	aacttgacca	ggaaggacaa	atacgccaca	1440
cttgccggaa	cagacggcag	tgttccgttc	aagaacgaaa	aggaaaagga	aaaggcttga	1500
gaactgaaag	tgattcaaag	taaattaatt	agtctcatga	ttttctattg	ttttacatat	1560
attatatgcc	ggaatcatta	tcatgacaat	gaactccgat	tgtgatgata	ataaattggt	1620
gtgattataa	aaataatagt	ttcttatatt	gtcgtggaga	caagtcaacg	gtcatttttt	1680
ttaagttttt	gaaaataaat	tttatctgga	caagtttcac	cttaga		1726

<210> 122
 <211> 1504
 <212> DNA
 <213> Populus trichocarpa

<400> 122	
tcctctctca	tggtaatggg ggctcatcgc cattttgaac caatagcaaa atgtagcaca 60
gagggacgat	caaatcagac tgtggctgcc gacctggatg gaacacttct tgtatcaaga 120
agtgcctttc	catacttctt gcttggtgct attgaagctg ggagtctctt aagaggacta 180
attctcttag	catcagtccc atttgatac ttacgtact tatttatatc agaggcaatg 240
gcaatcaaga	ctttcatctt cattgccttt gcgggactta aaataagaga cattgagctt 300
gtttcaaggt	ctgtcttgcc aaagttttat gctgatgatg ttcattccaga aacttggagg 360
gtgttcaatt	cttttgaaa aagatacatt gttactgcta atcctaggat tatggtcgag 420
ccatttgtca	agaccttttt aggggctgac aaggttcttg gcactgaatt agaggctaca 480
aaatctggaa	gagcaactgg gtttatcaag aaacctggag ttcttggttg agatcataaa 540
agagatgctc	tcttaaaaga gtttggcaca aatttgcctg atttaggcct gggagaccgc 600
gaaactgacc	atgacttcat gtccatttgc aaggaaggat atatggtgcc aggaaccaag 660
tgtgagcctc	tagcaagaaa caagcttcta agccctgtca tattccacga tggccggtta 720
gttcaaaggc	ctactccttt ggctgccctt ttgaccttct tatggatgcc aattggtatt 780
attctctcca	tacttagagt ctaccttaac atccctttgc ccgagagact tgcttggtac 840
acttataagc	tactaggaat tagagttatt gtcaagggtta cccctccacc ccctccagga 900
aaaggccaca	gtggagtcct ctttgtttgc aatcatcgta ctgtcttaga cccagttgtc 960

pctau2012001598-seq1

actgctgttg cactaggaag aaaaattagt tgtgtcacct atagcataag caagttcact	1020
gaaattatatt caccatttaa agctgttgct ttatcaaggg agagagacaa agatgctgct	1080
aacattaagc gactgcttga agaaggtgac ctggctcattt gccctgaagg aaccacatgt	1140
agagagccat ttctcttgag atttagtgct ctttttgctg agcttactga taggattgtg	1200
ccagttgcta tcaatacaaa acaaagtgtg tttcatggca cgacggttcg agggcacaaa	1260
ctgttgacc cttattttgt gttcatgaat ccaatgccaa catatgaggt caccttcttg	1320
aatcagctgc ccaaagagct tacttgcaaa ggaggcaaat cagccattga agttgcaaat	1380
tacatacaaa ggggtgctggc tgggacactt ggatttgagt gcactaactt gaccaggaag	1440
gacaagtatg ctaagcttgc aggaacagac ggccgtgttc tatctaagaa ggaaaatgct	1500
tgag	1504

<210> 123
 <211> 1506
 <212> DNA
 <213> Populus trichocarpa

<400> 123	
atggtttagct tcccaactgt tgacaaatgt gcatccatcg gcagagaaaa gcacagtgtgta	60
gttgctgaca tggatgggac cttgcttaga ggccgtagct cattccctta ctttgctcta	120
cttgccctcg aagctggtgg gatTTTTtagg ctgcttttct tgctcttgaa ctcaccacta	180
gcaggacttc tgtattactt tgtgtccgag tctgctggta ttaaagtctt tatctttgct	240
acatgtgccg gaatgaagtt atcagatatc gagtctgtag cacgtgctgt gctgccaaag	300
ttttactcta gtgaccttca ttctgagtcg tggcgtgtgt tctcttcttg cggaaaacgt	360
tgcgttctta ctgcaaatcc aagaattatg gtggaagcat ttttgaaaga tttcttgga	420
gctgatttgg ttttggggac tgagatatta acttataaag gtagagcaac agggtttgtt	480
cagagcccgg gactacttgt ggggaagaac aaggcagatg cccttaaaaa ggcttttggg	540
aagacacaac cagagattgg acttgagat cgctacacag atgcaccctt catggctcta	600
tgcaaggagg gctacatagt gccacctaag ccagaggttg aagctgtaac tactgacaag	660
ctcccaaaac cagtcacttt ccatgatggc cggctagtcc aaaaaccaac accattatca	720
gcactactca taattctctg gatccctata ggcttcattt tagcctgctt gaggattgca	780
gcagggtcac tcctgcctat gccaatggtc tactatgctt tctgggcact tgggtgtccgt	840
gtcaacataa aggggaacccc acctcccca gccaaaaaat caattggtca atctggtgtc	900
ctctttgttt gtcacacag aaccttgctt gatcccatat ttctctccac tgcccttggg	960
cgccctattc cagcagttac atactcactt tctcgtctct ctgaaattat ctcaccaatc	1020
aaaactgtta gactcagccg tgaccgagct gcagacgcat caatgatcaa gaagttatta	1080
gaagaagggtg acctagctat atgtcccag cgaactactt gccggaacc gtttctttta	1140
aggttctcgg ctttgtttgc tgaattaaca gacgaactcg ttccagtggc tatggcgaat	1200
cgcatgagta tgtttcacgg aaccactgca agaggctgga aagggatgga tccattttac	1260

pctau2012001598-seql

ttcttcatga accctagccc agcatatgaa gtaactttcc tgaacaagtt gccactagag	1320
ctaacattaa gtgggtgggaa atctagccat gaggtggcaa actacatata aagagtgatt	1380
gctgctactc tctcctatga atgcacttcc ttctactagga gagacaagta cagggcactt	1440
gctgggaatg atgggaatgt ggttgaaaag acgaaggtcc aagccaacaa agtaatgggg	1500
tgtaa	1506

<210> 124
 <211> 1506
 <212> DNA
 <213> Populus trichocarpa

<400> 124	
atggatagct ttccaactgt tgacaaatgt gcatccatag gtagagaaaa gcacactgtg	60
gtagctgaca tggatggaac cctgcttaga ggtcgtagct cattccctta ctttgctcta	120
cttgcctttg aagctggtag gatttttagg ctgattttct tgctattggc ctcaccgcta	180
gcaggacttc tatactactt tgtgtccgag tctgctggta ttaaggttct tatatttgct	240
acatgtgccg gaatgaaggt ttccgatatt gagtctgtgg ctcgtgctgt actgcctaag	300
ttttactcta gcgatcttca ttctgagtcg tggcgtgtgt tttcttcttg tggaaaacga	360
tgtgttctta ctgcaaatcc aagaattatg gtggaagcat ttttgaaaga tttcttggga	420
gctgatttgg ttttggggac agagatgtct acttataaag gtagagcaac agggtttgtt	480
cagagcccgg gactacttgt ggggaagaac aaggcagatg cccttaaaaa ggcttttggg	540
atgacacaac cagagattgg acttgagat cgtcatacag atgcaccctt catggctcta	600
tgcaaggagg gctacatagt gccacctaag ccagaggttg aagctgtaac tactgacaag	660
ctcccaaac cagtcatctt ccatgatggc cggctagtcc aaaaaccaac accattatca	720
gcactactca caattctctg gatcccaata ggcttcattt tagcctgctt gagaattgca	780
gccgggtcac tcctgcctat gccaatggtc tactatgctt tctgggcact tgggtgtccgt	840
gtcatcataa agggaacccc acctccccg gccaaagaaat caatcgggtca atctggtgtc	900
ctctttatct gctctcaccg aaccttgctt gatccaatat ttctctccac tgccctcggc	960
cgtcctatcc cagcagtcac gtactcagtt tctcgtctct ctgaaattat ctcacccatc	1020
aaaactgtta gactcagccg tgaccgtgct acagacgcat ccatgatcaa gaagttgtta	1080
gaagaagggtg acctagctat atgccctgag ggaactactt gccgcgaacc ctttctttta	1140
aggttctcag ctttgtttgc tgaattaaca gaccaacttg ttccagtggc tgtggtgaat	1200
cgcattgagta tgtttcatgg gacaacagca agagggtgga aagggatgga tccgttttac	1260
ttcttcatga accctagccc agcatatgaa gtaactttct tgaacagggt accacaggag	1320
ctaacctgca ctgggtgggaa atctagtcac gaggtggcaa actacatata aagagtgatt	1380
gctgccaccc tgtcctatga atgcacttcc ttctactagga gagataagta cagggcactt	1440
gctgggaacg atggtactgt ggttgaaaaa acgaagctcc aagccaacaa agtaatgggg	1500

1506

```
<210> 125
<211> 1797
<212> DNA
<213> Populus trichocarpa
```

400>	125					
acaacataac	tcacacttcc	taagaactct	caaaataaac	actatctatc	cctctccctc	60
tccctctcta	tccccctctg	tgatgggtcat	ggggggtcat	cgccattttg	aaccaatagc	120
aaaatgtagc	acagaggggc	gatcgaatca	gaccgtggct	gcagacctag	atggaacact	180
tcttgatca	agaagtgtt	ttccttactt	cttgcttggt	gctattgaag	ctgggagtct	240
cctgagagga	ctaattcttt	tatcatcagt	cccattttgtg	tactttacgt	acttaattat	300
atcagaggaa	atggcaatca	agaccttcac	cttcattgcc	ttcgctggac	ttaaaataag	360
agacattgag	cttgtttcaa	ggtctgtctt	gccaaagttt	tatgctgatg	atgtccatcc	420
agaaacttgg	agggtgttca	attctttttg	aaagagatac	attgttacag	caaatcctag	480
gattatggta	gagccatttg	ccaagaactt	tttaggtgct	gacaaggttc	ttggcactga	540
aatacaggct	acaaaatctg	gaagggcaac	tgggtttgtg	aagaaacctg	gagttcttgt	600
tggagatcat	aagaggaatg	ctctcttaaa	ggagtttggt	acaagtttgc	ctgatttagg	660
cctgggagac	cgcgaaactg	accatgactt	catgtccatc	tgcaaggaag	gatatatattg	720
gccaagatcc	aagtgcgagc	cactaccaag	aaacaagctt	ctaagccctg	tcatattcca	780
tgagggccgg	ctggttcaaa	gacctactcc	cttggttgcc	cttctgacgt	tcttatggat	840
gccaattgggt	attattctct	ccgtccttag	ggtctacctt	aacatccctt	tgccctgagag	900
aattgcctgg	tacaattata	agctgctagg	aattagagta	gttggttaagg	gtaatcctcc	960
accccctcca	agaaaaggcc	aaagtggagt	cctcttttgtt	tgcaatcacc	gtactgtctt	1020
agaccctggt	gtcactgctg	tcgccctaag	aagaaaaatt	agctgtgtca	catacagcat	1080
aagcaagttc	actgagatta	ttcacctat	caaggctggt	gctttatcaa	gggagagaga	1140
cagagatgct	gcaaataata	agcgtctact	tgaagaagg	gacttggtaa	tttgcccgga	1200
aggaaccact	tgtagagagc	catttctctt	gagatttagt	gctctttttg	cagaacttac	1260
agataggatt	gtgccagtcg	ctatcaatac	aaaacaatct	gtcttcaatg	gcacgacggt	1320
tcgagggcac	aaactgttgg	acccttattt	tgtgttcattg	aatccaatgc	caacatatga	1380
gatcaccttc	ttgaatcagc	tgcccacaga	gcttacttgc	aaaggaggca	aatcatccat	1440
tgaagttgca	aattacatac	aaaggggtgct	ggctgggaca	cttggatttg	agtgcactaa	1500
tttgactagg	aaggacaagt	acgctatcct	tgcaggaaca	gacggccgtg	ttgcaactaa	1560
gaaggaaaaag	gcttcagaaa	attaatctct	ccaaaaggcg	gcataaaaaga	aaatccatga	1620
tttaattttct	ctattgtttt	tttttatata	taaaatattg	ctataacaat	aaaatttgac	1680
aatcgctatt	gtgattaatt	actataaaat	aacttttagag	tatatttgct	ttttttattt	1740
tttaagtatg	tataccgaca	tcaaattact	ctaattgcaat	gtttatttta	tcttcaa	1797

pctau2012001598-seq1

<210> 126
 <211> 1575
 <212> DNA
 <213> Sorghum bicolor

```

<400> 126
atggagcgtc ccagcccaac gtgcttagcg gcgaacggcc cacgtagaaa agcccgaaca      60
atcccgttcc cgccgcccc ggctgatctt cgtacgggga ggttgcggtc gcggcagacg      120
gtcgccgccg acctcgacgg cacgctgctc ctgtcccgaa gcgcgttccc gtactacctc      180
ctcgtggcgc tggaggccgg cagcctcctc cgcgccgtgg cgcttctcat gtccgtgccc      240
ttcgtgtacc tcacctacgt caccgtgtcc gagaccctgg ccgtgcgcgc gttcctctac      300
gtcgccgtgg cggggctcga cgtgagcgac atcgaggccg tcgcccggac cgtgctcccc      360
aggttctacg ccggcgacgt ccaccgggag ggctggcgcg tcttccgctc gttcgggagg      420
aggtgcgtcg tcaccgccag ccgcgggtg atggtggagc cgttcgccag ggcgttcctt      480
ggcgccgacg tggtcacggt gacggagatg gaggtgggcg cgtccgggaa ggccacgggg      540
ttcgtcgccg ggcccggggg gctcgtcggc gaacacaaga ggcgggcggt ggtgagggag      600
ttcgggcagc cgctgccgga cgtcggcatg ggggatcgcg agagcgactt cgacttcatg      660
gccatctgca aggaagcgta catggttaaca cggcagaagt accgcgcgtt gccccgggag      720
cagctgcaga gccgggtcat cctccacgac ggccgcctgg ccgggcgacc gacggcgacc      780
aacacgcttc tcactctcct gtggatgccg ctcggtctcg cgctcgcgct ggcgcgcgtg      840
catctccacc tgctcctccc ggcgcgcgcc ctctcctacg cctacaagct tatgggcgtc      900
aagctggtgg tgcgcggcaa ccggccgccg ccgccgccgc cgtccaagaa gggcgggcggc      960
cccggggttc tcttcgtgtg caaccaccgc accacgctcg accctgtcgc ggtggccgtc     1020
gcgctgggcc gcaaggtgag atgggtcgtc acggacggcg cctccagcag gttctcggag     1080
gccgtgtcgc ctgtcatgac aggcgtggcg ctgcccgtgc ccgtgccag ccgcgagagc     1140
gacgacgacg cggacgccgc cgcccgcatc cggcggtgc tcgaggaggg agacgacgtc     1200
gtcatcttcc ccgagggcac catctgccgc gagcccttcc tgctgcggtt cggcgcgctc     1260
ttcgtgaggg tcaccgaccg catcgtgccc gtggccatcg gcgctaggga gggcatgttc     1320
cacgggtcca cggcgcgcgg gctcaggaga atggatccct acttcttttt catgaaccca     1380
cggccgacgt acgaggtgac gttcctgaac cagcttccaa gggagctcac ctgcggcggc     1440
ggaaggtcgc cggtcgaggt ggccaactat gtccaggagg tcctggccgc gcagcttgga     1500
ttcgactgca cttgcaccag caagcaggca aaaacaggat ggtctccggc ggccccgacg     1560
gttgcgtggc gttga                                         1575
    
```

<210> 127
 <211> 1967
 <212> DNA
 <213> Sorghum bicolor

pctau2012001598-seql

<400> 127
gcttcccata tccatactcc cccgcgcccc tcaccaccaa cctcccgtcg cgtatattat 60
aactacacct cgccgcccc a tcttccactg ccaactcgca gcagcagtct ctctctgtct 120
cccacacact ccttccccgg ggcctggcag tgagatcgag cctcgatcca tcgagcacta 180
tagtagcaaa gacacaagat ggtggcgctc tatcgcagtc gcaggttcaa gcccatcgag 240
gagtgtgtct cggagggggc gtcggagcag acggtggccg ccgacctgga cggcacgctg 300
ctcatctccc ggagcgcgtt cccctactac ctctctgtgg cgctcgaggc cgggagcgtc 360
ctccgcgccc tgctgtgtct cctgtccgtg cctttcgtct acgtcaccta catcttcttc 420
tccgagtcgc tggccatcag cacgctcgtc tacatctccg tcgcggggct caaggtgcgc 480
agcatcgaga tgggtggcgc ctccgtgtct cccaggttct acgctggcga cgtgcacccg 540
gagagctgga ggggtgtcaa ctcttccggc aagcgtaca tcatcaccgc cagccccagg 600
atcatggtcg agcccttcgc cagggccttc ctccgcccgc acaaggtcgt cgggaccgag 660
ctcgaggtcg gcaagaacgg caaggccacg gggttcatgg tcaagcccgg cgtgtctcgtc 720
ggcgaccaca agaagaaggc cgtcgtcaag gagctgggag acgccgtgcc cgacgtcggc 780
atgggggata gggagaccga cttcgacttc atgtccatct gcaaggaggc gtacctggtg 840
acgtcgagga agtacagccc ggtgggcaag aaccagctgc tgagcccgt gatcctacac 900
gacggtcgcc tgggtgcagc cccgacgccg ctcgctcgcgc tcgtcacctt cctctggatg 960
ccgttcggct tcgcgctcgc gctcatgcgc gtgtacatca acctgccgt gcccgagcgc 1020
atcgtctact acacctaaa gctcatgggc atcaggctga tcgtcaaggg caaccgcgcg 1080
ccgccgcca agaaggcca cccgggcgtc ctcttcgtct gcaaccaccg caccgtgctc 1140
gaccccgctg aggtcgccgt ggcgctgcgc cgcaaggta gctgcgtgac ctacagcatc 1200
tccaagttct ccgagctcat ctgcccac aaggccgtcg cgctgtcgcg ggagcgagag 1260
aaggacgccg agaacatccg ccgcctgtg gaggaaggcg acctggtgat ctgccccgag 1320
ggcaccacct gccgcgagcc cttcctgtg cggttcagcg cgctcttcgc cgagctcacc 1380
gaccgcatcg tgccggtggc gatcaacacc aaggagagca tgttccacgg ctccaccgtg 1440
cgcggggttca agtcatgga cccttacttc ttcttcatga acccgccggc gacgtacgag 1500
atcacgttcc tgaccagct cccaaggac ctacatgca gcggcgga gtcgcccac 1560
gaggtggcca actacatcca gaagacgtc agcggacagc ttggcttcga gtgcaccacc 1620
atcacgcgca aggagaagta cggcactc gctggaacgg acggccgtgt cccgtccaag 1680
aacaaggaca aggagaagga gaagaactaa cggccaacgt cgatcatcgt gctgcccgc 1740
cagcacactg ccgttacgag tactgtgaa atttattcca attaattgtt ccttgtcgag 1800
aagttggcca attactcatt gtttttaaaa ggaccatgtg atgtgggccc tgtactagta 1860
caagcttgat gatcatataa ttattatcta caagttcttc ctcttcggaa gaagggcaaa 1920
agtgttgta ggcccgtctt aatacaaagt ttcgttgcat agtttcc 1967

pctau2012001598-seq1

<210> 128
 <211> 1518
 <212> DNA
 <213> Sorghum bicolor

<400> 128
 atggctgcct cgccgttccc cacggtggag aagtgtcctt ccaccgaccg gtccggcgac 60
 acggtggtgg cggacctgga cggcacgctg ctgtgcggcc ggagctcctt cccgtacttc 120
 gcgcacatgg cgttcgagac cggcggcgctg ctgcgcctgc tgcgtctcat cgcgctggcg 180
 ccgctcgcgg gcctcctcta ctacttcgtg tccgagccgg cgggcatcca ggtgctcatc 240
 ttcgcgtcca tggcgggggc caggggtggc gacatcgagg ccgtggcgcg ggcggtgctg 300
 cccaagttct actgcgccga cctccaccgc gagtcgtggc gcgtcttctc ggcgtgcggc 360
 cgccggtgcg tgctcaccgc gaaccgcgg atcatggtgg aggcgttcct caaggagtac 420
 gtcggcacgg acgtcgtcgt cggcacggag ctgcgtcgtg ggcgaggccg cgcgacgggg 480
 ctggtgcgct ccccgggcgt cctggtcggc gagcagaagg cggacgcgct ccggaggacg 540
 ttcggcgacg acgtcgcgcc cgaggtcggc ctcggcgaca ggaagacgga ctaccggttc 600
 atgaggctgt gcaaggaggg ctacgtcgtg ccggcgacgc ccaagctgaa gcccgtgccg 660
 cgcgagaacc tgccgaagcc ggtggtcttc cagcagggc ggctcgtcca gaagccgtcc 720
 ccggcgctcg cgctgctcac cgtgctctgg atcccgatcg ggttcctgct cgcgtgcctg 780
 cgcacgcggc cgggcgcgct cctgccgatg cgcattggtg accacgcgtt ccgcgccctc 840
 ggcgtccgtg tcaccatcag gggaaaccct ccgccgccgg caagccgcga gacgggccag 900
 accggcgtgc tcttcatctg ctcccaccgc accctcctcg acccatctt cctctccacc 960
 gccctggggc gcccacacac cgccgtcacc tactcggctt cgcggctgtc ggaaattctg 1020
 tcgccaatcc gcacggtgcg cctgaccgga gaccgcgcgg cggacgcggc catgatccgg 1080
 cggctcctga ccgagggaga cctggtgatc tgcccggagg ggacgacgtg ccgggagccg 1140
 ttcctgctgc gtttctcggc gctgttcgag gagctgacgg acgagatcgt gccggtggcg 1200
 atggagaacc agatgagcat gttccacggg acgacggcgc gagggtggaa agggctggac 1260
 cctttctact tcttcatgaa cccgagcccc gggtagctgg tcacgttcct caacaagctc 1320
 ccagcagagc tcacctgcaa tggcggcggc aagagcagcc acgaggtggc caactacatc 1380
 cagcgggtca tcgctccac gctgtcctac gagtgcacca acttcacacg gaaggacaag 1440
 tacaaggcgc tcgctggcaa cgacggcacc gtcgtgtcca agcccaacat tgacaagaac 1500
 aaggtcatgg gctgctag 1518

<210> 129
 <211> 1776
 <212> DNA
 <213> Sorghum bicolor

<400> 129
 gcagtgcgca gtgcgagaga gagacagagt gactgacagc tccgatccgc cgcacacgcg 60
 cgtgccacag cagatagcag cgagccatcc atcccgccgg ccggcatgga ggaggcgggtg 120

pctau2012001598-seq1

gaggctgtgg	ttgccgccgg	agtggagccg	ttcccacagg	tggacaagtg	cgacgcctcg	180
ggcctcggcg	cgcactcctc	cgtgggtggcg	gaccttgacg	gcacgctgct	ccgggtcccgc	240
agcgcgttcc	cgtactacgc	gctgggtggcc	ttcgagaccg	gcggcgctgcc	acgcctgctc	300
ctcctcctgc	tcctctcccc	gctggccgcc	ttgctgcgct	tcctggcgtc	ggagtccgcg	360
tgcctgcgcg	tgctgggtgtt	cggtgccacg	gcggggcgccc	gtgtccgcga	cgtcgagtcc	420
gcggcgcgcg	ccgtgctccc	caggttctac	gccgccgacg	tgcacccggc	ggcctggcgc	480
gtcttctccg	cgtgctcccg	gaggcgcgtc	gtgctcacgg	ccacgcccag	gctcttcgcc	540
gagccgttcc	tcagggactg	cctcggcgcc	gacgccgtcg	cgggcaccga	gctcgccacg	600
tggcgcgga	gcgccacggg	cttgggtggac	gcgcgcaggg	gcggcgctcct	cgtcggcagg	660
agcaaggccc	aggcgctgcg	ggagatcttg	gccgccgacg	gcggcgacgc	gcccgcgctc	720
ggactcgggg	acagccgctc	cgactacccc	ttcatgagca	tgtgcaagga	ggcgtacatc	780
gtaccgcgcg	cgccggtgga	gccggcattg	cccatggacc	agctgcctcg	tccggtcatc	840
ttccacgacg	ggcgcttggg	gcgccggccg	acgccgctgg	tggcgctggg	ggtggtgctg	900
tgggtccccc	tggggcttgc	gctggcctgc	ctccgcatcg	cggcgggggc	gctgctgccc	960
atgccgctgg	tgtactacgc	cttctggggc	ctgggcgtgc	gcgtcgctgg	gcggggcgcg	1020
ccgccgccgc	cgcgcgccga	gcgcgccacg	gggcggaggg	gcgtcctctt	cgctgctcg	1080
caccggacgc	tgctggacct	catcttctcg	tccacggcgc	tgggccgccc	cgtcgccgcc	1140
gtcacctact	cgctctccag	gctctccgag	ttcctgtcgc	ccatccggac	ggtgcgcctg	1200
tcgcgggacc	gcgccacgga	cgccgccatg	atcagggacc	tgctggcgga	gggagacctg	1260
gtcatctgcc	ccgaggggac	cacgtgccgg	gagcccttcc	tgctgcgctt	ctcggcgctc	1320
ttcgccgagc	tcaccagcga	ggtggtgccc	gtcgccatgg	agaacaggat	gagcatgttc	1380
cacggcacca	ccgccagggg	gtggaaaggg	atggaccctt	tctacttctt	catgaacccg	1440
agcccgccct	acgtcgtcac	cttcctcaac	aagctcccgc	cggagctcac	ctgcgccggc	1500
ggcaggacca	gccacgaggt	ggccaactac	atccagaggg	tcacgcgccg	cacgctctcc	1560
tacgagtgca	caagcctcac	caggaaggac	aagtaccggg	cgcttgagg	caacgacgga	1620
gtcgtcgcca	cccaaagcc	gccgccaacg	gttgctagct	agctagctaa	ggttattatt	1680
attattattt	tttttaaaaa	aaaagagaga	gaaagaaaat	tgggttatta	tattgtggtg	1740
atgatgctac	tgaattcatt	gtgattgtta	tgatgc			1776

<210> 130

<211> 1744

<212> DNA

<213> Ricinus communis

<400> 130

catttataac	attttcagag	agagagagag	agagagtaat	ataaatatgc	cggcgatgaa	60

acaggctcaa	aagtttccgt	caataacatc	atgcagcggc	acctcatagc	aatcaatagc	120

pctau2012001598-seq1

agctgatctg gatggtactc tgctcgtctc cagtagctcc tttccttact tcatgctcgt	180
ggccgtggaa gccggtagtc tgttgctggtg tctagtgttg ctgctttcgc tgccttttat	240
catcatctct tatttcttca tatcagaagc catcgggacg caaatcctaa tcttcatctc	300
tttcgctggt ctcaaaatcc gtgacatcga gcttgtttcc cgcgctgttc tgcctagatt	360
ttatgctgcg gatgtgagga aagaaagcta cgagggtgttc gatagggtgca agaggaaagt	420
ggtggtgacg gcgaatccga cgattatggt ggagccgttt gtgaaggatt ttcttggcgg	480
agacaaagt ttgggaacgg agattgaagt aaatcctaag acaaagaggg ccaactggttt	540
cgtgaagaag cctggcgtgt tggttggcaa atggaagaag ttggccattt tgaaagagtt	600
tggagaggac gtcctgatc ttgggatcgg agaccgcaag actgatcatg acttcatgtc	660
catttgcaag gagggctaca tggtttatca cagcaaaaaa gcagctactc cactgccacg	720
ggatcgctg aaaagtccta taatcttcca cgatggccgt tttgttcaac gaccggaccc	780
actcaatgcc ctgctacct atctttggtt gccatttgga ttcattgctct ccatctttcg	840
ggtatacttc aatctaccac tccccgaacg catcgtacgg tacacatacg agatgttagg	900
catccacctt gtgatccgtg ggtaccacc tccagcgcca tcacgtggaa cccctggaaa	960
tctatatgtg tgcaaccacc gtacagcttt ggatcctatc gtgattgcaa tcgcactcgg	1020
acgcaaagt tcgtgtgtaa catacagcgt tagccgtctc tcgagattcc tatcaccaat	1080
cccagccatc gctttgacac gtgatcgtgc ggctgacgca gaacgcatca cagctctact	1140
ccaaaagggt gatctttag tgtgtcctga gggaactaca tgtcgcgagc agtttttgtt	1200
gagattcagt gcaactgttg ctgaaatgag cgataggatt gtgcctgtgg cggttaattg	1260
caagcaaagt atgttctatg gcaccactgt acgtgggggt aaattttggg acccttattt	1320
tttctttatg aatccaaggc caacttatga ggtgacattt cttgatcggg taccgaaga	1380
gatgacggt aaagcgggag ggaaatcatc gattgaggtg gctaattatg tgcagaagg	1440
tttgggtgat gtgttgggtt ttcagtgtac tggattaacg aggaaggata agtacttgtt	1500
gcttggaggg aatgatggca aggtggagtc tatgtacaac tccaaaaagt aagcatttgg	1560
tttttcgttt ctataagatt tttagttacg gatgtttgta aggccagaca tatcaaatat	1620
tttagacatt gatgtcgtcg tgctgaattt gcgcacagct ttatatgcaa gactctttta	1680
agaatttgga tgtccgattt catggacgtg actgggtaag atttcatgat gttataagg	1740
aatt	1744

<210> 131
 <211> 1539
 <212> DNA
 <213> Ricinus communis

<400> 131	
atggcaaacc ctaaagaaac atatgctcct tttagctcat tcccaaatat agaaaagtgc	60
acatccgtag gtcgacaaaa cgacactgta gtgaccgata tggatggaac cctactttgt	120
ggacgtagcc cgtttcctta ctttgcctta gttgcctatg aagctagtgg aatattaagg	180

pctau2012001598-seq1

cttctattct	tgctattagg	ttcgcctatt	gctggaattc	tccgctactt	catctccgaa	240
tctgctggaa	tccgagtact	tgtatttgca	acctttgcag	ggatgcgagt	gtctgatata	300
gaatcgggtg	caagagctgt	tttgcccaag	ttttattcta	gtgacttgca	cccagagact	360
tggcgggtgt	tctcttcgtg	tgggaagaag	tgcgtaatca	cggaaaatcc	taggattatg	420
gtagaggcat	tcttgaaaga	gtttgttggg	gctgatatgg	ttattggcac	tgagattggt	480
gtgcataaag	ggagagcgac	tgggtttatt	aggagtcccg	gtgttcttgt	tggacagaac	540
aaggctgatg	cgctaaaaat	ggctttctat	gatgcgccgg	tgccaaatat	tgggatcggg	600
gacaggagga	ctgatcgtcc	attcatgaaa	ttgtgcaagg	agagctacat	agtcccaagc	660
aattctcaag	tagagccagt	tggctctaaa	aagttaccaa	aacccatagt	tttccatgac	720
gggcggcttg	tccaaaaacc	aacacattta	atcgcgcttg	ctactcttct	ttggattcca	780
gtaggcttta	ttctcgctg	tctacgctta	gccgccgggtg	cgctccttcc	aatgccatta	840
gtgtactacg	ccttgtgggc	acttgggtgt	cgtattcaca	taaaaggcac	accaccgccg	900
ccagccaaga	aatccactgg	ccaagctgga	gtcctcttcg	tatgttccca	tagaactctt	960
cttgaccggg	tcttcctctc	catcgcttta	ggacgtccaa	tcccagccgt	cacttactcc	1020
ctttcgcggc	tatcgaggtt	catttcaccc	atcaagactg	tcaggctcac	tcgagaccgc	1080
gttactgatg	caaacatgat	caaggaacta	ttacaagaag	gtgacttagt	catatgtcct	1140
gaagggacta	cttgtagga	accatttttg	cttcggtttt	cagctttggt	cgcgaggtta	1200
accgatgaac	ttgtccctgt	cgccatggct	aataagatga	ctatgtttca	tggaacaaca	1260
gctagaggat	ggaaagggat	ggaccccttc	tatttcttga	tgaaccctag	ccctgcatat	1320
gaagtgacat	ttttgaacaa	gttgcccttac	aatttaacat	gcgagagcagg	aaaatctagc	1380
catagtgttg	ctaattatat	acaaaggacg	attgctgcaa	ctttgtctta	ccaatgcacc	1440
agctttacta	ggaaggataa	atatagagca	cttgctggca	atgacggaac	tgctgtggag	1500
gagcccaagc	aggctcgtga	tgaagtcatt	gattcctag			1539

<210> 132
 <211> 1515
 <212> DNA
 <213> Ricinus communis

<400> 132	
atggttatgg	60
tggaagcgtt	
cccagaagtt	
agcaaattgt	
cgtccaaagg	
ccgagaaaag	
cacacgggtg	120
tcgctgatat	
ggatggcacc	
ctgtttattg	
gtcgtagctc	
cttcccttat	
ttcgtcttaa	180
ttgccttcga	
ggctgggtgg	
attttttaggc	
ttcttttctt	
gctcttggca	
tcaccacttg	240
cggcacttct	
ttattacttt	
gtttcagagt	
ctgctgggat	
tcaagttctt	
atctttgcat	300
cctttgctgg	
tatgaaagta	
tcggatattg	
aatctgtagc	
acgcgcagtg	
ctgccaaagt	360
tctactccag	
cgatctccat	
cctgagtcac	
ggcgtgtgtt	
ctcttcttgt	
gggaaacgct	420
gcgttcttac	
agcgaacca	
agaattatgg	
tggaagcatt	
tctgaaggat	

pctau2012001598-seq1

ttcttgggag	ctgatttagt	ttttggtact	gaaatatcaa	cttataaagg	tagagctaca	480
ggttttggtc	gtacccttg	tgtacttgta	ggcaagaaca	aggcaaagtc	tcttaaaaag	540
gcttttgag	acgcaaggcc	agaagttggg	cttgagata	ggcgactga	tgctcccttt	600
atggctttgt	gcaaggagg	atacctggtg	ccaccaagc	cggaagtcag	ggcagtaaca	660
ggtgacaagc	tcccaaaacc	catcgtcttc	catgatggca	ggctagtcca	gaaaccaaca	720
ccattaatgg	cacttctcat	cattctctgg	ataccaatag	gcttcattct	agcttgcttg	780
agaattgcag	caggatcact	cctgcctatg	ccaatggtct	actatgcctt	cttagcactc	840
ggtgtccgtg	tcacagtaaa	aggcaacca	cctcctccag	ccaagaaatc	aactgggtcaa	900
tcagggtgcc	tctttgtttg	ctccacaga	actctccttg	accccatatt	cctctccact	960
gcccttggcc	gccctattgc	cgcagtcaca	tactcagtat	ctcgcctctc	ggaaattata	1020
tcacccatca	aaacagttag	gctcagccgc	gatagagcaa	cagatgcagc	aatgatcaag	1080
aagttgtag	tagaaggtga	cttagcgata	tgtcctgaag	ggaccacttg	cagggaaccc	1140
tttcttttaa	ggttttcagc	tttgttcgca	gaattaaccg	atcaaattgt	gccggtagct	1200
atggtgaaca	gaatgagcat	gttccatgga	actaccgcta	gaggggtggaa	agggatggat	1260
ccattttact	tcttcatgaa	ccctagtcca	gcttatgaag	taactttctt	gagtaagttg	1320
ccccaagaat	taacgtgcag	ttcaggcaag	tctagccacg	aagtggctaa	ttacatacaa	1380
agggtgattg	ctgctactct	ttcctatgaa	tgcactggat	ttactagaaa	agataagtac	1440
agagcacttg	ctggcaacga	tggcagtgtt	gttgagaaac	ctaagctcga	agccaacaaa	1500
gtaatgggtt	gctaa					1515

<210> 133
 <211> 1804
 <212> DNA
 <213> Arabidopsis lyrata

<400> 133	
tttgtgctct	ctgttcttta
gagccgcat	taacgagaga
gagagagaga	gagagagact
60	
tggcgagtta	taatgtctcc
cgcaagaag	agtggaaact
ttcctccgat	aacggaatgc
120	
agagacggag	agtacgattc
gatagccgcc	gatcttgacg
ggactctgct	tctctcaaga
180	
agctcgttcc	cttacttcat
gctggtcgcc	gttgaagctg
gaagcctttt	ccgtggacta
240	
atactgcttc	tctcattacc
aatcgtcatt	atttcttacc
tattcgtatc	cgaatctctt
300	
ggaatccaga	tcctcatctt
catctccttc	gctggtctca
aaatccgcga	tatcgaactc
360	
gtctctcgcg	cagttcttcc
acggttttat	gcggcggatg
tgaggaaaga	cagttttgat
420	
gtgtttgata	agtgtaagag
gaaagtggta	gtgacggcga
atccgatagt	gatggtggag
480	
gcgtttgtga	aggattatct
tggagggtgat	aaagtttttg
gaacagagat	tgaagttaac
540	
cctaaatcca	atagagccac
tggattttgtg	aagaagcctg
gtgttcttgt	tggatgatctt
600	
aagaggtag	ccattttgaa
agagtttggt	gacgaatcac
ctgatctcgg	cctcggtgat
660	
cgaacctctg	atcacgattt
catgtccctc	tgcaagaaag
gttacatggt	tcatgcgtgc
720	

pctau2012001598-seq1

aagtcagcca	caacaattcc	aaaagaacgt	ttaaagaacc	gcatagtctt	ccacgacggg	780
cgtttagcgc	aacgtccaac	gcccttaaac	gccattgtca	catacctatg	gcttcccttt	840
ggtttcatac	tctccatcat	tcgctgtctac	ttcaacctcc	ctttacctga	acgctttgtc	900
cgttacacct	acgagattct	cggaattcac	ttaaccattc	gtggccatcg	tcctccacct	960
ccttcccccg	gaactcttgg	caacctctat	gtccttaacc	accgtaccgc	gcttgatccc	1020
atcatcgttg	ctattgctct	tggacgcaag	atctgttgtg	ttacttacag	tgtctctcgt	1080
ctctccctta	tgctttctcc	tattcctgct	gttgccctca	cccgtgaccg	tgtcaccgat	1140
gctgccaaaca	tgagaaaact	tctcgagaaa	ggcgacttgg	tgatatgtcc	ggaaggcaca	1200
acgtgcagag	aagagtatct	actgagattt	agcgctctat	tctcagagct	aagtgaccgg	1260
attgtgccag	tagcgatgaa	ctgtaaacia	ggaatgttca	acgggaccac	agtgaggggt	1320
gtgaagtctt	gggaccctta	cttctttttc	atgaacccaa	gaccaagcta	tgaagccact	1380
ttcttggatc	gtttgcctga	agaaatgact	gtcaacggtg	gcggaagac	tcctatagag	1440
gtggccaatt	acgtccagaa	agttatcggc	gcggtcttgg	gcttcgaatg	caccgaactt	1500
actcgcaagg	ataaatatct	tttgcttggg	ggtaatgacg	gcaaggtgga	gtctatcatg	1560
ttgggataaa	cttgaaataa	cttaagagat	aagttactaa	aatgtaaaag	agataaagct	1620
gaatatctta	ggagtattct	agaataaatg	tattaggagt	tttaattgat	gatgtattgg	1680
tacgttattg	ttatgttaat	tctatttttc	cctaaaacta	ttttgtgatc	acatttctat	1740
acgcatgtgc	ttaagggata	ataagaactt	tatatttaaa	gagcaaaatt	ttattggagg	1800
aagt						1804

<210> 134
 <211> 1746
 <212> DNA
 <213> Arabidopsis lyrata

<400> 134	
atatccattc	tcaccttcca aatctctctc ttctcttctc tttccggctt cttcactaat 60
gggagctcag	gagaaacgac gtcgtttcga gcctatatca aagtgcgacg ttaaggaccg 120
gtccaacccat	accgtggccg ctgatctaga cgggacacta ctgatctcac gtagcgcctt 180
cccttactat	ttcctcgtag ccctcgaggc agggggcttg ctccgagcat tgatcctact 240
cgtgtccgta	ccgttcgtct atcttacgta cttgaccgtc tcggagactt tggccatcaa 300
cgtttttgtc	ttcatcacgt ttgcgggtct caagatccga gacgttgagc tagtggtccg 360
ttccgtcctc	ccgaggttct atgcggagga cgtgaggccc gatacgtggc gtatctttaa 420
cacgtttggg	aaacggtaca tagtaacggc gagtcctcga atcatggtcg agccgttcgt 480
gaagacatac	ctaggagtgt ataaagttct cggaacagaa ctagaggtct ccaaatacggg 540
tcgggcaacc	gggttcacca gaaaaccggg tattctcgtg ggtcagcaca aacgtgaagt 600
cgttttgaga	gagtttggtg gccttgctgc tgatttacct gatttggggc tcggcgatag 660

<400>	135						
atgtcgc	aaa	cga	aaccggc	tcc	aaaattc	ccttcaataa	catcatgtac
cg	gctcggcg						60
tatcag	tcaa	tagcggctga	tcttgatggc	acgctcctcg	tgtccagtag	ctccttccct	120
tacttcatgt	tagtggctgt	cgaagctgg	t	agtctcctcc	gtggctctgg	gtcgccttctg	180
tctctacctc	ttgtgatcat	ctcctat	ttt	ttcatatctg	aagccgctcg	aatccaaatc	240
cta	atctata	tctcctttgc	tgggctcaag	atccgcgaca	tcgagctgg	ttcccgcgct	300
gttttgccta	ggttttatgc	tgctgatgtg	aggaaggaga	gttttgaggt	ttttgacaaa		360
tgcaagagga	aggtgggtgg	gacggcaaat	ccgacgatta	tgg	tggagcc	attt	gtgaag
gattttcttg	gcggagacaa	ggtgttaggt	acggagattg	aagtaaatcc	aaagactaag		480
agggccaccg	gattcgtgaa	gaatcctggc	gtgttagtg	ggaaatggaa	gaaattgtcc		540
at	tttgaaag	agtttgga	ggaatcaccg	gatcttgga	tcggagaccg	caagacggat	600
catgatttca	tgtccatttg	caaggagggc	tacatggttc	aacgcagcaa	atcagctact		660
ccaataccac	tggatcgtct	caaaagtcgt	ataattttcc	atgatggccg	gttcgttcag		720

pctau2012001598-seq1

cgacctgacc	cactgaacgc	tcttgtcacc	tatctttggc	tgccatttgg	attcatcctc	780
tctattattc	gcgtatactt	caatctccca	ctcccagaac	gcatcgtagc	gtacacatac	840
gagatgctag	gcatccacct	tgtgatccgt	gggaacccgc	ctcctgcccc	atcacctgga	900
accccaggaa	atctctatgt	ttgcaaccac	cgctcagctt	tgatcctat	tgtcattgca	960
atagcacttg	gacgcaaagt	ttcctgtgta	acatacagcg	taagccgtct	ctcgagggttc	1020
ctatctccga	ttccagctat	cgctttgact	cgtgatcgtg	ccgctgatgc	agcacgcatac	1080
tcagagctac	tccaaaaggg	tgatcttgta	gtgtgtccag	aggggactac	gtgtcgcgag	1140
ccgttcctgt	tgcgattcag	tgcattgttc	gcagaaatga	gcgataggat	tgtgccggtg	1200
gcggttaact	gtaagcaaaa	tatgttctat	ggtaccaccg	tacgtggtgt	caagttttgg	1260
gatccttatt	acttcttcat	gaaccaaga	ccaacctatg	aggtgacatt	ccttgatcgg	1320
ttgccggaag	agatgacagc	taaggcggga	gggaaatcgt	cgatcgaggt	ggctaattat	1380
gtgcagaagg	tgttgggtga	tgtgttgga	ttcgaatgca	ctggattaac	tagaaaggat	1440
aagtatatgt	tgctaggagg	aatgatggc	aaggtggagt	caatgtacaa	caccaaaaaa	1500
gccggttaa						1509

<210> 136
 <211> 1545
 <212> DNA
 <213> Oryza sativa

<400> 136	
atgggggacg	ccacgatgat
cgcgctgcgcg	ttcccggcgg
tgagagaagtg	ttcgtccaag
	60
gaccgcggcg	gcgacacgat
cgtcgcgcgac	ctcgacggca
cgctgctgtg	cggccggagc
	120
tcgttcccgt	acttcgcgca
catggcggttc	gagacggggg
gcgtcctccg	gctgctggcg
	180
ctgatcctgc	tggcgccgct
cgccggcggtg	ctctactacc
tggtgtccga	gtcggccggc
	240
atccaggtgc	tcattcttcgc
gtccatggcc	ggcgcgaggg
tcgccgacgt	tgaggcggtg
	300
gcgcgcgcgcg	tgctgcccga
gttctactgc	tccgacatcc
acccggagtc	gtggcgcggtg
	360
ttctcggcgt	gcgggcgcgcg
gtgcgtgctc	accgcgaacc
cgaggatgat	ggtggaggcg
	420
ttcttgaagg	agtacctcgg
cgccgacatc	gtcgtgggca
cggagctcgt	ggtctggcgt
	480
ggccgcgcca	cggggctcgt
ccggagcccc	ggcgtgctcg
tcggcgagaa	caaggcggcc
	540
gcgctccggg	aggcgttcgg
cgacgcctcg	ccggagatcg
gcatcggcga	ccggaggacg
	600
gactaccctg	tcattgaggct
atgcaaggag	gggtacgtgg
tgccggtgtc	gccgacgccg
	660
aggcccgtgc	cgcgggagga
gctcccgaag	ccggtggtgt
tccacgacgg	ccgcctggtc
	720
cagaagccgt	cgccggcgct
cgcgctgctc	accgtgctct
ggatcccat	cggcttcgtg
	780
ctgcctgcc	tccgcatcgc
cgccggcgcg	ctcctccga
tgcgctgggt	gtaccacgcc
	840
ttccgcgccc	tcggcggtgcg
cgtcacgcgtc	aagggcaccc
cgccgccacc	ggccagccgc
	900
gagaccggcc	agtccggcgt
gctcttcatac	tgctcccacc
gcaccctcct	cgaccccatc
	960

pctau2012001598-seq1

ttcctctcca	ccgccctcgg	ccgccccatc	accgccgtca	cctactcggg	gtcgcgggctg	1020
tcggagatat	tgtcgccgat	caggacgggtg	cggctgacgc	gggacagggc	ggcggacgcg	1080
gcgatgatcc	ggcgggtgct	ggcggagggc	gacctggtga	tctgcccgga	ggggacgacg	1140
tgccgggagc	cgttcctcct	ccggttctcg	gcgctgttcg	cggagctcac	cgacgagatc	1200
gtgccgggtg	cgatggagaa	ccagatgagc	atgttccacg	gcacgacggc	gaggggggtg	1260
aaggcgctgg	accctttcta	cttcttcatg	aaccctagcc	ccggatacgt	ggtcaccttc	1320
ctcagcaagc	tccccggcga	gctcacctgc	aacggcgggtg	gcggcgggcg	cggcgggccg	1380
agcagccacg	aggtggccaa	ctacatccag	aggctcatcg	cctccacgct	ctcctacgag	1440
tgcaccagct	tcaccaggaa	ggacaagtac	agggcgctcg	ccggcaacga	cggcaccgtc	1500
gtgtccaagc	ccaacatcga	cgagaagaag	gccatgggat	gttaa		1545

<210> 137
 <211> 3343
 <212> DNA
 <213> Brassica napus

<400> 137

aaattaaaca	aatttgatag	aatatacatt	ttttttatca	gatctttatt	attcaaaatc	60
attaattgtc	atatatactt	tatccacatt	aggcaattcc	gtaatcttta	tttaaggaaa	120
taataaatga	cattaataat	gaatttatgg	ttagtttaat	aaaaaaactt	attatataat	180
tagatgacca	acctatttct	ctaataattc	taagaatcat	cctagtata	acatgtgtct	240
acaaaaagaa	gttgtaatgt	ttcacaata	atatacaggg	gatcagttat	tttcatcaac	300
atagttggat	acacagctgt	atagacatta	attggtaaag	ctatacatag	caaagtata	360
ttaatgagtc	actcaactaa	tgagacggcc	ttctatttag	gtgagctggg	aagctacttc	420
ctttcatgtc	gatcaacaaa	aaataaataa	ggcttttctt	ttagtcctaa	ctccgaagag	480
aaccgcgtga	cctaaactat	agtatttatt	acacatattg	atccacgatt	gatagataat	540
attcaccact	tttcattata	tttacatact	aatgcaaacc	tatgtagtaa	aattaatcac	600
atattgaatt	atgacttata	aaatgtaatc	gtataacaaa	aagatggatg	gtttaagcca	660
aagactttgt	tacgactctt	gtaatgggga	atttatatat	ttaagtgaga	atgatgacaa	720
agtgagaaa	aatcagggtta	taattgttag	aaatgatagt	ttaaggataa	atttgtaaag	780
aagtcataat	ggataaggag	tgacaatatt	ataaatcggg	aaggaatgca	gcgcatttaa	840
aacaagaaa	accctttccc	tctctttctt	tccaagattc	tctctctctg	tttaccatca	900
aagagagatc	gaaaaatgtc	tccggcaaag	aagaagagca	gaagctttcc	tccgataagc	960
gagtgtaaac	gtggagagta	tgattcgatc	gccgcagatc	tcgacgggac	tatgctcctc	1020
tcgagaagct	ccttccctta	tttcatgtc	gtcgccatag	aagcaggcag	cctcttccgt	1080
ggattgatcc	tccttctatc	tctcccgatc	gtcatcatcg	cctacctctt	cgtctccgaa	1140
gctctcggca	tccaaatcct	catctacatc	tccttcgccg	gcatcaaaat	ccgcgacatc	1200
gagctcgtct	cccgcgccgt	tcttccacgg	ttctacgcgg	cggatgtgag	gaaggacagt	1260

pctau2012001598-seql

ttcgaggtgt	ttgataaatg	caagaggaaa	gtggtggtca	ctgctaatacc	catagtgatg	1320
gttgagcctt	tcgtcaagga	ttacttgggg	ggagataaag	ttttgggaac	agagatcgaa	1380
gtgaatccca	aaacgatgaa	agccacgggg	tttgtgaaga	agcctggtgt	tcttgttgga	1440
gacttgaaga	gattagccat	cttgaaggag	ttcggggaag	aatcaccgga	tcttggcctc	1500
ggtgaccgca	cctccgatca	tgatttcatg	tccatttgca	aggtactata	aacatatacg	1560
tttaagtttc	ttgaattttt	aagaaccccg	gtttaggtga	aaatcgagtg	cataaacgtc	1620
tgatactttg	aaacatttcg	taatcaaaga	agtaagtact	tgaaaactac	acataaatgc	1680
agtaggacct	atttaatgcc	gtgttgctg	ttcacatgcc	ttcacatgat	ctattcactt	1740
ttttttacgt	ccaaattatt	tgtatgtgtg	tgtatgtata	ataaaaacac	gcatgaaata	1800
tctagattaa	gtagttttgt	tatttgaatt	caagaaagga	acgtaataag	tttccggtga	1860
caaaaaaatg	tagataatag	aatagaatag	tagtagtgat	tttcataaat	gcttcctacc	1920
ttatcaccaa	ccgcccattc	ccacataagt	tgaatggtcc	catttaatga	ccgaacccat	1980
atgtcaacat	tcctttcttt	ctttttctcc	taaccactga	aactttcagt	aattttacgag	2040
ctttcttact	ataattgtgt	aactttggaa	tcttttacta	catgattggt	aacgacttgc	2100
ttttctcttg	taacaggaag	gttacatggt	gcatgagtcc	aaatcagcca	caacagtccc	2160
catagaacgt	ctcaagaacc	gtataatctt	ccacgacggc	cgtctagtcc	aacgccttac	2220
gccgttaaac	gccataatca	tttacctttg	gcttcctttc	ggtttcatgc	tctctatctt	2280
ccgcgtttac	ttcaacctcc	ctttaccgga	acgcttcggt	cgctataact	acgagatcct	2340
cggtattcac	ctaactatcc	gtggccaccg	tcctcctcct	ccttccccag	gcacacctgg	2400
aaacctctac	gtccttaacc	accgtactgc	tcttgacccc	attatcatcg	ccattgctct	2460
cggccgtaag	atctcttgcg	tcacttatag	tgtctctcgc	ctctcccgga	tgctctcacc	2520
catccctgcc	gtcgctctta	cccagagaccg	cgccgctgat	gctgctcgca	tgagaaaact	2580
cctcgagaaa	ggtgacatca	acccttggtta	ctatcattct	atatcattgg	tttcatcatc	2640
tcatgagttc	caatatgaaa	ataatatagg	ttttgaaagc	cttcagtcaa	aaaataaaaa	2700
tattgtaggt	ttccatgtca	cggttacttt	taagaatcaa	ctgcataact	agaaaacata	2760
aaacatagca	tctttgtgtg	ttttctatac	atattatcac	gctaggcaca	aaagatagtt	2820
ggagaatcag	atagcgttcg	tctatataaa	agtgggaaac	agagcatcta	tgtttactag	2880
atataaggtc	catcacataa	gtctttttatt	atttttgtaa	caggtgattt	ggtgatctgc	2940
cctgaaggca	ccactttag	agaaccgtat	ctactacgat	ttagcgctct	atttgcagag	3000
ctcagcgacc	ggatcgtgcc	cgtggccatg	aactgcaagc	aagggatggt	caacggaacc	3060
acagtacgag	gtgtaaaatt	ctgggaccct	tacttcttct	tcatgaaccc	gagacctagc	3120
tacgaggcca	ctttcttgga	ccgtttgcc	gaagaatgga	cagtaaaccg	cggcggtaaa	3180
actccaatag	aggtggcaaa	ttacgttcag	aaggtgatcg	gtggagtatt	gggattcgag	3240
tgactgagc	tcacaaggaa	ggataagtat	cttttgcttg	gaggcaacga	cgggaagggt	3300

pctau2012001598-seql

gagtctatca ataagagcaa ggccaaggag taataaatcc agc 3343

<210> 138
 <211> 1503
 <212> DNA
 <213> Populus trichocarpa

<400> 138
 atgtcgccaa cgaaacccgc aaaaaaattt ccgccgataa cggcctgtaa cggcacgacg 60
 caccaatcta tagcagcggg tcttgatggc acactccttg tgtcgagtag ctcgtttcca 120
 tactttatga ttgttgctgt tgaagctgga agcctctttc gtggctcttg tttacttctt 180
 tctactgccta tcgtcatcgt gtcttacctc ttcatatctg aagccttggg gatacaaatg 240
 ttaatcttca tttctatgtc tgggctcaag atccgtgata ttgaacttgt ttcacgcgct 300
 gtattgccta ggttttatgc tgcagatgtt agatcggaga gttttgaggt gtttgatagg 360
 tgcaagagaa aggtgggtgt gacagcaa atccaactataa tgggtgagcc ttttgtgaag 420
 gattttctcg gtggagataa ggttttgggc acggagattg aagtgaacc gaagaccaag 480
 agggctacag gatttggtta gaagcctggg gttctggtgg gcaagtggaa agaattggct 540
 gttttaaagg agtttgaga ggaagcaccg gatcttggtt ttggtgaccg caagactgat 600
 catgatttca tgtccctttg caaggagggc tacatggtgc atcgagcaa atcagcgact 660
 ccactaccgc gggatcgctt caaaaatcgt atcatcttcc atgatggccg tcttggttcaa 720
 cgtccagacc cactcaatgc cttgatcact tacatttggg tgccggttg attcatcctc 780
 tccatcattc gcgtctactt caacctcca ctcccagaac gcatcgtag gtacacatac 840
 gaaatgcttg gcattcacct ggtcattcgt gggacccac ctctgcccc atcacctggg 900
 accccaggaa atctctatgt ttgcaatcac cgtacagccc tggatccaat tgtcatagcc 960
 atagcacttg gacgcaaagt ctcgtgtgtt acgtacagcg taagccgtct ctcgaggttt 1020
 ctttcaccaa tcccagccat cgctttgact cgtgatcgcg cggctgatgc tgcaagaata 1080
 tcatccatac ttcaaaaggg tgatctagt gtgtgtccgg agggaaccac ttgtcgtgag 1140
 gagttcttgc tgcgatttag tgcattgttt gcagagttga gtgataggat tgtgcccggtg 1200
 gcagtcaatt gtaaacaaaa catgttttat ggcacgaccg tgcgtggtgt caagttctgg 1260
 gatccttatt ttttcttcat gaacccgagg ccaacttacg aagtacatt ccttgatcgc 1320
 ttgccagagg aaatgacggt caaggctggg ggcaaactct cgatcgaggt ggcaaactat 1380
 gtgcagaaaag tgttggtgga agtgtagga ttcgagaaca ctgggttgac caggaaggac 1440
 aagtacctgt tgcttggtgg gaatgatggg aaggctgagt caatgcacaa tgccaaaaaa 1500
 taa 1503

<210> 139
 <211> 2117
 <212> DNA
 <213> Sorghum bicolor

pctau2012001598-seql

<400> 139
gccgagctcg actcgggaatc cgatccccc a cccaactcca tgctccgatt cctcctccct 60
tctctgtctc gactccgacc accaccacct ctgtctgccag ccccgctcta catataaatc 120
cctccgtcat catcgtcaac cgcagcggtg ggccgctcaa caaacgcagc caccgccgcc 180
gccgccacca cgacaccaac catataatcg cctgccc aag tcccaacctt ttccgacca 240
gccagacacc accaccaccg tcgtcttcct cgccatggcc cccgcggcga gcgcgaggtt 300
cccaccggtg tcattcctac acgcgtccgc acgcacccgc cgaaccgccg ccgcggacct 360
ggacggcacg ctgtctcgct cctcctcggc gttcccctac tacttcctcg tggcgctcga 420
ggccggctcg tgctcccgcg cggcggcgct cctcctcgtg gctccgctcc tgctagcgct 480
ctacacgctg gtgtccgagg cggcggccat cgcgctgctg gcgttcgtca cgttcgcggtg 540
gctgcgcgtg cgcgacgtcg aggccgtcgc ccgcggcgct ctgccgcggc actacgcggc 600
cggcgtgcgc gccgacacct ggaccgtgtt ccgcggctgc ggcgagggga ggagggctgt 660
cgtcaccgcg tccccgccg tcatggtggg cgagttcgta cgcgagttcc ttggtgctga 720
ggtcgtggc accgagctcg agaccttctc cgcgttcggc gccgccaggt ttacggggag 780
gatcaaggcc gtgctcgtcg gggagaggaa ggcggaggtc gtccggcggc tcttcgccgg 840
tggggagatg cccgacgtcg ggcttgggga ccgcgagagt gaccatgact tcatggccat 900
ctgcaaggaa gcttacatgg tgctccgga caagcgcgcg gcgcgtgcgg ccgccgacac 960
gctgctgtcc cgctccgtct tccacgacgg ccgcctcgtc cagcgcccgg acccggcga 1020
ggcgtgttc gcgtggcct acctcccgct gggcttcctg ctggcgctgt tccgcgtctt 1080
cttcaacctc atgatgccg tacgcctggt ccgccacacg taccgcctga cggggatccg 1140
cctccgcgtc cggggcacgc cgccgccgcc cccggcgccg ggcgccccgg gttcgtgtct 1200
ggtctgcaac caccgcacgg cgctggacct gatcatcctg tccgtggcgc tgggtcgccc 1260
cgtgtcgtgc gtcacgtaca gcgcgagccg cctctcgacg gcgatctcgc cgatccgggc 1320
cgtcgcgtg tcccgggacc gcgccaccga cgcggcgcg c atggcggcgc tgctggcgga 1380
gggcgacgtc gtcgtgtgcc ccgaggggac cacgtgccg gagccctgcc tgctccggtt 1440
ctcggcgctg ttccgccgag tcacggaccg gatcgtgcc gtggccatgg aggcgcggca 1500
gggcacctac tacgggtcga cggcgagggg gtggaagtgg ctcgaccct acttcttcta 1560
catgaacctg cggccaggct acgacgtcac gttcctgccg ccgctgcggc cggaggagac 1620
gtcggcgcc ggccgggagga gcgccgtcga cgtggctaac cacgtgcaga gggatcatgc 1680
caaggagctc gggttccagt gcaccacgt caccaggaag gacaagtaca tgaagctcgc 1740
cggcaacgac ggctcggtcg ccgccagggc taaaaaggac gccgccgatg ataacgtgc 1800
cacggtcacc accaccaaga agtttgtatg aagagatcga ttatatatat tggcaacct 1860
gctcgactat atcggagcag cattgtactg acgtaaaaaa aaaaaaaca tgcattgcaca 1920
aaaggcttag ttaggcttat gtatattgga gtatcgtgtt cttttattta tactacgtat 1980
acgtagaaaag aatggacgag tgaatggagt ccaattttta aggcgcaatt tatatatgat 2040

pctau2012001598-seq1

tatgattatt ttgggtagtg tatttgctta aaaatctaaa cacatactaa attgcattgc	2100
aacaatttaa atttttt	2117

<210> 140
 <211> 1547
 <212> DNA
 <213> Ricinus communis

<400> 140	
acacatatatt cccactgcta ctagcttccc aaaactcaca ccttaatcta tcaaattggt	60
catggggggt catcggcatt tcgagtcaat caagaaatgc agcaccgagg gacggtcgaa	120
tcagacggtg gctgcagacc tggacggcac acttcttggt tcatcaagtg cttttccata	180
cttcttgctc gttgcacttg aagctggaag cgtgttaaga gctatactcc ttttagcatc	240
tggtcctttc gtgtatttta cgtacctggt tttatcagag gcaatagcta tcaaaacctt	300
catctttatt gcctttgcag gacttaaaat aaggggacgtt gagcttgttt cgcggtctgt	360
cttgcctaga ttctacgagg atgatgtcca tcctgaaact tggagagtgt tcaattcttt	420
tgggaaaaga tatattgtca cgcgcagtcc taggattatg gttgaacctt ttgttaagac	480
cttttttaggt gctgacaaag ttcttggcac tgaactagag gttacaaaat cgggaaaggc	540
aactgggttc gtcaagaaac ctggaattct tgttgctgaa cataaaagag atgccatttt	600
aaaagaactg ggcgctgatt tgccagattt aggcctgtgg agaccgagaa actgaccaca	660
atttcatgtc tatttgcaag gaaggatata tgggtgccaag aacaaagtgt gagccactac	720
caagaaacaa gcttctaagc cctgtaatat tccatgaagg acggttagtt caaaggccaa	780
cacctctagt tgccctcttg acctttttat ggatgccaat cgggtataatt ctctcaatcc	840
ttaggggtata ccttaatatc cctttacctg aaagaattgc ctattacaat tacaaaattc	900
ttggaattaa agttatagtt aagggcaccc ctctctctgc tccaagaaaa ggccaaagtg	960
gagtcctttt tgtatgcaac catcgtactg ttttagacct agtagtgaca gcagttgcct	1020
taggcagaaa gatcagttgt gttacatata gcataagcaa attcacagaa atcatatcgc	1080
caatcaaagc agtcgcatta tcaagagaaa gagagaaaga tgctgctaatt attaacggtt	1140
tattagaaga aggtgacttg gtcatttgcc ctgaaggaac aacttgtaga gagccatttc	1200
ttttgagatt tagtgccctt tttgctgaac ttaccgacag aattgtgcct gttgctatca	1260
ataccaagca aacagttttc cacggcacga ctgtccgagg gcataaatta ttggaccctt	1320
attttggtgt catgaatcca atgccaacct atgagatcac attcttgaat cagttgccta	1380
tagagcttac ttgcaaagggt ggtaaatacgt cgatcgaagt tgccaattac attcagaggg	1440
ttcttgctgg tacattgggg tttgaatgta ctaatttgac aaggaaggac aagtatgcta	1500
tccttgctgg tacagacggt cgtgttccat ctaagaagga aaagtaa	1547

<210> 141
 <211> 1549
 <212> DNA

<213> *Arabidopsis lyrata*

<400> 141

```

aagattcgcc gcaagaaaac atgtctccgg cgaagaagag ccgaagtttt cctccgataa      60
gcgagtgtaa aagcagagag tatgattcga tcgctgcgga tctcgacggc actctgcttc      120
tctcaagaag ctctttccct tacttcatgc tcgtcgccat tgaagcaggt agtctcttcc      180
gtggattgat ctttcttctt tctcttccca tcgtcatcat cgcttatctc ttctgtctccg      240
aatctctcgg aatccaaatc cttatctaca tctccttcgc cggatatcaa atcaagaaca      300
tcgaactcgt ctctcgcgcc gttcttccac ggttttacgc ggcggatgtg aggaaagata      360
gttttgaggt gtttgataaa tgcaagagga aagtgggtgg tactgcgaat ccgatagtga      420
tggttgagcc attcgtcaaa gattatttgg gaggagaaaa agttttggga acagagattg      480
aagttaaccc caaaacgatg aaagctactg gttttgttaa aaagcctggg gttcttgttg      540
gtgatttgaa gagattagcc atcttgaaag agttcggcga agaatacccg gatattggac      600
tcggtgaccg gacctccgat cacgatttca tgtccatttg caaggaaggt tacatggtgc      660
atgagaccaa atcagccaca acagtcccta tagagcgtct caagaaccgc ataatcttcc      720
acgatggccg tcttgccaa cgtccgacgc cgttgaacgc cttaatcatt tacctttggc      780
ttcctttcgg ctttatgctc tctatcttcc gcgtctactt caacctccct ttaccggaac      840
gcttcgttcg ctacacttac gagatcctcg gcattcacct cacaatccgt ggccaccgtc      900
ctccccctcc ttccccggc aaacctggaa acctttacgt ccttaaccac cgaaccgccc      960
ttgaccccat aattatcgcc attgctctcg gccgtaagat cacatgtgtc acttatagtg     1020
tctctcgctt ctccctgatg ctctcaccca ttctgtctgt tgctctgacc cgagaccgtg     1080
tcgctgatgc agcccgcatg agacaactcc tcgagaaagg tgatttggtg atctgtcctg     1140
aaggcacaac gtgtagagaa ccatacttac taagattcag tgctctatct gcggagctga     1200
gcgaccggat cgtgcctgtg gccatgaact gcaagcaagg gatgttcaac gggacgacag     1260
ttagaggtgt gaaattctgg gatccgtact tcttcttcat gaaccctcga cctagctacg     1320
aggccacttt cttggaccgt ttgccgaag aaatgacggg aaacgggtgg ggcaagactc     1380
cttttgaggt ggctaattac gttcagaagg tgatcgggtg agttttgggg ttcgaatgca     1440
ctgaacttac aaggaaggat aagtatctgt tgcttgaggg caacgacgga aaggttgagt     1500
ctatcaataa gaccaagtcc atggactaat aatctagcta gttactggt      1549

```

<210> 142

<211> 1995

<212> DNA

<213> *Arabidopsis thaliana*

<400> 142

```

ccaacacatt tacttttctt cttcttctct ctctacgcca tagctatggg ttaccagag      60
cttctcgtga ttctagcaga atgggttctt taccgtcttc ttgctaaatc atgttacaga     120
gctgctagaa agctcagagg ctatggtttc cagttaaaaa atctcctaag cctctccaaa     180

```

pctau2012001598-seq1

acacaatctc	tacacaacaa	ttctcagcat	catcttcata	accatcatca	acaaaatcac	240
cctaaccaaa	cccttcaaga	ttctctagac	cctctgttcc	caagccttac	aaaataccaa	300
gaacttcttc	tagacaaaaa	cagagcatgc	tctgtttcgt	ctgatcatta	tcgtgatacc	360
tttttctgcg	atatcgatgg	tgttttatta	agacaacatt	cgagcaaaca	cttccacacg	420
ttcttccctt	acttcatgct	cgtggctttt	gaaggaggat	cgatcatcag	agcgatcctt	480
ctcttgctct	cctgttcttt	cctttggact	ctgcaacaag	aaaccaaact	cagagttctc	540
tctttcataa	ccttctctgg	tcttagagtc	aaagacatgg	acaatgtttc	tagatcgggt	600
ctaccaaaat	tcttcctcga	gaacttgaat	atccaagtgt	acgacatctg	ggctcgaact	660
gagtactcca	aagttgtggt	cacgagcttg	ccgcaagtgc	tggtggagag	gttcctcaga	720
gaacatctca	acgctgatga	tgtgattggg	actaaactgc	aagaaataaa	agtcattggga	780
aggaaattct	acacaggcct	agcttctggg	tctggtttcg	tcctgaaaca	taaatcagct	840
gaagactact	ttttgatag	caagaagaag	cctgcacttg	gaattggaag	ctcttcaagc	900
ccacaagatc	acatcttcat	ctccatttgc	aaagaagctt	atttttggaa	tgaagaagag	960
tccatgagca	agaacaacgc	attgcctaga	gaaagatacc	caaagccact	gatattccac	1020
gacgggagat	tggctttctt	gccaacacca	ttagctacat	tagccatggt	catatggctc	1080
cctatcggat	ttttactcgc	tgtcttcagg	atctcggtag	gagtctttct	tccttaccat	1140
gtagccaatt	tcttggcttc	tatgtccggg	gttaggatca	ctttcaagac	ccataatttg	1200
aataacggtc	gtccagagaa	aggtaatagt	ggagtcttgt	atgtatgcaa	ccataggacc	1260
ctcctcgacc	cagttttcct	caccacttcg	ctgggcaagc	ccttgaccgc	ggttacttac	1320
agcctaagca	agttctccga	gttcatagca	cctctcaaga	cggttagcct	taaaagagat	1380
aggaagaaa	acgggtgaggc	aatgcagaga	ctattaagca	aaggagatct	agtggtttgt	1440
ccggagggaa	ccacatgtag	agagccttat	ctactaaggt	tcagtccttt	gttcgctgag	1500
ttactgaag	acattgtttc	cgtggctgta	gatgctcgag	tgagcatggt	ctatgggaca	1560
actgcgagcg	ggctcaagtg	tttagaccct	atcttcttcc	tcatgaacc	taggccgggt	1620
tattgtcttg	agattttaaa	gaagttgcca	aaagagatga	cttgtgcagg	aggcaagtca	1680
agctttgaag	ttgcgaattt	catacaggga	gaacttgcta	gggttttggg	atttgaatgc	1740
actaatctca	caaggaggga	taagtacttg	gttcttgctg	gaaatgaagg	cattgtcaga	1800
taaactattc	cgaattttca	ttagtttttt	tgtttgtttg	aaaagctaaa	tattattggt	1860
atttcaggac	ttagatgttt	aattctatat	ggttgattca	ttggcatcat	ttctatttct	1920
tatcactcgg	gaccattgat	atgtattgaa	aatgtttgct	ttatatgatt	ttgtttctgg	1980
atttgtatag	tccat					1995

<210> 143
 <211> 1828
 <212> DNA
 <213> Arabidopsis thaliana

pctau2012001598-seql

<400> 143
gtctgaaacc aaacaaagaa aatacacgaa ggaaggaaaa gaaaagtaaa aagaaaagct 60
ttctgcaaaa tccaaacatg tccgccaaga tttcaatatt ccaagctctt gtctttctat 120
tctaccggtt tatcctccgg cgatatcgga actctaaacc aaaataccaa aatggccctt 180
cttctctcct ccaatccgac ctatcacgcc acacattgat cttcaacgta gaaggagctc 240
tttcaaatc cgactctctc ttcccttact tcatgttagt agcatttgag gcgggaggcg 300
taataagggtc atttctcctc ttcattctct atccattgat aagcttgatg agccatgaga 360
tgggtgtcaa agtgatggta atggtagagct tcttcgggat caaaaaagaa ggttttcgag 420
cggggagagc ggttttgcct aaatactttc tagaagatgt cggactcgag atgttcgaag 480
tgttgaagag aggagggag aaatcggag tgagtgatga tcttcctcaa gttatgatcg 540
aagggttctt gagagattac ttggagattg acgttgtggt cgggagagaa atgaaagtcg 600
ttggaggtta ttatctaggt atcatggagg ataaaaccaa acatgatctt gtctttgatg 660
agttagtctg taaagagaga ctaaaccacc gtcgtgttat tggcatcact tccttcaata 720
catctcttca ccgatatcta ttctctcagt tttgccagga aatttatttc gtgaagaaat 780
cagacaagcg aagctggcaa accctaccac gaagccagta ccctaaacca ttgattttcc 840
atgatggccg tctcgcgatc aaaccaaccc taatgaacac tttggtcttg ttcatgtggg 900
gtcctttcgc agccgcagcc gcagcagcca gactcttcgt ctctctttgc atcccttact 960
ctttatcaat cccgatcctc gccttttccg gttgcagact aaccgtcact aacgactacg 1020
tttcatctca aaaacaaaaa ccaagtcaac gcaaagggtg tctctttgtt tgtaaccata 1080
ggactttatt ggaccctctc tatgttgcac tcgctttgag aaagaaaaac atcaaaactg 1140
taacgtatag tttgagtagg gtatctgaga ttttggtcc gatcaagacg gtgagactga 1200
cccgtgatcg ggtgagcgac ggtcaagcca tggagaaatt gttaaccgaa ggagatctcg 1260
ttgtttgtcc tgaaggaacc actttagtag aaccttacct gcttaggttt agccctttgt 1320
tcaccgaggt tagtgatgtc atcgttcccg tggctgtgac ggtacacgtg accttcttct 1380
acggtacaac ggcgagtggc ctttaaggcac ttgaccgct tttcttcctc ttggatcctt 1440
atcctaccta caccatccaa tttctcgacc ctgtctccgg tgccacgtgc caagatcctg 1500
atggaaagtt gaagtttagg gtggccaaca atgttcagag tgatattggg aaggcggttg 1560
atttcgagtg cacaagtctc actagaaaag acaagtatct gatcttggcc ggtaataatg 1620
gagtagttaa gaaaaattaa aatttgtttt aggatgatga gtgagtgatg agttcggttc 1680
tttgggtatg taagtaacca atcgattatt tcatgtgttt attaataatg tctacgtaaa 1740
tgtaactaca atcttttttt cttggtaatg tacacatatg ttctatatca tatccaaaat 1800
ctttgtcaga caattttgcc tttcatgc 1828

<210> 144
<211> 503
<212> PRT
<213> Arabidopsis thaliana

pctau2012001598-seql

<400> 144

```

Met Ser Pro Ala Lys Lys Ser Arg Ser Phe Pro Pro Ile Ser Glu Cys
 1      5      10      15

Lys Ser Arg Glu Tyr Asp Ser Ile Ala Ala Asp Leu Asp Gly Thr Leu
 20      25      30

Leu Leu Ser Arg Ser Ser Phe Pro Tyr Phe Met Leu Val Ala Ile Glu
 35      40      45

Ala Gly Ser Leu Phe Arg Gly Leu Ile Leu Leu Leu Ser Leu Pro Ile
 50      55      60

Val Ile Ile Ala Tyr Leu Phe Val Ser Glu Ser Leu Gly Ile Gln Ile
 65      70      75      80

Leu Ile Phe Ile Ser Phe Ala Gly Ile Lys Ile Lys Asn Ile Glu Leu
 85      90      95

Val Ser Arg Ala Val Leu Thr Arg Phe Tyr Ala Ala Asp Val Arg Lys
100      105      110

Asp Ser Phe Glu Val Phe Asp Lys Cys Lys Lys Arg Lys Val Val Val
115      120      125

Thr Ala Asn Pro Ile Val Met Val Glu Pro Phe Val Lys Asp Tyr Leu
130      135      140

Gly Gly Asp Lys Val Leu Gly Thr Glu Ile Glu Val Asn Pro Lys Thr
145      150      155      160

Met Lys Ala Thr Gly Phe Val Lys Lys Pro Gly Val Leu Val Gly Asp
165      170      175

Leu Lys Arg Leu Ala Ile Leu Lys Glu Phe Gly Asp Asp Ser Pro Asp
180      185      190

Leu Gly Leu Gly Asp Arg Thr Ser Asp His Asp Phe Met Ser Ile Cys
195      200      205

Lys Glu Gly Tyr Met Val His Glu Thr Lys Ser Ala Thr Thr Val Pro
210      215      220

Ile Glu Ser Leu Lys Asn Arg Ile Ile Phe His Asp Gly Arg Leu Val
225      230      235      240

Gln Arg Pro Thr Pro Leu Asn Ala Leu Ile Ile Tyr Leu Trp Leu Pro
245      250      255

Phe Gly Phe Met Leu Ser Val Phe Arg Val Tyr Phe Asn Leu Pro Leu

```

pctau2012001598-seql

260

265

270

Pro Glu Arg Phe Val Arg Tyr Thr Tyr Glu Ile Leu Gly Ile His Leu
 275 280 285

Thr Ile Arg Gly His Arg Pro Pro Pro Pro Ser Pro Gly Lys Pro Gly
 290 295 300

Asn Leu Tyr Val Leu Asn His Arg Thr Ala Leu Asp Pro Ile Ile Ile
 305 310 315 320

Ala Ile Ala Leu Gly Arg Lys Ile Thr Cys Val Thr Tyr Ser Val Ser
 325 330 335

Arg Leu Ser Leu Met Leu Ser Pro Ile Pro Ala Val Ala Leu Thr Arg
 340 345 350

Asp Arg Val Ala Asp Ala Ala Arg Met Arg Gln Leu Leu Glu Lys Gly
 355 360 365

Asp Leu Val Ile Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro Tyr Leu
 370 375 380

Leu Arg Phe Ser Ala Leu Phe Ala Glu Leu Ser Asp Arg Ile Val Pro
 385 390 395 400

Val Ala Met Asn Cys Lys Gln Gly Met Phe Asn Gly Thr Thr Val Arg
 405 410 415

Gly Val Lys Phe Trp Asp Pro Tyr Phe Phe Phe Met Asn Pro Arg Pro
 420 425 430

Ser Tyr Glu Ala Thr Phe Leu Asp Arg Leu Pro Glu Glu Met Thr Val
 435 440 445

Asn Gly Gly Gly Lys Thr Pro Phe Glu Val Ala Asn Tyr Val Gln Lys
 450 455 460

Val Ile Gly Gly Val Leu Gly Phe Glu Cys Thr Glu Leu Thr Arg Lys
 465 470 475 480

Asp Lys Tyr Leu Leu Leu Gly Gly Asn Asp Gly Lys Val Glu Ser Ile
 485 490 495

Asn Lys Thr Lys Ser Met Glu
 500

<210> 145

<211> 501

<212> PRT

<213> Arabidopsis thaliana

pctau2012001598-seql

<400> 145

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

pctau2012001598-seql

Pro Leu Pro Glu Arg Ile Ala Arg Tyr Asn Tyr Lys Leu Thr Gly Ile
275 280 285

Lys Leu Val Val Asn Gly His Pro Pro Pro Pro Pro Lys Pro Gly Gln
290 295 300

Pro Gly His Leu Leu Val Cys Asn His Arg Thr Val Leu Asp Pro Val
305 310 315 320

Val Thr Ala Val Ala Leu Gly Arg Lys Ile Ser Cys Val Thr Tyr Ser
325 330 335

Ile Ser Lys Phe Ser Glu Leu Ile Ser Pro Ile Lys Ala Val Ala Leu
340 345 350

Thr Arg Gln Arg Glu Lys Asp Ala Ala Asn Ile Lys Arg Leu Leu Glu
355 360 365

Glu Gly Asp Leu Val Ile Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro
370 375 380

Phe Leu Leu Arg Phe Ser Ala Leu Phe Ala Glu Leu Thr Asp Arg Ile
385 390 395 400

Val Pro Val Ala Ile Asn Thr Lys Gln Ser Met Phe Asn Gly Thr Thr
405 410 415

Thr Arg Gly Tyr Lys Leu Leu Asp Pro Tyr Phe Ala Phe Met Asn Pro
420 425 430

Arg Pro Thr Tyr Glu Ile Thr Phe Leu Lys Gln Ile Pro Ala Glu Leu
435 440 445

Thr Cys Lys Gly Gly Lys Ser Pro Ile Glu Val Ala Asn Tyr Ile Gln
450 455 460

Arg Val Leu Gly Gly Thr Leu Gly Phe Glu Cys Thr Asn Phe Thr Arg
465 470 475 480

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

Lys Lys Glu Lys Thr
500

<210> 146
<211> 289
<212> PRT
<213> Arabidopsis thaliana

<400> 146

pctau2012001598-seql

```

Met Val His Ala Thr Lys Ser Ala Thr Thr Ile Pro Lys Glu Arg Leu
1      5      10      15

Lys Asn Arg Ile Val Phe His Asp Gly Arg Leu Ala Gln Arg Pro Thr
20      25      30

Pro Leu Asn Ala Ile Ile Thr Tyr Leu Trp Leu Pro Phe Gly Phe Ile
35      40      45

Leu Ser Ile Ile Arg Val Tyr Phe Asn Leu Pro Leu Pro Glu Arg Phe
50      55      60

Val Arg Tyr Thr Tyr Glu Met Leu Gly Ile His Leu Thr Ile Arg Gly
65      70      75      80

His Arg Pro Pro Pro Pro Ser Pro Gly Thr Leu Gly Asn Leu Tyr Val
85      90      95

Leu Asn His Arg Thr Ala Leu Asp Pro Ile Ile Val Ala Ile Ala Leu
100     105     110

Gly Arg Lys Ile Cys Cys Val Thr Tyr Ser Val Ser Arg Leu Ser Leu
115     120     125

Met Leu Ser Pro Ile Pro Ala Val Ala Leu Thr Arg Asp Arg Ala Thr
130     135     140

Asp Ala Ala Asn Met Arg Lys Leu Leu Glu Lys Gly Asp Leu Val Ile
145     150     155     160

Cys Pro Glu Gly Thr Thr Cys Arg Glu Glu Tyr Leu Leu Arg Phe Ser
165     170     175

Ala Leu Phe Ala Glu Leu Ser Asp Arg Ile Val Pro Val Ala Met Asn
180     185     190

Cys Lys Gln Gly Met Phe Asn Gly Thr Thr Val Arg Gly Val Lys Phe
195     200     205

Trp Asp Pro Tyr Phe Phe Phe Met Asn Pro Arg Pro Ser Tyr Glu Ala
210     215     220

Thr Phe Leu Asp Arg Leu Pro Glu Glu Met Thr Val Asn Gly Gly Gly
225     230     235     240

Lys Thr Pro Ile Glu Val Ala Asn Tyr Val Gln Lys Val Ile Gly Ala
245     250     255

Val Leu Gly Phe Glu Cys Thr Glu Leu Thr Arg Lys Asp Lys Tyr Leu
260     265     270

```

pctau2012001598-seql

Leu Leu Gly Gly Asn Asp Gly Lys Val Glu Ser Ile Asn Asn Thr Lys
275 280 285

Lys

<210> 147
<211> 503
<212> PRT
<213> Arabidopsis thaliana
<400> 147

Met Ser Pro Ala Lys Lys Ser Arg Ser Phe Pro Pro Ile Ser Glu Cys
1 5 10 15

Lys Ser Arg Glu Tyr Asp Ser Ile Ala Ala Asp Leu Asp Gly Thr Leu
20 25 30

Leu Leu Ser Arg Ser Ser Phe Pro Tyr Phe Met Leu Val Ala Ile Glu
35 40 45

Ala Gly Ser Leu Phe Arg Gly Leu Ile Leu Leu Leu Ser Leu Pro Ile
50 55 60

Val Ile Ile Ala Tyr Leu Phe Val Ser Glu Ser Leu Gly Ile Gln Ile
65 70 75 80

Leu Ile Phe Ile Ser Phe Ala Gly Ile Lys Ile Lys Asn Ile Glu Leu
85 90 95

Val Ser Arg Ala Val Leu Thr Arg Phe Tyr Ala Ala Asp Val Arg Lys
100 105 110

Asp Ser Phe Glu Val Phe Asp Lys Cys Lys Lys Arg Lys Val Val Val
115 120 125

Thr Ala Asn Pro Ile Val Met Val Glu Pro Phe Val Lys Asp Tyr Leu
130 135 140

Gly Gly Asp Lys Val Leu Gly Thr Glu Ile Glu Val Asn Pro Lys Thr
145 150 155 160

Met Lys Ala Thr Gly Phe Val Lys Lys Pro Gly Val Leu Val Gly Asp
165 170 175

Leu Lys Arg Leu Ala Ile Leu Lys Glu Phe Gly Asp Asp Ser Pro Asp
180 185 190

Leu Gly Leu Gly Asp Arg Thr Ser Asp His Asp Phe Lys Ser Ile Cys
195 200 205

pctau2012001598-seql

Lys Glu Gly Tyr Met Val His Glu Thr Lys Ser Ala Thr Thr Val Pro
 210 215 220

Ile Glu Ser Leu Lys Asn Arg Ile Ile Phe His Asp Gly Arg Leu Val
 225 230 235 240

Gln Arg Pro Thr Pro Leu Asn Ala Leu Ile Ile Tyr Leu Trp Leu Pro
 245 250 255

Phe Gly Phe Met Leu Ser Val Phe Arg Val Tyr Phe Asn Leu Pro Leu
 260 265 270

Pro Glu Arg Phe Val Arg Tyr Thr Tyr Glu Ile Leu Gly Ile His Leu
 275 280 285

Thr Ile Arg Gly His Arg Pro Pro Pro Pro Ser Pro Gly Lys Pro Gly
 290 295 300

Asn Leu Tyr Val Leu Asn His Arg Thr Ala Leu Asp Pro Ile Ile Ile
 305 310 315 320

Ala Ile Ala Leu Gly Arg Lys Ile Thr Cys Val Thr Tyr Ser Val Ser
 325 330 335

Arg Leu Ser Leu Met Leu Ser Pro Ile Pro Ala Val Ala Leu Thr Arg
 340 345 350

Asp Arg Val Ala Asp Ala Ala Arg Met Arg Gln Leu Leu Glu Lys Gly
 355 360 365

Asp Leu Val Ile Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro Tyr Leu
 370 375 380

Leu Arg Phe Ser Ala Leu Phe Ala Glu Leu Ser Asp Arg Ile Val Pro
 385 390 395 400

Val Ala Met Asn Cys Lys Gln Gly Met Phe Asn Gly Thr Thr Val Arg
 405 410 415

Gly Val Lys Phe Trp Asp Pro Tyr Phe Phe Phe Met Asn Pro Arg Pro
 420 425 430

Ser Tyr Glu Ala Thr Phe Leu Asp Arg Leu Pro Glu Glu Met Thr Val
 435 440 445

Asn Gly Gly Gly Lys Thr Pro Phe Glu Val Ala Asn Tyr Val Gln Lys
 450 455 460

Val Ile Gly Gly Val Leu Gly Phe Glu Cys Thr Glu Leu Thr Arg Lys
 465 470 475 480

pctau2012001598-seql

Asp Lys Tyr Leu Leu Leu Gly Gly Asn Asp Gly Lys Val Glu Ser Ile
485 490 495

Asn Lys Thr Lys Ser Met Glu
500

<210> 148
<211> 467
<212> PRT
<213> Oryza sativa

<400> 148

Met Ala Phe Glu Thr Gly Gly Val Leu Arg Leu Leu Ala Leu Ile Leu
1 5 10 15

Leu Ala Pro Leu Ala Gly Val Leu Tyr Tyr Leu Val Ser Glu Ser Ala
20 25 30

Gly Ile Gln Val Leu Ile Phe Ala Ser Met Ala Gly Ala Arg Val Ala
35 40 45

Asp Val Glu Ala Val Ala Arg Ala Val Leu Pro Lys Phe Tyr Cys Ser
50 55 60

Asp Ile His Pro Glu Ser Trp Arg Val Phe Ser Ala Cys Gly Arg Arg
65 70 75 80

Cys Val Leu Thr Ala Asn Pro Arg Met Met Val Glu Ala Phe Leu Lys
85 90 95

Glu Tyr Leu Gly Ala Asp Ile Val Val Gly Thr Glu Leu Val Val Trp
100 105 110

Arg Gly Arg Ala Thr Gly Leu Val Arg Ser Pro Gly Val Leu Val Gly
115 120 125

Glu Asn Lys Ala Ala Ala Leu Arg Glu Ala Phe Gly Asp Ala Ser Pro
130 135 140

Glu Ile Gly Ile Gly Asp Arg Arg Thr Asp Tyr Pro Phe Met Arg Leu
145 150 155 160

Cys Lys Glu Gly Tyr Val Val Pro Val Ser Pro Thr Pro Arg Pro Val
165 170 175

Pro Arg Glu Glu Leu Pro Lys Pro Val Val Phe His Asp Gly Arg Leu
180 185 190

Val Gln Lys Pro Ser Pro Ala Leu Ala Leu Leu Thr Val Leu Trp Ile
195 200 205

Pro Ile Gly Phe Val Leu Ala Cys Leu Arg Ile Ala Ala Gly Ala Leu
Page 159

pctau2012001598-seql

210

215

220

Leu Pro Met Arg Val Val Tyr His Ala Phe Arg Ala Leu Gly Val Arg
 225 230 235 240

Val Thr Val Lys Gly Thr Pro Pro Pro Ala Ser Arg Glu Thr Gly
 245 250 255

Gln Ser Gly Val Leu Phe Ile Cys Ser His Arg Thr Leu Leu Asp Pro
 260 265 270

Ile Phe Leu Ser Thr Ala Leu Gly Arg Pro Ile Thr Ala Val Thr Tyr
 275 280 285

Ser Val Ser Arg Leu Ser Glu Ile Leu Ser Pro Ile Arg Thr Val Arg
 290 295 300

Leu Thr Arg Asp Arg Ala Ala Asp Ala Ala Met Ile Arg Arg Leu Leu
 305 310 315 320

Ala Glu Gly Asp Leu Val Ile Cys Pro Glu Gly Thr Thr Cys Arg Glu
 325 330 335

Pro Phe Leu Leu Arg Phe Ser Ala Leu Phe Ala Glu Leu Thr Asp Glu
 340 345 350

Ile Val Pro Val Ala Met Glu Asn Gln Met Ser Met Phe His Gly Thr
 355 360 365

Thr Ala Arg Gly Trp Lys Ala Leu Asp Pro Phe Tyr Phe Phe Met Asn
 370 375 380

Pro Ser Pro Gly Tyr Val Val Thr Phe Leu Ser Lys Leu Pro Gly Glu
 385 390 395 400

Leu Thr Cys Asn Gly Gly Gly Gly Gly Gly Gly Arg Ser Ser His
 405 410 415

Glu Val Ala Asn Tyr Ile Gln Arg Leu Ile Ala Ser Thr Leu Ser Tyr
 420 425 430

Glu Cys Thr Ser Phe Thr Arg Lys Asp Lys Tyr Arg Ala Leu Ala Gly
 435 440 445

Asn Asp Gly Thr Val Val Ser Lys Pro Asn Ile Asp Glu Lys Lys Ala
 450 455 460

Met Gly Cys
 465

<210> 149

pctau2012001598-seql

<211> 512

<212> PRT

<213> Picea sitchensis

<400> 149

Met Ala Lys Gly Gln Ser Gly Leu Asp Cys Gln Gly Leu Phe Asp Val
1 5 10 15

Ile Ser Lys Cys Ser Ser Lys Gly Arg Glu Asn Gln Thr Val Ala Ala
20 25 30

Asp Leu Asp Gly Thr Met Leu Arg Ser Arg Ser Ser Phe Pro Tyr Phe
35 40 45

Met Leu Val Ala Met Glu Gly Gly Ser Leu Leu Arg Gly Met Val Leu
50 55 60

Leu Ala Cys Ala Pro Met Ala Trp Phe Leu Tyr Asn Phe Ile Ser Glu
65 70 75 80

Ala Ala Gly Ile Gln Leu Leu Ile Phe Ile Ser Phe Ala Gly Leu Lys
85 90 95

Ile Arg Asp Ile Glu Leu Val Ser Arg Ala Val Leu Pro Lys Phe Tyr
100 105 110

Ala Glu Asp Val His Pro Glu Ser Trp Arg Val Phe Ser Ala Phe Gly
115 120 125

Lys Arg Tyr Ile Val Thr Ala Asn Pro Arg Ile Met Val Glu Pro Phe
130 135 140

Val Lys Gly Tyr Leu Gly Ala Asp Lys Val Leu Gly Thr Glu Leu Gln
145 150 155 160

Val Ser Arg Ser Gly Arg Ala Thr Gly Phe Val Lys Lys Pro Gly Val
165 170 175

Leu Val Gly Asp Leu Lys Lys Ala Ala Val Glu Ala Glu Phe Gly Asp
180 185 190

Lys Leu Pro Glu Leu Gly Leu Gly Asp Arg Glu Thr Asp His Pro Phe
195 200 205

Met Ser Leu Cys Lys Glu Gly Tyr Met Val Pro Lys Met Lys Val Asp
210 215 220

Glu Val Pro Thr Asn Lys Leu Met Ser Pro Ile Val Phe His Asp Gly
225 230 235 240

Arg Leu Val Gln Arg Pro Asn Pro Gly Ala Ala Leu Leu Thr Phe Leu
245 250 255

pctau2012001598-seql

Trp Leu Pro Ile Gly Phe Phe Leu Ala Gln Phe Arg Val Phe Gly Asn
260 265 270

Ile Pro Ile Pro Glu Lys Tyr Val Lys Ile Ser Tyr Lys Ile Met Gly
275 280 285

Ile Lys Leu Val Val Lys Gly Thr Pro Pro Pro Ala Pro Lys Lys Lys
290 295 300

Gly Glu Arg Gly Val Leu Phe Val Cys Asp His Arg Thr Leu Leu Asp
305 310 315 320

Pro Val Ile Val Ala Val Ala Leu Gly Arg Lys Val Ser Ala Val Thr
325 330 335

Tyr Ser Ile Ser Arg Phe Ser Glu Ile Ile Ser Pro Ile Lys Thr Val
340 345 350

Arg Leu Thr Arg Asp Arg Glu Arg Asp Ala Ala Asn Ile Lys Arg Leu
355 360 365

Leu Glu Glu Gly Asp Leu Val Ile Cys Pro Glu Gly Thr Thr Cys Arg
370 375 380

Glu Pro Phe Leu Leu Arg Phe Ser Ala Leu Phe Ala Glu Leu Thr Asp
385 390 395 400

Arg Ile Val Pro Val Ala Ile Cys Asn Lys Met Ser Met Phe His Gly
405 410 415

Thr Thr Val Arg Gly Trp Lys Gly Phe Asp Pro Phe Phe Phe Phe Met
420 425 430

Asn Pro Val Pro Thr Tyr Glu Val Thr Phe Leu Asp Gln Leu Pro Thr
435 440 445

Glu Leu Thr Cys Ala Gly Gly Lys Ser Pro Ile Glu Val Ala Asn Tyr
450 455 460

Ile Gln Arg Val Leu Ala Ala Thr Leu Gly Tyr Glu Cys Thr Asn Phe
465 470 475 480

Thr Arg Lys Asp Lys Tyr Arg Met Leu Ala Gly Asn Asp Gly Ile Val
485 490 495

Ser Leu Asn Lys Thr Ala Thr Ala Arg Ser Asn Asn His Lys Gln Asn
500 505 510

<210> 150

<211> 502

pctau2012001598-seql

<212> PRT
 <213> Zea mays

<400> 150

```

Met Val Ala Ser Pro Arg Phe Lys Pro Ile Glu Glu Cys Cys Ser Glu
1          5          10          15

Gly Arg Ser Glu Gln Thr Val Ala Ala Asp Leu Asp Gly Thr Leu Leu
          20          25          30

Ile Ser Arg Ser Ala Phe Pro Tyr Tyr Leu Leu Val Ala Leu Glu Ala
          35          40          45

Gly Ser Val Leu Arg Ala Ala Leu Leu Leu Leu Ser Val Pro Phe Val
          50          55          60

Tyr Val Thr Tyr Ala Phe Phe Ser Glu Ser Leu Ala Ile Ser Thr Leu
65          70          75          80

Val Tyr Ile Ser Val Ala Gly Leu Lys Val Arg Ser Ile Glu Met Val
          85          90          95

Ala Arg Ser Val Leu Pro Arg Phe Tyr Ala Gly Asp Val His Pro Glu
          100          105          110

Ser Trp Arg Val Phe Ser Ser Phe Gly Arg Arg Tyr Val Val Thr Ala
          115          120          125

Ser Pro Arg Val Met Val Glu Pro Phe Ala Arg Ala Phe Leu Gly Ala
          130          135          140

Asp Lys Val Val Gly Thr Glu Leu Glu Val Gly Arg Asp Gly Lys Ala
145          150          155          160

Thr Gly Phe Met Ala Arg Pro Gly Val Leu Val Gly Asp His Lys Lys
          165          170          175

Lys Ala Val Val Lys Glu Leu Gly Asp Ala Leu Pro Asp Val Gly Met
          180          185          190

Gly Asp Arg Glu Thr Asp Phe Asp Phe Met Ser Ile Cys Lys Glu Ala
          195          200          205

Tyr Leu Val Thr Ser Arg Lys Tyr Ser Pro Val Pro Arg Asn Gln Leu
210          215          220

Leu Ser Pro Leu Ile Val His Asp Gly Arg Leu Val Gln Arg Pro Thr
225          230          235          240

Pro Leu Val Ala Leu Val Thr Phe Leu Trp Met Pro Phe Gly Phe Ala
          245          250          255

```

pctau2012001598-seql

Leu Ala Leu Met Arg Val Tyr Ile Asn Leu Pro Leu Pro Glu Arg Ile
 260 265 270
 Val Tyr Tyr Thr Tyr Lys Leu Met Gly Ile Arg Leu Val Val Lys Gly
 275 280 285
 Thr Pro Pro Pro Pro Pro Lys Lys Gly His Pro Gly Val Leu Phe Val
 290 295 300
 Cys Asn His Arg Thr Val Leu Asp Pro Val Glu Val Ala Val Ala Leu
 305 310 315 320
 Arg Arg Lys Val Ser Cys Val Thr Tyr Ser Ile Ser Lys Phe Ser Glu
 325 330 335
 Leu Ile Ser Pro Ile Lys Ala Val Ala Leu Ser Arg Glu Arg Asp Lys
 340 345 350
 Asp Ala Glu Asn Ile Arg Arg Leu Leu Glu Glu Gly Asp Leu Val Ile
 355 360 365
 Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro Phe Leu Leu Arg Phe Ser
 370 375 380
 Ala Leu Phe Ala Glu Leu Thr Asp Arg Ile Val Pro Val Ala Ile Asn
 385 390 395 400
 Thr Lys Glu Thr Met Phe His Gly Ser Thr Val Arg Gly Phe Lys Leu
 405 410 415
 Met Asp Pro Tyr Phe Phe Phe Met Asn Pro Arg Pro Thr Tyr Glu Ile
 420 425 430
 Thr Phe Leu Thr Gln Leu Pro Lys Asp Leu Thr Cys Ser Gly Gly Lys
 435 440 445
 Ser Pro Ile Glu Val Ala Asn Tyr Ile Gln Lys Thr Leu Ser Gly Gln
 450 455 460
 Leu Gly Phe Glu Cys Thr Ser Ile Thr Arg Lys Glu Lys Tyr Gly Met
 465 470 475 480
 Leu Ala Gly Thr Asp Gly Arg Val Pro Ser Lys Asn Lys Glu Lys Glu
 485 490 495
 Lys Glu Lys Asp Lys Asn
 500

<210> 151
 <211> 500
 <212> PRT

<213> Arabidopsis thaliana

<400> 151

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

pctau2012001598-seql

Gly Phe Ile Leu Ser Ile Ile Arg Val Tyr Phe Asn Leu Pro Leu Pro
260 265 270

Glu Arg Phe Val Arg Tyr Thr Tyr Glu Met Leu Gly Ile His Leu Thr
275 280 285

Ile Arg Gly His Arg Pro Pro Pro Pro Ser Pro Gly Thr Leu Gly Asn
290 295 300

Leu Tyr Val Leu Asn His Arg Thr Ala Leu Asp Pro Ile Ile Val Ala
305 310 315 320

Ile Ala Leu Gly Arg Lys Ile Cys Cys Val Thr Tyr Ser Val Ser Arg
325 330 335

Leu Ser Leu Met Leu Ser Pro Ile Pro Ala Val Ala Leu Thr Arg Asp
340 345 350

Arg Ala Thr Asp Ala Ala Asn Met Arg Lys Leu Leu Glu Lys Gly Asp
355 360 365

Leu Val Ile Cys Pro Glu Gly Thr Thr Cys Arg Glu Glu Tyr Leu Leu
370 375 380

Arg Ser Ser Ala Leu Phe Ala Glu Leu Ser Asp Arg Ile Val Pro Val
385 390 395 400

Ala Met Asn Cys Lys Gln Gly Met Phe Asn Gly Thr Thr Val Arg Gly
405 410 415

Val Lys Phe Trp Asp Pro Tyr Phe Phe Phe Met Asn Pro Arg Pro Ser
420 425 430

Tyr Glu Ala Thr Phe Leu Asp Arg Leu Pro Glu Glu Met Thr Val Asn
435 440 445

Gly Gly Gly Lys Thr Pro Ile Glu Val Ala Asn Tyr Val Gln Lys Val
450 455 460

Ile Gly Ala Val Leu Gly Phe Glu Cys Thr Glu Leu Thr Arg Lys Asp
465 470 475 480

Lys Tyr Leu Leu Leu Gly Gly Asn Asp Gly Lys Val Glu Ser Ile Asn
485 490 495

Asn Thr Lys Lys
500

<210> 152
<211> 486
<212> PRT
<213> Oryza sativa

pctau2012001598-seql

<400> 152

Met Ser Pro Phe Lys Pro Ile Glu Gln Cys Ser Thr Glu Gly Arg Ser
1 5 10 15

Gln Gln Thr Val Ala Ser Asp Leu Asp Gly Thr Leu Leu Leu Ser Arg
20 25 30

Ser Ala Phe Pro Tyr Tyr Leu Leu Val Ala Leu Glu Ala Gly Gly Pro
35 40 45

Leu Arg Ala Val Ala Leu Leu Met Ser Val Pro Phe Val Tyr Leu Thr
50 55 60

Tyr Val Thr Ile Ser Glu Ser Leu Ala Val Arg Ala Leu Leu Tyr Ile
65 70 75 80

Ala Val Ala Gly Leu Glu Val Arg Asp Val Glu Ser Val Ala Arg Ser
85 90 95

Val Leu Pro Arg Phe Tyr Ala Gly Asp Val His Pro Glu Gly Trp Arg
100 105 110

Val Phe Ser Ser Phe Gly Arg Arg Cys Val Val Thr Ala Ser Pro Arg
115 120 125

Val Met Val Glu Pro Phe Ala Arg Ala Phe Leu Gly Ala Asp Arg Val
130 135 140

Ile Gly Thr Glu Leu Glu Val Gly Glu Asp Gly Arg Ala Thr Gly Phe
145 150 155 160

Val Ala Lys Pro Gly Val Leu Ile Arg Glu His Lys Arg Asn Ala Val
165 170 175

Val Arg Glu Phe Gly Asp Ala Leu Pro Asp Val Gly Met Gly Asp Arg
180 185 190

Glu Ser Asp Phe Asp Phe Met Ala Ile Cys Lys Asp Ala Tyr Val Val
195 200 205

Thr Thr Ser Arg Lys His Arg Pro Val Pro Glu Ser Gln Leu Leu Arg
210 215 220

Thr Val Val Leu His Asp Gly Arg Leu Ala Gln Arg Pro Thr Ala Ile
225 230 235 240

Asn Thr Leu Leu Val Phe Leu Trp Met Pro Val Gly Phe Ala Leu Ala
245 250 255

Leu Leu Arg Ala Cys Leu Ser Leu Leu Leu Pro Glu Arg Val Leu Ser
Page 167

pctau2012001598-seql

260

265

270

Tyr Ala Tyr Lys Leu Thr Gly Val Gly Leu Val Val Arg Gly Arg Pro
 275 280 285

Pro Pro Pro Asp Gly Ser Pro Gly Val Leu Phe Val Cys Asn His Arg
 290 295 300

Thr Val Leu Asp Gln Val Ala Val Ala Ala Ala Leu Gly Arg Lys Val
 305 310 315 320

Ile Cys Val Thr Tyr Ser Val Pro Arg Lys Thr Tyr Gly Met Ser Ser
 325 330 335

Arg Leu Pro Glu Ala Leu Thr Ala Ser Pro Val Lys Ala Ala Val Ala
 340 345 350

Leu Cys Arg Glu Arg Asp Arg Asp Ala Asp Arg Val Arg Arg Leu Leu
 355 360 365

Glu Glu Gly Val Asp Ile Val Ala Phe Pro Glu Gly Thr Thr Cys Arg
 370 375 380

Gly Ala Phe Leu Leu Arg Phe Ser Ser Leu Phe Ala Glu Leu Thr Asp
 385 390 395 400

Arg Ile Val Pro Val Ala Ile Ala Thr Arg Glu Thr Met Phe His Gly
 405 410 415

Ser Thr Ala Arg Gly Phe Lys Gly Met Asp Pro Tyr Phe Phe Phe Met
 420 425 430

Asn Pro Arg Pro Ala Tyr Glu Val Thr Phe Leu Ser Gln Leu Pro Ser
 435 440 445

Glu Leu Thr Ser Gly Gly Gly Gly Lys Ser Pro Val Glu Val Ala Asn
 450 455 460

Tyr Val Gln Lys Ala Leu Ala Gly Gln Leu Gly Ser Glu His Ile Gly
 465 470 475 480

Ile Thr Arg Lys Glu Lys
 485

<210> 153
 <211> 506
 <212> PRT
 <213> Oryza sativa

<400> 153

Met Ala Ala Ala Ala Ala Ala Gly Arg Arg Arg Phe Phe Pro Pro Val
 1 5 10 15

pctau2012001598-seql

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

pctau2012001598-seql

Ile Arg Leu Ala Val Arg Gly Ala Pro Pro Pro Pro Pro Arg Pro Gly
290 295 300

Thr Pro Gly Ser Leu Leu Val Cys Asn His Arg Thr Ala Leu Asp Pro
305 310 315 320

Ile Ile Val Ser Ile Ala Leu Gly Arg Pro Val Thr Cys Val Thr Tyr
325 330 335

Ser Val Ser Arg Leu Ser Thr Ala Ile Ser Pro Ile Arg Ala Ala Ala
340 345 350

Leu Thr Pro Asp Arg Ala Ala Asp Ala Ala Arg Ile Ala Ala Leu Leu
355 360 365

Glu Glu Gly Asp Val Val Val Cys Pro Glu Gly Thr Thr Cys Arg Glu
370 375 380

Pro Tyr Leu Leu Arg Phe Ser Ala Leu Phe Ala Glu Leu Thr Ala Arg
385 390 395 400

Ile Val Pro Val Ala Val Glu Ala Arg Gln Gly Thr Tyr Tyr Gly Ser
405 410 415

Thr Ala Arg Gly Trp Lys Phe Leu Asp Pro Tyr Phe Phe Tyr Met Asn
420 425 430

Pro Arg Pro Gly Tyr Glu Val Thr Phe Leu Pro Ala Leu Arg Pro Glu
435 440 445

Glu Thr Cys Val Ala Gly Gly Arg Ser Ala Val Glu Val Ala Asn His
450 455 460

Val Gln Arg Val Ile Ala Lys Glu Leu Gly Phe Gln Cys Thr Thr Leu
465 470 475 480

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

Ala Ala Ala Ala Asp Lys Pro Lys Ala Asn
500 505

<210> 154
<211> 486
<212> PRT
<213> Oryza sativa

<400> 154

Met Ser Pro Phe Lys Pro Ile Lys Gln Cys Ser Thr Glu Gly Arg Ser
1 5 10 15

pctau2012001598-seql

Gln Gln Thr Val Ala Ser Asp Leu Asp Gly Thr Leu Leu Leu Ser Arg
20 25 30

Ser Ala Phe Pro Tyr Tyr Leu Leu Val Ala Leu Glu Ala Gly Gly Pro
35 40 45

Leu Arg Ala Val Ala Leu Leu Met Ser Val Pro Phe Val Tyr Leu Thr
50 55 60

Tyr Val Thr Ile Ser Glu Ser Leu Ala Val Arg Ala Leu Leu Tyr Ile
65 70 75 80

Ala Val Ala Gly Leu Glu Val Arg Asp Val Glu Ser Val Ala Arg Ser
85 90 95

Val Leu Pro Arg Phe Tyr Ala Gly Asp Val His Pro Glu Gly Trp Arg
100 105 110

Val Phe Ser Ser Phe Gly Arg Arg Cys Val Val Thr Ala Ser Pro Arg
115 120 125

Val Met Val Glu Pro Phe Ala Arg Ala Phe Leu Gly Ala Asp Arg Val
130 135 140

Ile Gly Thr Glu Leu Glu Val Gly Glu Asp Gly Arg Ala Thr Gly Phe
145 150 155 160

Val Ala Lys Pro Gly Val Leu Ile Arg Glu His Lys Arg Asn Ala Val
165 170 175

Val Arg Glu Phe Gly Asp Ala Leu Pro Asp Val Gly Met Gly Asp Arg
180 185 190

Glu Ser Asp Phe Asp Phe Met Ala Ile Cys Lys Asp Ala Tyr Val Val
195 200 205

Thr Thr Ser Arg Lys His Arg Pro Val Pro Glu Ser Gln Leu Leu Arg
210 215 220

Thr Val Val Leu His Asp Gly Arg Leu Ala Gln Arg Pro Thr Ala Ile
225 230 235 240

Asn Thr Leu Leu Val Phe Leu Trp Met Pro Val Gly Phe Ala Leu Ala
245 250 255

Leu Leu Arg Ala Cys Leu Ser Leu Leu Leu Pro Glu Arg Val Leu Ser
260 265 270

Tyr Ala Tyr Lys Leu Thr Gly Val Gly Leu Val Val Arg Gly Arg Pro
275 280 285

pctau2012001598-seql

Pro Pro Pro Asp Gly Ser Pro Gly Val Leu Phe Val Cys Asn His Arg
290 295 300

Thr Val Leu Asp Pro Val Ala Val Ala Ala Ala Leu Gly Arg Lys Val
305 310 315 320

Ile Cys Val Thr Tyr Ser Val Pro Arg Lys Thr Tyr Gly Met Ser Ser
325 330 335

Arg Leu Pro Glu Ala Leu Thr Ala Ser Pro Val Lys Ala Ala Val Ala
340 345 350

Leu Cys Arg Glu Arg Asp Arg Asp Ala Asp Arg Val Arg Arg Leu Leu
355 360 365

Glu Glu Gly Val Asp Ile Val Ala Phe Pro Glu Gly Thr Thr Cys Arg
370 375 380

Gly Ala Phe Leu Leu Arg Phe Ser Ser Leu Phe Ala Glu Leu Thr Asp
385 390 395 400

Arg Ile Val Pro Val Ala Ile Ala Thr Arg Glu Thr Met Phe His Gly
405 410 415

Ser Thr Ala Arg Gly Phe Lys Gly Met Asp Pro Tyr Phe Phe Phe Met
420 425 430

Asn Pro Arg Pro Ala Tyr Glu Val Thr Phe Leu Ser Gln Leu Pro Ser
435 440 445

Glu Leu Thr Ser Gly Gly Gly Gly Lys Ser Pro Val Glu Val Ala Asn
450 455 460

Tyr Val Gln Lys Ala Leu Ala Gly Gln Leu Gly Ser Glu His Ile Gly
465 470 475 480

Ile Thr Arg Lys Glu Lys
485

<210> 155
<211> 507
<212> PRT
<213> Oryza sativa

<400> 155

Met Ala Ala Ala Ala Ala Ala Gly Arg Arg Arg Phe Phe Pro Pro Val
1 5 10 15

Thr Ala Tyr Asp Ala Ala Ala Gly Ala Arg Arg Thr Val Ala Ala Asp
20 25 30

pctau2012001598-seql

Leu Asp Gly Thr Leu Leu Val Ser Ser Ser Ala Phe Pro Tyr Tyr Phe
 35 40 45
 Leu Val Ala Leu Glu Ala Gly Ser Tyr Leu Arg Ala Leu Ala Leu Leu
 50 55 60
 Leu Ala Ala Pro Trp Leu Leu Ala Leu Tyr Val Gly Val Ser Glu Ala
 65 70 75 80
 Ala Ala Ile Ala Leu Leu Val Phe Ile Thr Phe Ala Gly Leu Arg Val
 85 90 95
 Arg Asp Val Glu Ala Val Ala Arg Ala Val Leu Pro Arg His Tyr Ala
 100 105 110
 Ala Gly Val Arg Ala Glu Thr Trp Ala Val Phe His Gly Trp Ala Lys
 115 120 125
 Arg Arg Val Leu Val Thr Ala Leu Pro Ala Val Met Val Gly Glu Phe
 130 135 140
 Val Arg Glu Phe Leu Gly Ala Glu Val Ala Gly Thr Glu Leu Glu Thr
 145 150 155 160
 Phe Ala Ser Gly Lys Arg Phe Thr Gly Arg Ile Lys Ala Val Leu Val
 165 170 175
 Gly Glu Lys Lys Arg Glu Val Val Glu Arg Leu Phe Ala Gly Gly Asp
 180 185 190
 Met Pro Asp Val Gly Leu Gly Asp Arg Glu Ser Asp His Asp Phe Met
 195 200 205
 Ala Ile Cys Lys Glu Ala Tyr Met Val Pro Lys Asn Lys Arg Ala Pro
 210 215 220
 Arg Ala Ala Ala Asp Glu Leu Leu Ser Arg Ala Ile Phe His Asp Gly
 225 230 235 240
 Arg Leu Val Arg Arg Pro Glu Pro Ala Ser Ala Leu Phe Ala Leu Ala
 245 250 255
 Tyr Leu Pro Val Gly Phe Ala Val Ala Leu Leu Arg Val Phe Leu Asn
 260 265 270
 Leu Pro Val Pro Ala Arg Leu Val Arg His Thr Tyr Arg Leu Thr Gly
 275 280 285
 Ile Arg Leu Ala Val Arg Gly Ala Arg Arg Arg Pro Leu Arg Pro Gly
 290 295 300

pctau2012001598-seql

Asp Ala Gly Ala Pro Ser Cys Cys Ala Thr Thr Glu Pro Pro Arg Thr
305 310 315 320

Pro Ser Ser Cys Pro Ser Ala Leu Ala Ala Pro Leu Arg Ala Leu Arg
325 330 335

Thr Ala Leu Ala Pro Leu Ser Thr Ala Ile Ser Pro Ile Arg Ala Ala
340 345 350

Ala Leu Thr Arg Asp Arg Ala Ala Asp Ala Ala Arg Ile Ala Ala Leu
355 360 365

Leu Glu Glu Gly Asp Val Val Val Cys Pro Glu Gly Thr Thr Cys Arg
370 375 380

Glu Pro Tyr Leu Leu Arg Phe Ser Ala Leu Phe Ala Glu Leu Thr Ala
385 390 395 400

Arg Ile Val Pro Val Ala Val Glu Ala Arg Gln Gly Thr Tyr Tyr Gly
405 410 415

Ser Thr Ala Arg Gly Trp Lys Phe Leu Asp Pro Tyr Phe Phe Tyr Met
420 425 430

Asn Pro Arg Pro Gly Tyr Glu Val Thr Phe Leu Pro Ala Leu Arg Pro
435 440 445

Glu Glu Thr Cys Val Ala Gly Gly Arg Ser Ala Val Glu Val Ala Asn
450 455 460

His Val Gln Arg Val Ile Ala Lys Glu Leu Gly Phe Gln Cys Thr Thr
465 470 475 480

Leu Thr Arg Lys Asp Lys Tyr Met Lys Leu Ala Gly Asn Asp Gly Arg
485 490 495

Val Ala Ala Ala Ala Asp Lys Pro Lys Ala Asn
500 505

<210> 156
<211> 497
<212> PRT
<213> Oryza sativa

<400> 156

Met Val Ser Arg Arg Phe Lys Pro Val Glu Glu Cys Ser Ser Asp Gly
1 5 10 15

Arg Ser Glu Gln Thr Val Ala Ala Asp Phe Asp Gly Thr Leu Val Arg
20 25 30

Ser Arg Ser Ala Phe Pro Tyr Tyr Leu Leu Val Ala Leu Glu Ala Gly
Page 174

35

40

45

Ser Val Leu Arg Ala Val Val Leu Leu Leu Ser Val Pro Phe Val Tyr
 50 55 60
 Val Thr Tyr Ile Phe Phe Ser Glu Ser Leu Ala Ile Ser Thr Leu Val
 65 70 75 80
 Tyr Ile Ser Val Ala Gly Leu Lys Val Arg Asn Ile Glu Met Val Ala
 85 90 95
 Arg Ser Val Leu Pro Lys Phe Tyr Ala Glu Asp Val His Pro Glu Ser
 100 105 110
 Trp Arg Val Phe Asn Ser Phe Gly Lys Arg Tyr Ile Ile Thr Ala Ser
 115 120 125
 Pro Arg Ile Met Val Glu His Phe Ala Lys Thr Phe Leu Gly Ala Asp
 130 135 140
 Lys Val Val Gly Thr Glu Leu Glu Val Gly Lys Asn Gly Lys Ala Thr
 145 150 155 160
 Gly Phe Met Val Lys Pro Gly Val Leu Val Gly Asp His Lys Arg Gln
 165 170 175
 Ala Val Val Lys Glu Leu Arg Asp Ala Val Pro Asp Val Gly Leu Gly
 180 185 190
 Asp Arg Glu Thr Asp Phe Asp Phe Met Ser Ile Cys Lys Glu Ala Tyr
 195 200 205
 Leu Val Thr Ser Arg Lys Tyr Ser Ala Val Pro Lys Asn Gln Leu Leu
 210 215 220
 Ser Pro Leu Ile Leu His Asp Gly Arg Leu Val Gln Arg Pro Thr Pro
 225 230 235 240
 Leu Val Ala Leu Val Thr Phe Leu Trp Met Pro Phe Gly Phe Val Leu
 245 250 255
 Ala Leu Leu Arg Val Tyr Val Asn Leu Pro Leu Pro Glu Arg Ile Val
 260 265 270
 Phe Tyr Thr Tyr Lys Leu Met Gly Ile Arg Leu Ile Val Lys Gly Asn
 275 280 285
 Pro Pro Pro Pro Pro Lys Lys Gly His Pro Gly Val Leu Phe Val Cys
 290 295 300
 Asn His Arg Thr Val Leu Asp Pro Val Glu Val Ala Val Ala Leu Arg
 Page 175

pctau2012001598-seql

305 310 315 320

Arg Lys Val Ser Cys Val Thr Tyr Ser Ile Ser Lys Phe Ser Glu Leu
 325 330 335

Ile Ser Pro Ile Lys Ala Val Ala Leu Ser Arg Glu Arg Glu Lys Asp
 340 345 350

Ala Glu Asn Ile Arg Arg Leu Leu Glu Glu Gly Asp Leu Val Ile Cys
 355 360 365

Pro Glu Gly Thr Thr Cys Arg Glu Pro Phe Leu Leu Arg Phe Ser Ala
 370 375 380

Leu Phe Ala Glu Leu Thr Asp Arg Ile Val Pro Val Ala Ile Asn Thr
385 390 395 400

Lys Glu Ser Met Phe His Gly Ser Thr Val Arg Gly Phe Lys Leu Met
 405 410 415

Asp Pro Tyr Phe Phe Phe Met Asn Pro Arg Pro Thr Tyr Glu Ile Thr
 420 425 430

Phe Leu Asn Gln Leu Pro Lys Glu Leu Thr Cys Ser Gly Gly Lys Ser
 435 440 445

Pro Ile Glu Val Ala Asn Tyr Ile Gln Lys Thr Leu Ser Gly Gln Leu
 450 455 460

Gly Phe Glu Cys Thr Ala Ile Thr Arg Lys Glu Lys Tyr Ser Ile Leu
465 470 475 480

Ala Gly Thr Asp Gly Arg Val Pro Ser Lys Asn Lys Glu Lys Glu Lys
 485 490 495

Asn

<210> 157
<211> 505
<212> PRT
<213> Oryza sativa

<400> 157

Met Ile Ala Ser Pro Phe Pro Ala Val Glu Lys Cys Ser Ser Lys Asp
1 5 10 15

Arg Gly Gly Asp Thr Ile Val Ala Asp Leu Asp Gly Thr Leu Leu Cys
 20 25 30

Gly Arg Ser Ser Phe Pro Tyr Phe Ala His Met Ala Phe Glu Thr Gly
 35 40 45

pctau2012001598-seql

Gly Val Leu Arg Leu Leu Ala Leu Ile Leu Leu Ala Pro Leu Ala Gly
 50 55 60
 Val Leu Tyr Tyr Leu Val Ser Glu Ser Ala Gly Ile Gln Val Leu Ile
 65 70 75 80
 Phe Ala Ser Met Ala Gly Ala Arg Val Ala Asp Val Glu Ala Val Ala
 85 90 95
 Arg Ala Val Leu Pro Lys Phe Tyr Cys Ser Asp Ile His Pro Glu Ser
 100 105 110
 Trp Arg Val Phe Ser Ala Cys Gly Arg Arg Cys Val Leu Thr Ala Asn
 115 120 125
 Pro Arg Met Met Val Glu Ala Phe Leu Lys Glu Tyr Leu Gly Ala Asp
 130 135 140
 Ile Val Val Gly Thr Glu Leu Val Val Trp Arg Gly Arg Ala Thr Gly
 145 150 155 160
 Leu Val Arg Ser Pro Gly Val Leu Val Gly Glu Asn Lys Ala Ala Ala
 165 170 175
 Leu Arg Glu Ala Phe Gly Asp Ala Ser Pro Glu Ile Gly Ile Gly Asp
 180 185 190
 Arg Arg Thr Asp Tyr Pro Phe Met Arg Leu Cys Lys Glu Gly Tyr Val
 195 200 205
 Val Pro Val Ser Pro Thr Pro Arg Pro Val Pro Arg Glu Glu Leu Pro
 210 215 220
 Lys Pro Val Val Phe His Asp Gly Arg Leu Val Gln Lys Pro Ser Pro
 225 230 235 240
 Ala Leu Ala Leu Leu Thr Val Leu Trp Ile Pro Ile Gly Phe Val Leu
 245 250 255
 Ala Cys Leu Arg Ile Ala Ala Gly Ala Leu Leu Pro Met Arg Val Val
 260 265 270
 Tyr His Ala Phe Arg Ala Leu Gly Val Arg Val Thr Val Lys Gly Thr
 275 280 285
 Pro Pro Pro Pro Ala Ser Arg Glu Thr Gly Gln Ser Gly Val Leu Phe
 290 295 300
 Ile Cys Ser His Arg Thr Leu Leu Asp Pro Ile Phe Leu Ser Thr Ala
 305 310 315 320

pctau2012001598-seql

Leu Gly Arg Pro Ile Thr Ala Val Thr Tyr Ser Val Ser Arg Leu Ser
325 330 335

Glu Ile Leu Ser Pro Ile Arg Thr Val Arg Leu Thr Arg Asp Arg Ala
340 345 350

Ala Asp Ala Ala Met Ile Arg Arg Leu Leu Ala Glu Gly Asp Leu Val
355 360 365

Ile Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro Phe Leu Leu Arg Phe
370 375 380

Ser Ala Leu Phe Ala Glu Leu Thr Asp Glu Ile Val Pro Val Ala Met
385 390 395 400

Glu Asn Gln Met Ser Met Phe His Gly Thr Thr Ala Arg Gly Trp Lys
405 410 415

Ala Leu Asp Pro Phe Tyr Phe Phe Met Asn Pro Ser Pro Gly Tyr Val
420 425 430

Val Thr Phe Leu Ser Lys Leu Pro Gly Glu Leu Thr Cys Asn Gly Gly
435 440 445

Gly Gly Arg Ser Ser His Glu Val Ala Asn Tyr Ile Gln Arg Leu Ile
450 455 460

Ala Ser Thr Leu Ser Tyr Glu Cys Thr Ser Phe Thr Arg Lys Asp Lys
465 470 475 480

Tyr Arg Ala Leu Ala Gly Asn Asp Gly Thr Val Val Ser Lys Pro Asn
485 490 495

Ile Asp Glu Lys Lys Ala Met Gly Cys
500 505

<210> 158
<211> 474
<212> PRT
<213> Oryza sativa

<400> 158

Met Gly Asp Pro Pro Met Ile Ala Ser Pro Phe Pro Ala Val Glu Lys
1 5 10 15

Phe Ser Ser Lys Asp Arg Gly Gly Asp Thr Ile Val Pro Asp Leu Asp
20 25 30

Gly Thr Leu Leu Cys Gly Arg Ser Ser Phe Pro Tyr Phe Gly His Met
35 40 45

pctau2012001598-seql

Ala Phe Lys Thr Gly Gly Val Leu Arg Leu Leu Ala Leu Ile Leu Leu
50 55 60

Ala Pro Phe Ala Gly Leu Leu Tyr Tyr Leu Val Ser Asp Ser Ala Gly
65 70 75 80

Ile Gln Val Leu Ile Phe Ala Ser Met Ala Gly Pro Arg Val Pro Asp
85 90 95

Phe Lys Ala Ala Phe Leu Lys Glu Tyr Leu Gly Ala Asp Ile Val Val
100 105 110

Gly Thr Glu Leu Val Val Trp Arg Gly Arg Ala Thr Gly Leu Val Arg
115 120 125

Ser Pro Gly Val Leu Val Gly Glu Asn Lys Ala Ala Ala Leu Arg Glu
130 135 140

Ala Phe Gly Asp Ala Ser Pro Glu Ile Gly Ile Gly Asp Arg Arg Thr
145 150 155 160

Asp Tyr Pro Phe Met Arg Leu Cys Lys Glu Gly Tyr Val Val Pro Val
165 170 175

Ser Pro Thr Pro Arg Pro Val Pro Arg Glu Glu Leu Pro Lys Pro Val
180 185 190

Val Phe His Asp Gly Arg Leu Val Gln Lys Pro Ser Pro Ala Leu Ala
195 200 205

Leu Leu Thr Val Leu Trp Ile Pro Ile Gly Phe Val Leu Ala Cys Leu
210 215 220

Arg Ile Ala Ala Gly Ala Leu Leu Pro Met Arg Val Val Tyr His Ala
225 230 235 240

Phe Arg Ala Leu Gly Val Arg Val Thr Val Lys Gly Thr Pro Pro Pro
245 250 255

Pro Ala Ser Arg Glu Thr Gly Gln Ser Gly Val Leu Phe Ile Cys Ser
260 265 270

His Arg Thr Leu Leu Asp Pro Ile Phe Leu Ser Thr Ala Leu Gly Arg
275 280 285

Pro Ile Thr Ala Val Thr Tyr Ser Val Ser Arg Leu Ser Glu Ile Leu
290 295 300

Ser Pro Ile Arg Thr Val Arg Leu Thr Arg Asp Arg Ala Ala Asp Ala
305 310 315 320

pctau2012001598-seql

Ala Met Ile Arg Arg Leu Leu Ala Glu Gly Asp Leu Val Ile Cys Pro
325 330 335

Glu Gly Thr Thr Cys Arg Glu Pro Phe Leu Leu Arg Phe Ser Ala Leu
340 345 350

Phe Ala Glu Leu Thr Asp Glu Ile Val Pro Val Ala Met Glu Asn Gln
355 360 365

Met Ser Met Phe His Gly Thr Thr Ala Arg Gly Trp Lys Ala Leu Asp
370 375 380

Pro Phe Tyr Phe Phe Met Asn Pro Ser Pro Gly Tyr Val Val Thr Phe
385 390 395 400

Leu Ser Lys Leu Pro Gly Glu Leu Thr Cys Asn Gly Gly Gly Gly Gly
405 410 415

Gly Gly Gly Arg Ser Ser His Glu Val Ala Asn Tyr Ile Gln Arg Leu
420 425 430

Ile Ala Ser Thr Leu Ser Tyr Glu Cys Thr Ser Phe Thr Arg Lys Asp
435 440 445

Lys Tyr Arg Ala Leu Ala Gly Asn Asp Gly Thr Val Val Ser Lys Pro
450 455 460

Asn Ile Asp Glu Lys Lys Ala Met Gly Cys
465 470

<210> 159
<211> 497
<212> PRT
<213> Selaginella moellendorffii

<400> 159

Met Ala Asp Leu Lys Tyr Ser Arg Leu Val Gln Phe Gly Leu Val Glu
1 5 10 15

Asn Cys Ser Ser Glu Gly Arg Gly Met His Thr Val Ala Ala Asp Leu
20 25 30

Asp Gly Thr Leu Leu Arg Arg Arg Ser Ser Phe Pro Tyr Phe Leu Leu
35 40 45

Val Ala Leu Glu Gly Gly Ser Leu Phe Arg Ala Val Val Leu Gly Leu
50 55 60

Leu Ala Pro Ile Ala Trp Leu Leu Tyr His Phe Val Ser Glu Ala Ala
65 70 75 80

pctau2012001598-seql

Ala Ile Gln Leu Leu Ile Tyr Val Ser Phe Ser Gly Gln Lys Val Ser
85 90 95

Glu Ile Glu Lys Val Ala Thr Ala Val Leu Pro Lys Phe Tyr Ser Gln
100 105 110

Asp Val His Ser Asp Ala Phe Arg Val Phe Ser Ser Cys Gly Lys Lys
115 120 125

Tyr Val Ile Thr Ala Asn Pro Arg Ile Met Val Glu Tyr Phe Cys Lys
130 135 140

Ala Phe Leu Gly Ala Asp Lys Val Ile Gly Thr Glu Ile Glu Val Asp
145 150 155 160

Asp Asp Gly Arg Ala Thr Gly Phe Val Lys Phe Pro Gly Val Leu Val
165 170 175

Ser Gln Asn Lys Arg Thr Ala Leu Lys Leu Glu Phe Val Asp Gly Glu
180 185 190

Leu Pro Glu Ile Gly Leu Gly Asp Arg Glu Thr Asp Phe Ala Phe Met
195 200 205

Ser Leu Cys Lys Glu Ala Tyr Val Val Pro Lys Arg Lys Val Asp Ala
210 215 220

Val Asp Gly Ser Ala Leu Ala Gln Arg Val Ile Phe His Asp Gly Arg
225 230 235 240

Leu Val Gln Leu Pro Thr Pro Phe Val Ala Leu Val Thr Phe Leu Trp
245 250 255

Leu Pro Ile Gly Phe Ala Leu Ala Val Leu Arg Ile Gly Leu Thr Ser
260 265 270

Arg Cys Pro Arg Lys Tyr Met Gly Met Ile Tyr Ala Ile Leu Gly Val
275 280 285

Arg Arg Ile Thr Lys Gly Lys Val Pro Arg Lys Gln Asp Gly Arg Gly
290 295 300

Leu Leu Phe Val Cys Asn His Arg Thr Leu Val Asp Pro Ile Phe Leu
305 310 315 320

Gly Leu Ala Ala Glu Arg Asp Val Thr Gly Leu Thr Tyr Ser Ile Ser
325 330 335

Arg Val Ser Glu Val Leu Ser Pro Ile Val Thr Val Arg Leu Thr Arg
340 345 350

pctau2012001598-seql

Asp Lys Glu Arg Asp Phe Thr Lys Met Lys Ser Leu Leu Glu Arg Gly
355 360 365

Arg Asp Leu Cys Leu Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro Phe
370 375 380

Leu Leu Arg Phe Ser Ala Leu Phe Thr Gln Leu Thr His Arg Ile Val
385 390 395 400

Pro Ala Ala Val Lys Thr Arg Gln Ser Met Phe Asn Gly Thr Ala Val
405 410 415

Arg Gly Trp Lys Gly Met Asp Pro Phe Phe Phe Phe Met Asn Pro Phe
420 425 430

Pro Thr Tyr Glu Val Glu Phe Leu Glu Glu Leu Pro Ala Glu Met Ser
435 440 445

Val Gln Lys Gly Gly Lys Ser Ser Phe Glu Thr Ala Asn Phe Ile Gln
450 455 460

Lys Met Leu Gly Asp Lys Leu Gly Tyr Glu Cys Thr Thr Phe Thr Arg
465 470 475 480

Lys Asp Lys Tyr Leu Met Leu Asn Gly Ser Asp Gly Thr Val Pro Thr
485 490 495

Lys

<210> 160
<211> 534
<212> PRT
<213> Selaginella moellendorffii

<400> 160

Met Ser Leu Asp Gly Gly Lys Val Thr Asp Ser Phe Leu Val Asp Ala
1 5 10 15

Pro Phe Glu Pro Val Glu Arg Cys Ser Ser Ala Gly Arg Glu Lys Gln
20 25 30

Thr Cys Val Ala Asp Phe Asp Gly Thr Leu Ile Arg Gly Arg Ser Ser
35 40 45

Phe Pro Tyr Phe Met Leu Val Ala Leu Glu Gly Gly Gly Leu Ile Arg
50 55 60

Ser Val Ile Leu Gly Leu Phe Ala Pro Leu Ala Trp Ile Leu Tyr His
65 70 75 80

Phe Val Ser Glu Ser Ala Gly Ile Arg Leu Leu Ile Phe Ile Ser Phe
Page 182

pctau2012001598-seql

85

90

95

Ala Gly Met Lys₁₀₀ Ala Lys Gln Ile Glu₁₀₅ Ser Val Ser Arg Ala₁₁₀ Val Leu

Pro Lys Phe₁₁₅ Tyr Ser Glu Asp Val₁₂₀ His Ser Glu Ala Tyr₁₂₅ Arg Val Phe

Ser Ser₁₃₀ Cys Gly Lys Arg Val₁₃₅ Val Val Thr Ala Asn₁₄₀ Pro Arg Ile Met

Val Glu His Phe Ala Lys₁₅₀ Thr Cys Leu Gly Ala₁₅₅ Asp Lys Val Leu Gly₁₆₀

Thr Glu Ile Glu Val₁₆₅ Thr Arg Ser Gly Tyr₁₇₀ Ala Thr Gly Phe Leu₁₇₅ Lys

Arg Pro Gly Val₁₈₀ Leu Val Gly Val Asn₁₈₅ Lys Arg Lys Ala Val₁₉₀ Arg Gln

Glu Phe Gly₁₉₅ Glu His Leu Pro His₂₀₀ Val Gly Ile Gly Asp₂₀₅ Arg Ala Thr

Asp Phe Ala Phe Met Ala Leu₂₁₅ Cys Lys Glu Ala Tyr₂₂₀ Val Val Pro Pro

Ser Lys Val Glu Ala Leu₂₃₀ Ser Arg Asp Lys Leu₂₃₅ Val Lys Pro Val Val₂₄₀

Phe His Asp Gly Arg₂₄₅ Leu Val Gln Arg Pro₂₅₀ Thr Pro Leu Thr Ala₂₅₅ Leu

Val Thr Phe Leu₂₆₀ Trp Leu Pro Val Ser₂₆₅ Phe Val Leu Ala Val₂₇₀ Phe Arg

Ile Met Val₂₇₅ Thr Val Pro Cys Pro₂₈₀ Arg Glu Tyr Val Thr₂₈₅ Ile Val Tyr

Lys Leu₂₉₀ Val Gly Ile Arg Leu₂₉₅ Ile Val Lys Gly Pro₃₀₀ Ile Pro Pro Pro

Lys Lys Arg Gly Glu Ser₃₁₀ Gly Val Leu Phe Val₃₁₅ Ser Ser His Arg Thr₃₂₀

Leu Cys Asp Pro Val₃₂₅ Phe Val Gly Val Gly₃₃₀ Ala Arg Arg Glu Val₃₃₅ Thr

Ala Leu Thr Tyr₃₄₀ Ser Ile Ser Arg Val₃₄₅ Ser Glu Phe Leu Ser₃₅₀ Pro Ile

Lys Thr Val Gly Leu Ser Arg Asp Arg Glu Arg Asp Ala Ala Lys Ile

355

360

365

Lys Ala Leu Leu Gln Lys Gly Asp Leu Cys Ile Cys Pro Glu Gly Thr
 370 375 380

Thr Cys Arg Glu Pro Phe Leu Leu Arg Phe Ser Ala Leu Phe Ala Glu
 385 390 395 400

Leu Thr Asp Lys Ile Val Pro Val Ala Leu Cys Thr Lys Gly Ser Thr
 405 410 415

Phe His Gly Thr Thr Val Arg Gly Trp Lys Gly Leu Asp Pro Phe Phe
 420 425 430

Phe Phe Met Asn Pro Phe Pro Thr Tyr Glu Val Thr Phe Leu Gln Gln
 435 440 445

Leu Pro Pro Glu Leu Thr Val Gln Gln Gly Gly Lys Ser Ala Ile Glu
 450 455 460

Val Ala Asn His Ile Gln Arg Val Ile Ala Asp Thr Leu Gly Phe Glu
 465 470 475 480

Cys Thr Asn Phe Thr Arg Lys Asp Lys Tyr Gly Met Leu Ala Gly Asn
 485 490 495

Asp Gly Thr Val Pro Glu Arg Arg Arg Ser Gly Gly Gly Gly Gly Gly
 500 505 510

Ala Ala Lys Asp Leu Pro Lys Asp Glu Gln Ala Ala Gln Gly Gln Asp
 515 520 525

Gln Val Asn Gly Cys Ser
 530

<210> 161
 <211> 494
 <212> PRT
 <213> Selaginella moellendorffii

<400> 161

Met Ala Glu Phe Pro Pro Ile Glu Lys Cys Gly Thr Thr Ile Ser Arg
 1 5 10 15

Glu Ser Glu Ser Val Ala Ser Asp Leu Asp Gly Thr Leu Leu Arg Gly
 20 25 30

Arg Ser Ser Phe Pro Tyr Phe Phe Leu Thr Thr Phe Glu Thr Ser Gly
 35 40 45

Ile Leu Arg Ser Trp Leu Leu Leu Val Ser Pro Leu Val Trp Leu
 50 55 60

pctau2012001598-seql

```

Leu Tyr His Phe Ile Ser Glu Ala Ala Gly Ile Lys Leu Leu Ile Phe
65          70          75          80

Ala Ser Leu Ala Gly Val Pro Val Ser Gln Ile Glu Cys Ile Ala Arg
          85          90          95

Ser Val Leu Pro Arg Phe Tyr Leu Gln Asp Leu Asn Pro Glu Thr Trp
          100          105          110

Lys Val Phe Ser Ser Phe Gly Lys Lys Ile Val Val Thr Ala Asn Pro
          115          120          125

Arg Ile Leu Val Glu Ser Leu Cys Lys Glu Phe Leu Gly Ala Asp Glu
          130          135          140

Val Ile Gly Thr Glu Leu Glu Val Asp Asn Arg Gly Arg Ala Thr Gly
145          150          155          160

Phe Val Lys Ser Pro Gly Val Leu Val Gly Glu Asn Lys Ala Lys Ala
          165          170          175

Leu Leu Gln Ala Cys Asp Pro Glu Gln Leu Pro Asp Val Gly Leu Gly
          180          185          190

Asp Arg Gln Thr Asp Phe Pro Phe Met Lys Leu Cys Lys Glu Gly Tyr
          195          200          205

Val Val Pro Pro Ser Asn Asn Val Glu Pro Val Pro Lys Ser Gln Leu
          210          215          220

Leu Lys Pro Ile Ile Phe His Asp Gly Arg Leu Val Gln Leu Pro Thr
225          230          235          240

Pro Leu Val Ala Leu Val Thr Leu Leu Trp Ile Pro Val Gly Phe Val
          245          250          255

Leu Ala Cys Leu Arg Ile Ala Ala Gly Ser Leu Leu Pro Met Ser Ile
          260          265          270

Val Tyr His Ala Phe Trp Leu Leu Gly Val Arg Val Thr Val Lys Gly
          275          280          285

Thr Pro Pro Pro Arg Ala Thr Lys Asn Ser Arg Gly Val Leu Phe Ile
          290          295          300

Cys Ser His Arg Thr Leu Leu Asp Pro Ile Phe Leu Ser Cys Ala Leu
305          310          315          320

Gly Arg Gly Val Ser Ala Val Thr Tyr Ser Val Ser Arg Leu Ser Glu
          325          330          335

```

pctau2012001598-seql

Ile Leu Ser Pro Ile Lys Thr Val Ala Leu Ser Arg Asn Arg Tyr Lys
340 345 350

Asp Ala Ala Gln Ile Arg Glu Leu Leu Lys Glu Gly Asp Leu Val Ile
355 360 365

Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro Phe Leu Leu Arg Phe Ser
370 375 380

Ala Leu Phe Ala Glu Leu Ala Asp Asp Ile Val Pro Val Ala Met Asp
385 390 395 400

Asn Arg Met Ser Met Phe His Gly Thr Ala Ala Arg Ala Trp Lys Gly
405 410 415

Met Asp Pro Phe Tyr Phe Phe Met Asn Pro Ser Pro Ser Tyr Glu Val
420 425 430

Thr Phe Leu Gln Gln Leu Pro Glu Glu Leu Thr Cys Lys Gly Gly Lys
435 440 445

Ser Ser His Glu Val Ala Asn His Ile Gln Arg Ala Leu Ala Gly Val
450 455 460

Leu Asp Phe Glu Cys Thr Asn Leu Thr Arg Lys Asp Lys Tyr Lys Ala
465 470 475 480

Leu Val Gly Thr Asp Gly Ser Val Glu Thr Lys Lys Ser Thr
485 490

<210> 162
<211> 494
<212> PRT
<213> Selaginella moellendorffii

<400> 162

Met Ala Glu Phe Pro Pro Ile Glu Lys Cys Ser Ser Thr Ile Ser Arg
1 5 10 15

Glu Ser Glu Ser Val Ala Ser Asp Leu Asp Gly Thr Leu Leu Arg Gly
20 25 30

Arg Ser Ser Phe Pro Tyr Phe Phe Leu Thr Thr Phe Glu Thr Ser Gly
35 40 45

Ile Leu Arg Ser Trp Leu Leu Leu Leu Val Ser Pro Leu Val Trp Leu
50 55 60

Leu Tyr His Phe Ile Ser Glu Ala Ala Gly Ile Lys Leu Leu Ile Phe
65 70 75 80

pctau2012001598-seql

Ala Ser Leu Ala Gly Val Pro Val Ser Gln Ile Glu Cys Ile Ala Arg
85 90 95
Ser Val Leu Pro Arg Phe Tyr Leu Gln Asp Leu Asn Pro Glu Thr Trp
100 105 110
Lys Val Phe Ser Ser Phe Gly Lys Lys Ile Val Val Thr Ala Asn Pro
115 120 125
Arg Ile Leu Val Glu Ser Leu Cys Lys Glu Phe Leu Gly Ala Asp Glu
130 135 140
Val Ile Gly Thr Glu Leu Glu Met Asp Ser Arg Gly Arg Ala Thr Gly
145 150 155 160
Phe Val Lys Ser Pro Gly Val Leu Val Gly Glu Asn Lys Ala Lys Ala
165 170 175
Leu Leu Gln Ala Cys Asp Pro Glu Gln Leu Pro Asp Val Gly Leu Gly
180 185 190
Asp Arg Gln Thr Asp Phe Pro Phe Met Lys Leu Cys Lys Glu Gly Tyr
195 200 205
Val Val Pro Pro Ser Asn Asn Val Glu Pro Val Pro Lys Ser Gln Leu
210 215 220
Leu Lys Pro Ile Ile Phe His Asp Gly Arg Leu Val Gln Leu Pro Thr
225 230 235 240
Pro Leu Val Ala Leu Val Thr Leu Leu Trp Ile Pro Val Gly Phe Val
245 250 255
Leu Ala Cys Leu Arg Ile Ala Ala Gly Ser Leu Leu Pro Met Ser Ile
260 265 270
Val Tyr His Ala Phe Trp Leu Leu Gly Val Arg Val Thr Val Lys Gly
275 280 285
Thr Pro Pro Pro Arg Ala Thr Lys Asn Ser Arg Gly Val Leu Phe Ile
290 295 300
Cys Ser His Arg Thr Leu Leu Asp Pro Ile Phe Leu Ser Cys Ala Leu
305 310 315 320
Gly Arg Gly Val Ser Ala Val Thr Tyr Ser Val Ser Arg Leu Ser Glu
325 330 335
Ile Leu Ser Pro Ile Lys Thr Val Ala Leu Ser Arg Asn Arg Asp Lys
340 345 350

pctau2012001598-seql

Asp Ala Ala Gln Ile Arg Glu Leu Leu Lys Glu Gly Asp Leu Val Ile
355 360 365

Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro Phe Leu Leu Arg Phe Ser
370 375 380

Ala Leu Phe Ala Glu Leu Ala Asp Asp Ile Val Pro Val Ala Met Asp
385 390 395 400

Asn Arg Met Ser Met Phe His Gly Thr Ala Ala Arg Ala Trp Lys Gly
405 410 415

Met Asp Pro Phe Tyr Phe Phe Met Asn Pro Ser Pro Ser Tyr Glu Val
420 425 430

Thr Phe Leu Gln Gln Leu Pro Glu Glu Leu Thr Cys Lys Gly Gly Lys
435 440 445

Ser Ser His Glu Val Ala Asn His Ile Gln Arg Ala Leu Ala Gly Val
450 455 460

Leu Asp Phe Glu Cys Thr Asn Leu Thr Arg Lys Asp Lys Tyr Lys Ala
465 470 475 480

Leu Val Gly Thr Asp Gly Ser Val Glu Ser Lys Lys Ser Thr
485 490

<210> 163
<211> 490
<212> PRT
<213> Selaginella moellendorffii

<400> 163

Met Glu His Pro Tyr Asp Val Ile Asp Asn Cys Ser Ala Asp Asn Arg
1 5 10 15

Glu Lys Gln Ser Val Val Ala Asp Leu Asp Gly Ser Cys Val Arg Ser
20 25 30

Arg Ser Ser Phe Pro Tyr Phe Met Leu Val Ala Leu Glu Gly Gly Ser
35 40 45

Leu Leu Arg Ser Val Leu Leu Met Val Met Ala Pro Val Ala Trp Leu
50 55 60

Leu Tyr His Phe Val Ser Glu Ala Ala Gly Ile Lys Val Leu Ile Phe
65 70 75 80

Phe Ser Cys Ala Gly Gln Arg Ala Arg Asp Ile Glu Ala Val Ala Arg
85 90 95

pctau2012001598-seql

Ala Val Leu Pro Lys Phe Tyr Ser Glu Asp Val His Ser Ser Thr Trp
100 105 110

Arg Val Phe Ser Ser Cys Gly Lys Arg Tyr Ile Leu Thr Ala Ser Pro
115 120 125

Arg Leu Met Val Glu His Phe Ala Thr Thr Tyr Leu Gly Ala Asp Lys
130 135 140

Val Ile Gly Thr Glu Leu Gln Val Thr Arg Ser Gly Tyr Ala Thr Gly
145 150 155 160

Phe Val Lys Ser Pro Gly Val Leu Val Gly Val Asn Lys Arg Arg Ala
165 170 175

Leu Lys Ala Val Phe Gly Glu Glu Val Pro Asp Leu Gly Ile Gly Asp
180 185 190

Arg Ser Ser Asp Tyr Pro Phe Met Ala Phe Cys Lys Glu Ala Tyr Ile
195 200 205

Val Pro Ser Gln Pro Ala Asp Ala Val Pro Ala His Lys Leu Ala Arg
210 215 220

Lys Ile Ile Phe His Asp Gly Arg Leu Val Gln Arg Pro Thr Pro Phe
225 230 235 240

Val Ala Leu Val Thr Phe Leu Trp Leu Pro Ala Gly Phe Leu Leu Ala
245 250 255

Phe Phe Arg Ile Leu Val Ala Leu Pro Cys Pro Val Glu Tyr Val Pro
260 265 270

Leu Leu Tyr Lys Val Ile Gly Val Arg Leu Ile Val Lys Gly Lys Val
275 280 285

Pro Pro Pro Ser Lys Asn Gly Ile Leu Phe Val Cys Thr His Arg Thr
290 295 300

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

Ala Leu Thr Tyr Ser Ile Ser Arg Val Ser Glu Val Leu Ser Pro Ile
325 330 335

Lys Thr Val Pro Leu Ser Arg Asp Arg Glu Arg Asp Ala Ser Asn Ile
340 345 350

Lys Arg Leu Leu Leu Lys Glu Asn Leu Ala Ile Cys Pro Glu Gly Thr
355 360 365

pctau2012001598-seql

Thr Cys Arg Glu Pro Phe Leu Leu Arg Phe Ser Ala Leu Phe Ala Glu
370 375 380

Leu Ser Glu Asn Ile Val Pro Val Ala Thr Ser Thr Lys Thr Ser Ile
385 390 395 400

Phe His Gly Thr Thr Val Arg Gly Trp Lys Gly Met Asp Pro Phe Phe
405 410 415

Phe Phe Met Asn Pro Phe Pro Val Tyr Glu Val Thr Phe Leu Glu Arg
420 425 430

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

Ala Asn Tyr Ile Gln Arg Val Leu Ala Asp Thr Leu Gly Phe Glu Cys
450 455 460

Thr Asn Phe Thr Arg Lys Asp Lys Tyr Gly Met Leu Ile Gly Ser Asp
465 470 475 480

Gly Thr Val Pro Ala Arg Lys Lys Ser Phe
485 490

<210> 164
<211> 497
<212> PRT
<213> Selaginella moellendorffii

<400> 164

Met Ala Asp Leu Lys Tyr Ser Arg Leu Phe Arg Phe Gly Leu Val Glu
1 5 10 15

Asn Cys Ser Ser Glu Gly Arg Gly Met His Thr Val Ala Ala Asp Leu
20 25 30

Asp Gly Thr Leu Leu Arg Gly Arg Ser Ser Phe Pro Tyr Phe Leu Leu
35 40 45

Val Ala Leu Glu Gly Gly Ser Leu Phe Arg Ala Val Val Leu Gly Leu
50 55 60

Leu Ala Pro Ile Ala Trp Leu Leu Tyr His Phe Val Ser Glu Ala Ala
65 70 75 80

Ala Ile Gln Leu Leu Ile Tyr Val Ser Phe Ser Gly Gln Lys Val Ser
85 90 95

Glu Ile Glu Lys Val Ala Thr Ala Val Leu Pro Lys Phe Tyr Ser Gln
100 105 110

Asp Val His Ser Asp Ala Phe Arg Val Phe Ser Ser Cys Gly Lys Lys
Page 190

115

120

125

Tyr Val Ile Thr Ala Asn Pro Arg Ile Met Val Glu Tyr Phe Cys Lys
 130 135 140
 Ala Phe Leu Gly Ala Asp Lys Val Ile Gly Thr Glu Ile Glu Val Asp
 145 150 155 160
 Asp Asp Gly Arg Ala Thr Gly Phe Val Lys Phe Pro Gly Val Leu Val
 165 170 175
 Ser Gln Asn Lys Arg Thr Ala Leu Lys Leu Glu Phe Val Asp Gly Glu
 180 185 190
 Leu Pro Glu Ile Gly Leu Gly Asp Arg Glu Thr Asp Phe Ala Phe Met
 195 200 205
 Ser Leu Cys Lys Glu Ala Tyr Val Val Pro Lys Arg Lys Val Asp Ala
 210 215 220
 Val Asp Gly Ser Ala Leu Ala Gln Arg Val Ile Phe His Asp Gly Arg
 225 230 235 240
 Leu Val Gln Leu Pro Thr Pro Phe Val Ala Leu Val Thr Phe Leu Trp
 245 250 255
 Leu Pro Ile Gly Phe Ala Leu Ala Val Leu Arg Ile Gly Leu Thr Ser
 260 265 270
 Arg Cys Pro Arg Lys Tyr Met Gly Met Ile Tyr Ala Ile Leu Gly Val
 275 280 285
 Arg Arg Ile Thr Lys Gly Lys Val Pro Arg Lys Gln Asp Gly Arg Gly
 290 295 300
 Leu Leu Phe Val Cys Asn His Arg Thr Leu Val Asp Pro Ile Phe Val
 305 310 315 320
 Gly Leu Ala Ala Glu Arg Asp Val Thr Gly Leu Thr Tyr Ser Ile Ser
 325 330 335
 Arg Val Ser Glu Val Leu Ser Pro Ile Val Thr Val Arg Leu Thr Arg
 340 345 350
 Asp Lys Gln Arg Asp Phe Thr Lys Met Lys Ser Leu Leu Glu Arg Gly
 355 360 365
 Arg Asp Leu Cys Leu Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro Phe
 370 375 380
 Leu Leu Arg Phe Ser Ala Leu Phe Thr Glu Leu Thr His Arg Ile Val

pctau2012001598-seq1

385 390 395 400

Pro Ala Ala Val Lys Thr Arg Gln Ser Met Phe Asn Gly Thr Ala Val
405 410 415

Arg Gly Trp Lys Gly Met Asp Pro Phe Phe Phe Phe Met Asn Pro Phe
420 425 430

Pro Thr Tyr Glu Val Glu Phe Leu Glu Glu Leu Pro Val Glu Met Ser
435 440 445

Val Gln Lys Gly Gly Lys Ser Ser Phe Glu Thr Ala Asn Phe Ile Gln
450 455 460

Lys Met Leu Gly Asp Lys Leu Gly Tyr Glu Cys Thr Thr Phe Thr Arg
465 470 475 480

Lys Asp Lys Tyr Leu Met Leu Asn Gly Ser Asp Gly Thr Val Pro Thr
485 490 495

Lys

<210> 165

<211> 534

<212> PRT

<213> Selaginella moellendorffii

<400> 165

Met Ser Leu Asp Gly Gly Lys Val Thr Asp Ser Phe Leu Val Asp Ala
1 5 10 15

Pro Phe Glu Pro Val Glu Arg Cys Ser Ser Ala Gly Arg Glu Lys Gln
20 25 30

Thr Cys Val Ala Asp Phe Asp Gly Thr Leu Ile Arg Gly Arg Ser Ser
35 40 45

Phe Pro Tyr Phe Met Leu Val Ala Leu Glu Gly Gly Gly Leu Ile Arg
50 55 60

Ser Val Ile Leu Gly Leu Phe Ala Pro Leu Ala Trp Ile Leu Tyr His
65 70 75 80

Phe Val Ser Glu Ser Ala Gly Ile Arg Leu Leu Ile Phe Ile Ser Phe
85 90 95

Ala Gly Met Lys Ala Lys Gln Ile Glu Ser Val Ser Arg Ala Val Leu
100 105 110

Pro Lys Phe Tyr Ser Glu Asp Val His Ser Glu Ala Tyr Arg Val Phe
115 120 125

pctau2012001598-seql

```

Ser Ser Cys Gly Lys Arg Val Val Val Thr Ala Asn Pro Arg Ile Met
130      135      140

Val Glu His Phe Ala Lys Thr Cys Leu Gly Ala Asp Lys Val Leu Gly
145      150      155      160

Thr Glu Ile Glu Val Thr Arg Ser Gly Tyr Ala Thr Gly Phe Leu Lys
165      170

Arg Pro Gly Val Leu Val Gly Val Asn Lys Arg Lys Ala Val Arg Gln
180      185      190

Glu Phe Gly Glu His Leu Pro His Val Gly Ile Gly Asp Arg Ala Thr
195      200      205

Asp Phe Ala Phe Met Ala Leu Cys Lys Glu Ala Tyr Val Val Pro Pro
210      215      220

Ser Lys Val Glu Ala Leu Ser Arg Asp Lys Leu Val Lys Pro Val Val
225      230      235      240

Phe His Asp Gly Arg Leu Val Gln Arg Pro Thr Pro Leu Thr Ala Leu
245      250      255

Val Thr Phe Leu Trp Leu Pro Val Ser Phe Val Leu Ala Val Phe Arg
260      265      270

Ile Met Val Thr Val Pro Cys Pro Arg Glu Tyr Val Thr Ile Val Tyr
275      280      285

Lys Leu Val Gly Ile Arg Leu Ile Val Lys Gly Pro Ile Pro Pro Pro
290      295      300

Lys Lys Arg Gly Glu Ser Gly Val Leu Phe Val Ser Ser His Arg Thr
305      310      315      320

Leu Cys Asp Pro Val Phe Val Gly Val Gly Ala Arg Arg Glu Val Thr
325      330      335

Ala Leu Thr Tyr Ser Ile Ser Arg Val Ser Glu Phe Leu Ser Pro Ile
340      345      350

Lys Thr Val Gly Leu Ser Arg Asp Arg Glu Arg Asp Ala Ala Lys Ile
355      360      365

Lys Ala Leu Leu Gln Lys Gly Asp Leu Cys Ile Cys Pro Glu Gly Thr
370      375      380

Thr Cys Arg Glu Pro Phe Leu Leu Arg Phe Ser Ala Leu Phe Ala Glu
385      390      395      400

```

pctau2012001598-seql

Leu Thr Asp Lys Ile Val Pro Val Ala Leu Cys Thr Lys Gly Ser Thr
405 410 415

Phe His Gly Thr Thr Val Arg Gly Trp Lys Gly Leu Asp Pro Phe Phe
420 425 430

Phe Phe Met Asn Pro Phe Pro Thr Tyr Glu Val Thr Phe Leu Gln Gln
435 440 445

Leu Pro Pro Glu Leu Thr Val Gln Gln Gly Gly Lys Ser Ala Ile Glu
450 455 460

Val Ala Asn His Ile Gln Arg Val Ile Ala Asp Thr Leu Gly Phe Glu
465 470 475 480

Cys Thr Asn Phe Thr Arg Lys Asp Lys Tyr Gly Met Leu Ala Gly Asn
485 490 495

Asp Gly Thr Val Pro Glu Arg Arg Arg Ser Gly Gly Gly Gly Gly Gly
500 505 510

Ala Ala Lys Asp Leu Pro Glu Asp Glu Gln Ala Ala Gln Gly Gln Asp
515 520 525

Gln Val Asn Gly Cys Ser
530

<210> 166
<211> 497
<212> PRT
<213> Oryza sativa

<400> 166

Met Val Ser Arg Arg Phe Lys Pro Val Glu Glu Cys Ser Ser Asp Gly
1 5 10 15

Arg Ser Glu Gln Thr Val Ala Ala Asp Phe Asp Gly Thr Leu Val Arg
20 25 30

Ser Arg Ser Ala Phe Pro Tyr Tyr Leu Leu Val Ala Leu Glu Ala Gly
35 40 45

Ser Val Leu Arg Ala Val Val Leu Leu Leu Ser Val Pro Phe Val Tyr
50 55 60

Val Thr Tyr Ile Phe Phe Ser Glu Ser Leu Ala Ile Ser Thr Leu Val
65 70 75 80

Tyr Ile Ser Val Ala Gly Leu Lys Val Arg Asn Ile Glu Met Val Ala
85 90 95

pctau2012001598-seql

Arg Ser Val Leu Pro Lys Phe Tyr Ala Glu Asp Val His Pro Glu Ser
100 105 110
Trp Arg Val Phe Asn Ser Phe Gly Lys Arg Tyr Ile Ile Thr Ala Ser
115 120 125
Pro Arg Ile Met Val Glu His Phe Ala Lys Thr Phe Leu Gly Ala Asp
130 135 140
Lys Val Val Gly Thr Glu Leu Glu Val Gly Lys Asn Gly Lys Ala Thr
145 150 155 160
Gly Phe Met Val Lys Pro Gly Val Leu Val Gly Asp His Lys Arg Gln
165 170 175
Ala Val Val Lys Glu Leu Arg Asp Ala Val Pro Asp Val Gly Leu Gly
180 185 190
Asp Arg Glu Thr Asp Phe Asp Phe Met Ser Ile Cys Lys Glu Ala Tyr
195 200 205
Leu Val Thr Ser Arg Lys Tyr Ser Ala Val Pro Lys Asn Gln Leu Leu
210 215 220
Ser Pro Leu Ile Leu His Asp Gly Arg Leu Val Gln Arg Pro Thr Pro
225 230 235 240
Leu Val Ala Leu Val Thr Phe Leu Trp Met Pro Phe Gly Phe Ala Leu
245 250 255
Ala Leu Leu Arg Val Tyr Val Asn Leu Pro Leu Pro Glu Arg Ile Val
260 265 270
Phe Tyr Thr Tyr Lys Leu Met Gly Ile Arg Leu Ile Val Lys Gly Asn
275 280 285
Pro Pro Pro Pro Pro Lys Lys Gly His Pro Gly Val Leu Phe Val Cys
290 295 300
Asn His Arg Thr Val Leu Asp Pro Val Glu Val Ala Val Ala Leu Arg
305 310 315 320
Arg Lys Val Ser Cys Val Thr Tyr Ser Ile Ser Lys Phe Ser Glu Leu
325 330 335
Ile Ser Pro Ile Lys Ala Val Ala Leu Ser Arg Glu Arg Glu Lys Asp
340 345 350
Ala Glu Asn Ile Arg Arg Leu Leu Glu Glu Gly Asp Leu Val Ile Cys
355 360 365

pctau2012001598-seql

Pro Glu Gly Thr Thr Cys Arg Glu Pro Phe Leu Leu Arg Phe Ser Ala
370 375 380

Leu Phe Ala Glu Leu Thr Asp Arg Ile Val Pro Val Ala Ile Asn Thr
385 390 395 400

Lys Glu Ser Met Phe His Gly Ser Thr Val Arg Gly Phe Lys Leu Met
405 410 415

Asp Pro Tyr Phe Phe Phe Met Asn Pro Arg Pro Thr Tyr Glu Ile Thr
420 425 430

Phe Leu Asn Gln Leu Pro Lys Glu Leu Thr Cys Ser Gly Gly Lys Ser
435 440 445

Pro Ile Glu Val Ala Asn Tyr Ile Gln Lys Thr Leu Ser Gly Gln Leu
450 455 460

Gly Phe Glu Cys Thr Ala Ile Thr Arg Lys Glu Lys Tyr Ser Ile Leu
465 470 475 480

Ala Gly Thr Asp Gly Arg Val Pro Ser Lys Asn Lys Glu Lys Glu Lys
485 490 495

Asn

<210> 167
<211> 506
<212> PRT
<213> Oryza sativa

<400> 167

Met Ala Ala Ala Ala Ala Ala Gly Arg Arg Arg Phe Phe Pro Pro Val
1 5 10 15

Thr Ala Tyr Asp Ala Ala Ala Gly Ala Arg Arg Thr Val Ala Ala Asp
20 25 30

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

Leu Val Ala Leu Glu Ala Gly Ser Tyr Leu Arg Ala Leu Ala Leu Leu
50 55 60

Leu Ala Ala Pro Trp Leu Leu Ala Leu Tyr Val Gly Val Ser Glu Ala
65 70 75 80

Ala Ala Ile Ala Leu Leu Val Phe Ile Thr Phe Ala Gly Leu Arg Val
85 90 95

pctau2012001598-seql

Arg Asp Val Glu Ala Val Ala Arg Ala Val Leu Pro Arg His Tyr Ala
 100 105 110
 Ala Gly Val Arg Ala Asp Thr Trp Ala Val Phe His Gly Cys Ala Glu
 115 120 125
 Arg Arg Val Val Val Thr Ala Ser Pro Ala Val Met Val Gly Glu Phe
 130 135 140
 Val Arg Glu Phe Leu Gly Ala Glu Val Ala Gly Thr Glu Leu Glu Thr
 145 150 155 160
 Phe Ala Ser Gly Lys Arg Phe Thr Gly Arg Ile Lys Ala Val Leu Val
 165 170 175
 Gly Glu Lys Lys Arg Glu Val Val Glu Arg Leu Phe Ala Gly Gly Asp
 180 185 190
 Met Pro Asp Val Gly Leu Gly Asp Arg Glu Ser Asp His Asp Phe Met
 195 200 205
 Ala Ile Cys Lys Glu Ala Tyr Met Val Pro Lys Asn Lys Arg Ala Pro
 210 215 220
 Arg Ala Ala Ala Asp Glu Leu Leu Ser Arg Ala Ile Phe His Asp Gly
 225 230 235 240
 Arg Leu Val Arg Arg Pro Glu Pro Ala Ser Ala Leu Phe Ala Leu Ala
 245 250 255
 Tyr Leu Pro Val Gly Phe Ala Val Ala Leu Leu Arg Val Phe Leu Asn
 260 265 270
 Leu Pro Val Pro Ala Arg Leu Val Arg His Thr Tyr Arg Leu Thr Gly
 275 280 285
 Ile Arg Leu Ala Val Arg Gly Ala Pro Pro Pro Pro Arg Pro Gly
 290 295 300
 Thr Pro Gly Ser Leu Leu Val Cys Asn His Arg Thr Ala Leu Asp Pro
 305 310 315 320
 Ile Ile Val Ser Ile Ala Leu Gly Arg Pro Val Thr Cys Val Thr Tyr
 325 330 335
 Ser Val Ser Arg Leu Ser Thr Ala Ile Ser Pro Ile Arg Ala Ala Ala
 340 345 350
 Leu Thr Arg Asp Arg Ala Ala Asp Ala Ala Arg Ile Ala Ala Leu Leu
 355 360 365

pctau2012001598-seql

Glu Glu Gly Asp Val Val Val Cys Pro Glu Gly Thr Thr Cys Arg Glu
370 375 380

Pro Tyr Leu Leu Arg Phe Ser Ala Leu Phe Ala Glu Leu Thr Ala Arg
385 390 395 400

Ile Val Pro Val Ala Val Glu Ala Arg Gln Gly Thr Tyr Tyr Gly Ser
405 410 415

Thr Ala Arg Gly Trp Lys Phe Leu Asp Pro Tyr Phe Phe Tyr Met Asn
420 425 430

Pro Arg Pro Gly Tyr Glu Val Thr Phe Leu Pro Ala Leu Arg Pro Glu
435 440 445

Glu Thr Cys Val Ala Gly Gly Arg Ser Ala Val Glu Val Ala Asn His
450 455 460

Val Gln Arg Val Ile Ala Lys Glu Leu Gly Phe Gln Cys Thr Thr Leu
465 470 475 480

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

Ala Ala Ala Ala Asp Lys Pro Lys Ala Asn
500 505

<210> 168
<211> 515
<212> PRT
<213> Zea mays

<400> 168

Met Ala Ala Arg Phe Pro Pro Val Ser Ser Tyr Asp Ala Ser Ala Arg
1 5 10 15

Val Arg Arg Thr Ala Ala Ala Asp Leu Asp Gly Thr Leu Leu Ala Ser
20 25 30

Ser Ser Ala Phe Pro Tyr Tyr Phe Leu Val Ala Leu Glu Ala Gly Ser
35 40 45

Cys Ala Arg Ala Leu Leu Leu Leu Ala Ala Pro Leu Leu Leu Ala
50 55 60

Leu Tyr Thr Leu Val Ser Glu Ala Ala Ala Ile Ala Leu Leu Ala Phe
65 70 75 80

Val Thr Phe Ala Gly Leu Arg Val Arg Asp Val Glu Ala Val Ala Arg
85 90 95

Gly Val Leu Pro Arg His Tyr Ala Ala Gly Val Arg Ala Asp Thr Trp

pctau2012001598-seql

100 105 110
 Ala Val Phe Arg Asp Cys Gly Ala Ala Arg Arg Val Val Val Thr Ala
 115 120 125
 Ser Pro Ala Val Met Val Ala Asp Phe Val Arg Glu Phe Leu Gly Ala
 130 135 140
 Glu Leu Ala Ala Thr Glu Leu Glu Thr Cys Arg Ala Leu Gly Asp Ala
 145 150 155 160
 Trp Phe Thr Gly Arg Ile Lys Ala Val Leu Val Gly Glu Arg Lys Ala
 165 170 175
 Glu Val Val Arg Arg Leu Phe Ala Ala Gly Asp Leu Pro Asp Val Gly
 180 185 190
 Leu Gly Asp Arg Glu Ser Asp His Asp Phe Met Ala Ile Cys Lys Glu
 195 200 205
 Ala Tyr Met Val Pro Pro Asp Arg Arg Ala Pro Arg Ala Ala Ala Asp
 210 215 220
 Ala Leu Leu Ser Arg Ala Val Phe His Asp Gly Arg Leu Val Arg Arg
 225 230 235 240
 Pro Asp Pro Ala His Ala Leu Phe Ala Leu Ala Tyr Leu Pro Leu Gly
 245 250 255
 Phe Leu Leu Ala Leu Phe Arg Val Leu Phe Asn Leu Ala Met Pro Ala
 260 265 270
 Arg Leu Val Arg His Thr Tyr Arg Leu Thr Gly Ile Arg Leu Arg Val
 275 280 285
 Arg Gly Thr Pro Pro Pro Ala Pro Ala Pro Gly Ala Pro Gly Ser Leu
 290 295 300
 Leu Val Cys Asn His Arg Thr Ala Leu Asp Pro Ile Ile Val Ser Val
 305 310 315 320
 Ala Leu Gly Arg Pro Val Ala Cys Val Thr Tyr Ser Val Ser Arg Leu
 325 330 335
 Ser Thr Ala Val Ser Pro Val Arg Ala Val Ala Leu Ser Arg Asp Arg
 340 345 350
 Ala Ala Asp Ala Ala Arg Ile Ala Ala Leu Leu Ala Glu Gly Asp Val
 355 360 365
 Val Val Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro Cys Leu Leu Arg
 Page 199

pctau2012001598-seql

370

375

380

Phe Ser Ala Leu Phe Ala Glu Leu Thr Asp Arg Ile Val Pro Val Ala
385 390 395 400

Leu Asp Ala Arg Gln Gly Thr Tyr Tyr Gly Ser Thr Ala Arg Gly Trp
405 410 415

Lys Trp Leu Asp Pro Tyr Phe Phe Tyr Met Asn Pro Arg Pro Gly Tyr
420 425 430

Asp Val Thr Phe Leu Pro Pro Leu Arg Pro Glu Glu Thr Cys Gly Ala
435 440 445

Gly Gly Arg Ser Ala Val Asp Val Ala Asn His Val Gln Thr Leu Ile
450 455 460

Ala Lys Glu Leu Gly Phe Arg Cys Thr Lys Leu Thr Arg Lys Asp Lys
465 470 475 480

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

Asn Glu Lys Gly Ala Ala Asp Asp Ser Ala Ala Thr Ala Thr Thr Lys
500 505 510

Lys Leu Val
515

<210> 169
<211> 500
<212> PRT
<213> Zea mays

<220>
<221> misc_feature
<222> (92)..(92)
<223> Xaa can be any naturally occurring amino acid

<400> 169

Met Val Ala Ser Pro Arg Phe Lys Pro Ile Glu Glu Cys Cys Ser Glu
1 5 10 15

Gly Arg Ser Glu Gln Thr Val Ala Ala Asp Leu Asp Gly Thr Leu Leu
20 25 30

Ile Ser Arg Ser Ala Phe Pro Tyr Tyr Leu Leu Val Ala Leu Glu Ala
35 40 45

Gly Ser Val Leu Arg Ala Ala Leu Leu Leu Leu Ser Val Pro Phe Val
50 55 60

pctau2012001598-seql

Tyr Val Thr Tyr Ala Phe Phe Ser Glu Ser Leu Ala Ile Ser Thr Leu
 65 70 75 80
 Val Tyr Ile Ser Val Ala Gly Leu Lys Val Arg Xaa Ile Glu Met Val
 85 90 95
 Ala Arg Ser Val Leu Pro Arg Phe Tyr Ala Gly Asp Val His Pro Glu
 100 105 110
 Ser Trp Arg Val Phe Ser Ser Phe Gly Arg Arg Tyr Val Val Thr Ala
 115 120 125
 Ser Pro Arg Val Met Val Glu Pro Phe Ala Arg Ala Phe Leu Gly Ala
 130 135 140
 Asp Lys Val Val Gly Thr Glu Leu Glu Val Gly Arg Asp Gly Lys Ala
 145 150 155 160
 Thr Gly Phe Met Ala Arg Pro Gly Val Leu Val Gly Asp His Lys Lys
 165 170 175
 Lys Ala Val Val Lys Glu Leu Gly Asp Ala Leu Pro Asp Val Gly Met
 180 185 190
 Gly Asp Arg Glu Thr Asp Phe Asp Phe Met Ser Ile Cys Lys Glu Ala
 195 200 205
 Tyr Leu Val Thr Ser Arg Lys Tyr Ser Pro Val Pro Arg Asn Gln Leu
 210 215 220
 Leu Ser Pro Leu Ile Val His Asp Gly Arg Leu Val Gln Arg Pro Thr
 225 230 235 240
 Pro Leu Val Ala Leu Val Thr Phe Leu Trp Met Pro Phe Gly Phe Ala
 245 250 255
 Leu Ala Leu Met Arg Val Tyr Ile Asn Leu Pro Leu Pro Glu Arg Ile
 260 265 270
 Val Tyr Tyr Thr Tyr Lys Leu Met Gly Ile Arg Leu Val Val Lys Gly
 275 280 285
 Thr Pro Pro Pro Pro Pro Lys Lys Gly His Pro Gly Val Leu Phe Val
 290 295 300
 Cys Asn His Arg Thr Val Leu Asp Pro Val Glu Val Ala Val Ala Leu
 305 310 315 320
 Arg Arg Lys Val Ser Cys Val Thr Tyr Ser Ile Ser Lys Phe Ser Glu
 325 330 335

pctau2012001598-seql

Leu Ile Ser Pro Ile Lys Ala Val Ala Leu Ser Arg Glu Arg Asp Lys
340 345 350

Asp Ala Glu Asn Ile Arg Arg Leu Leu Glu Glu Gly Asp Leu Val Ile
355 360 365

Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro Phe Leu Leu Arg Phe Ser
370 375 380

Ala Leu Phe Ala Glu Leu Thr Asp Arg Ile Val Pro Val Ala Ile Asn
385 390 395 400

Thr Lys Glu Thr Met Phe His Gly Ser Thr Val Arg Gly Phe Lys Leu
405 410 415

Met Asp Pro Tyr Phe Phe Phe Met Asn Pro Arg Pro Thr Tyr Glu Ile
420 425 430

Thr Phe Leu Thr Gln Leu Pro Lys Asp Leu Thr Cys Ser Gly Gly Lys
435 440 445

Ser Pro Ile Glu Val Ala Asn Tyr Ile Gln Lys Thr Leu Ser Gly Gln
450 455 460

Leu Gly Phe Glu Cys Thr Ser Ile Thr Arg Lys Glu Lys Tyr Gly Met
465 470 475 480

Leu Ala Gly Thr Asp Gly Arg Val Pro Ser Lys Asn Lys Glu Lys Glu
485 490 495

Lys Asp Lys Asn
500

<210> 170
<211> 486
<212> PRT
<213> Zea mays

<400> 170

Met Val Ser Arg Arg Phe Lys Pro Ile Glu Leu Cys Asp Ser Glu Gly
1 5 10 15

Arg Ser Arg Gln Thr Val Ala Ala Asp Leu Asp Gly Thr Leu Leu Leu
20 25 30

Ser Arg Ser Ala Phe Pro Tyr Tyr Leu Leu Ile Ala Leu Glu Ala Gly
35 40 45

Ser Leu Leu Arg Ala Val Ala Leu Leu Leu Ser Val Pro Leu Val Tyr
50 55 60

Leu Thr Tyr Val Thr Val Ser Glu Thr Leu Ala Val Arg Ala Phe Ile
Page 202

pctau2012001598-seql

65					70					75					80
Tyr	Val	Ala	Met	Ala	Gly	Leu	Glu	Ala	Arg	Asp	Ile	Glu	Ala	Val	Ala
				85					90					95	
Arg	Ser	Val	Leu	Pro	Ser	Phe	Tyr	Ala	Gly	Asp	Val	His	Pro	Glu	Gly
			100					105					110		
Trp	Arg	Val	Phe	Arg	Ser	Phe	Gly	Arg	Arg	Cys	Val	Val	Thr	Ala	Ser
		115					120					125			
Pro	Arg	Leu	Met	Val	Glu	Pro	Phe	Ala	Lys	Ala	Phe	Leu	Gly	Ala	Asp
	130					135					140				
Val	Val	Ile	Gly	Thr	Glu	Val	Glu	Val	Ala	Gln	Ser	Gly	Lys	Ala	Thr
145					150					155					160
Gly	Phe	Val	Ala	Gly	Pro	Gly	Val	Leu	Val	Gly	Glu	His	Lys	Arg	Arg
				165					170					175	
Ala	Val	Ala	Arg	Glu	Phe	Gly	Asp	Ser	Leu	Pro	Asp	Val	Gly	Met	Gly
			180					185					190		
Asp	Arg	Glu	Ser	Asp	Phe	Asp	Phe	Met	Ser	Ile	Cys	Lys	Glu	Ala	Tyr
		195					200					205			
Thr	Val	Thr	Arg	Gln	Lys	Tyr	Arg	Ala	Leu	Pro	Arg	Asp	Gln	Leu	Gln
	210					215					220				
Gly	Arg	Ala	Ile	Leu	His	Asp	Gly	Arg	Leu	Ala	Arg	Arg	Pro	Thr	Ala
225					230					235					240
Thr	Asn	Ala	Leu	Leu	Thr	Phe	Leu	Trp	Met	Pro	Leu	Gly	Phe	Ala	Leu
				245					250					255	
Ala	Leu	Val	Arg	Ala	His	Leu	His	Leu	Leu	Leu	Pro	Val	Arg	Ala	Leu
			260					265					270		
Ala	Tyr	Ala	Tyr	Lys	Leu	Thr	Gly	Val	Lys	Leu	Val	Val	Arg	Gly	Asn
		275					280					285			
Arg	Pro	Pro	Pro	Ser	Lys	Lys	Lys	Lys	Gly	Asp	Gln	Leu	Gly	Val	Leu
	290					295					300				
Phe	Val	Cys	Asn	His	Arg	Ser	Thr	Leu	Asp	Ala	Val	Ala	Val	Gly	Val
305					310					315					320
Ala	Leu	Gly	Arg	Lys	Val	Arg	Trp	Val	Val	Thr	Asp	Gly	Gly	Ala	Gly
				325					330					335	
Ala	Ser	Arg	Phe	Ser	Glu	Pro	Val	Val	Ser	Pro	Ile	Met	Thr	Gly	Val

pctau2012001598-seql

340

345

350

Pro Leu Pro Val Pro Thr Ser Leu Glu Gly Asp Ala Asp Ala Asp Ala
 355 360 365

Ala Pro Arg Ile Arg Arg Leu Leu Glu Glu Gly Asp Asp Val Val Val
 370 375 380

Phe Pro Glu Gly Ala Ile Cys Arg Glu Pro Phe Leu Leu Arg Phe Gly
 385 390 395 400

Ala Leu Phe Ala Glu Val Thr Asp Arg Ile Val Pro Val Ala Ile Glu
 405 410 415

Pro Arg Glu Ser Met Phe His Gly Ser Thr Ala Arg Gly Leu Arg Gly
 420 425 430

Met Asp Pro Tyr Phe Phe Phe Met Asn Pro Arg Pro Thr Tyr Glu Val
 435 440 445

Thr Phe Leu Asn Gln Leu Pro Gly Glu Leu Thr Cys Gly Gly Gly Arg
 450 455 460

Ser Pro Val Glu Val Ala Ser Tyr Val Gln Glu Val Leu Ala Ala Gln
 465 470 475 480

Leu Gly Phe Glu Cys Thr
 485

<210> 171
 <211> 505
 <212> PRT
 <213> Brassica napus

<400> 171

Met Ser Pro Ala Lys Lys Lys Ser Thr Ser Phe Pro Pro Ile Ser Glu
 1 5 10 15

Cys Lys Ser Gly Glu Tyr Asp Ser Ile Ala Ala Asp Leu Asp Gly Thr
 20 25 30

Leu Leu Leu Ser Arg Ser Ser Phe Pro Tyr Phe Met Leu Val Ala Ile
 35 40 45

Glu Ala Gly Ser Leu Phe Arg Gly Leu Ile Leu Leu Leu Ser Leu Pro
 50 55 60

Ile Val Ile Ile Ala Tyr Leu Phe Val Ser Glu Ala Leu Gly Ile Gln
 65 70 75 80

Ile Leu Ile Tyr Ile Ser Phe Ala Gly Ile Lys Ile Arg Asp Ile Glu
 85 90 95

pctau2012001598-seql

Leu Val Ser Arg Ala Val Leu Pro Arg Phe Tyr Ala Ala Asp Val Arg
 100 105 110
 Lys Asp Ser Phe Glu Val Phe Asp Lys Cys Lys Arg Lys Val Val Val
 115 120 125
 Thr Ala Asn Pro Ile Val Met Val Glu Pro Phe Val Lys Asp Tyr Leu
 130 135 140
 Gly Gly Asp Lys Val Leu Gly Thr Glu Ile Glu Val Asn Pro Lys Thr
 145 150 155 160
 Met Lys Ala Thr Gly Phe Val Lys Lys Pro Gly Val Leu Val Gly Asp
 165 170 175
 Leu Lys Arg Leu Ala Ile Leu Lys Glu Phe Gly Asp Glu Ser Pro Asp
 180 185 190
 Leu Gly Leu Gly Asp Arg Thr Ser Asp His Asp Phe Met Ser Ile Cys
 195 200 205
 Lys Glu Gly Tyr Met Val His Glu Thr Lys Ser Ala Thr Thr Val Pro
 210 215 220
 Ile Glu Arg Leu Lys Asn Arg Ile Ile Phe His Asp Gly Arg Leu Val
 225 230 235 240
 Gln Arg Pro Thr Pro Leu Asn Ala Leu Ile Ile Tyr Leu Trp Leu Pro
 245 250 255
 Phe Gly Phe Met Leu Ser Ile Phe Arg Val Tyr Phe Asn Leu Pro Leu
 260 265 270
 Pro Glu Arg Phe Val Arg Tyr Thr Tyr Glu Ile Leu Gly Ile His Leu
 275 280 285
 Thr Ile Arg Gly His Arg Pro Pro Pro Pro Ser Pro Gly Thr Pro Gly
 290 295 300
 Asn Leu Tyr Val Leu Asn His Arg Thr Ala Leu Asp Pro Ile Ile Ile
 305 310 315 320
 Ala Ile Ala Leu Gly Arg Lys Ile Ser Cys Val Thr Tyr Ser Val Ser
 325 330 335
 Arg Leu Ser Arg Met Leu Ser Pro Ile Pro Ala Val Ala Leu Thr Arg
 340 345 350
 Asp Arg Ala Ala Asp Ala Ala Arg Met Arg Lys Leu Leu Glu Lys Gly
 355 360 365

pctau2012001598-seql

Asp Leu Val Ile Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro Tyr Leu
370 375 380

Leu Arg Phe Ser Ala Leu Phe Ala Glu Leu Ser Asp Arg Ile Val Pro
385 390 395 400

Val Ala Met Asn Cys Lys Gln Gly Met Phe Asn Gly Thr Thr Val Arg
405 410 415

Gly Val Lys Phe Trp Asp Pro Tyr Phe Phe Phe Met Asn Pro Arg Pro
420 425 430

Ser Tyr Glu Ala Thr Phe Leu Asp Arg Leu Pro Glu Glu Met Thr Val
435 440 445

Asn Gly Gly Gly Lys Thr Pro Ile Glu Val Ala Asn Tyr Val Gln Lys
450 455 460

Val Ile Gly Gly Val Leu Gly Phe Glu Cys Thr Glu Leu Thr Arg Lys
465 470 475 480

Asp Lys Tyr Leu Leu Leu Gly Gly Asn Asp Gly Lys Val Glu Ser Ile
485 490 495

Lys Lys Thr Lys Asp Asp Lys Ser Ser
500 505

<210> 172
<211> 500
<212> PRT
<213> Arabidopsis thaliana

<400> 172

Met Ser Pro Glu Lys Lys Ser Gln Asn Phe Pro Pro Ile Thr Glu Cys
1 5 10 15

Arg Asp Gly Glu Tyr Asp Ser Ile Ala Ala Asp Leu Asp Gly Thr Leu
20 25 30

Leu Leu Ser Arg Ser Ser Phe Pro Tyr Phe Met Leu Val Ala Val Glu
35 40 45

Ala Gly Ser Leu Leu Arg Gly Leu Ile Leu Leu Leu Ser Leu Pro Phe
50 55 60

Val Ile Ile Ser Tyr Leu Phe Val Ser Glu Ser Leu Gly Ile Gln Ile
65 70 75 80

Leu Ile Phe Ile Ser Phe Ala Gly Leu Lys Ile Arg Asp Ile Glu Leu
85 90 95

pctau2012001598-seql

Val Ser Arg Ala Val Leu Pro Arg Phe Tyr Ala Ala Asp Val Arg Lys
100 105 110
Asp Ser Phe Glu Val Phe Asp Lys Cys Lys Arg Lys Val Val Val Thr
115 120 125
Ala Asn Pro Ile Val Met Val Glu Ala Phe Val Lys Asp Tyr Leu Gly
130 135 140
Gly Asp Lys Val Leu Gly Thr Glu Ile Glu Val Asn Pro Lys Thr Asn
145 150 155 160
Arg Ala Thr Gly Phe Val Lys Lys Pro Gly Val Leu Val Gly Asp Leu
165 170 175
Lys Arg Leu Ala Ile Leu Lys Glu Phe Gly Asn Glu Ser Pro Asp Leu
180 185 190
Gly Leu Gly Asp Arg Thr Ser Asp His Asp Phe Met Ser Leu Cys Lys
195 200 205
Lys Gly Tyr Met Val His Ala Thr Lys Ser Ala Thr Thr Ile Pro Lys
210 215 220
Glu Arg Leu Lys Asn Arg Ile Val Phe His Asp Gly Arg Leu Ala Gln
225 230 235 240
Arg Pro Thr Pro Leu Asn Ala Ile Ile Thr Tyr Leu Trp Leu Pro Phe
245 250 255
Gly Phe Ile Leu Ser Ile Ile Arg Val Tyr Phe Asn Leu Pro Leu Pro
260 265 270
Glu Arg Phe Val Arg Tyr Thr Tyr Glu Met Leu Gly Ile His Leu Thr
275 280 285
Ile Arg Gly His Arg Pro Pro Pro Pro Ser Pro Gly Thr Leu Gly Asn
290 295 300
Leu Tyr Val Leu Asn His Arg Thr Ala Leu Asp Pro Ile Ile Val Ala
305 310 315 320
Ile Ala Leu Gly Arg Lys Ile Cys Cys Val Thr Tyr Ser Val Ser Arg
325 330 335
Leu Ser Leu Met Leu Ser Pro Ile Pro Ala Val Ala Leu Thr Arg Asp
340 345 350
Arg Ala Thr Asp Ala Ala Asn Met Arg Lys Leu Leu Glu Lys Gly Asp
355 360 365

pctau2012001598-seql

Leu Val Ile Cys Pro Glu Gly Thr Thr Cys Arg Glu Glu Tyr Leu Leu
370 375 380

Arg Phe Ser Ala Leu Phe Ala Glu Leu Ser Asp Arg Ile Val Pro Val
385 390 395 400

Ala Met Asn Cys Lys Gln Gly Met Phe Asn Gly Thr Thr Val Arg Gly
405 410 415

Val Lys Phe Trp Asp Pro Tyr Phe Phe Phe Met Asn Pro Arg Pro Ser
420 425 430

Tyr Glu Ala Thr Phe Leu Asp Arg Leu Pro Glu Glu Met Thr Val Asn
435 440 445

Gly Gly Gly Lys Thr Pro Ile Glu Val Ala Asn Tyr Val Gln Lys Val
450 455 460

Ile Gly Ala Val Leu Gly Phe Glu Cys Thr Glu Leu Thr Arg Lys Asp
465 470 475 480

Lys Tyr Leu Leu Leu Gly Gly Asn Asp Gly Lys Val Glu Ser Ile Asn
485 490 495

Asn Thr Lys Lys
500

<210> 173
<211> 514
<212> PRT
<213> Physcomitrella patens

<400> 173

Met Gln Gln Leu Asp Pro Leu Thr Ser Phe His Cys Ser Glu Ser Phe
1 5 10 15

Glu Gly Val Glu Thr Cys Arg Val Glu Gly Arg Ala Asn Gln Ser Ile
20 25 30

Val Ser Asp Leu Asp Gly Thr Leu Leu Arg Ser Arg Ser Ser Phe Pro
35 40 45

Tyr Phe Met Leu Leu Ala Phe Glu Ala Gly Ser Pro Leu Arg Ser Val
50 55 60

Val Leu Leu Leu Ile Ser Pro Val Val Trp Leu Val Tyr Asn Phe Val
65 70 75 80

Ser Glu Ala Ala Gly Ile Lys Met Leu Ile Phe Val Ser Leu Ala Gly
85 90 95

pctau2012001598-seql

Leu Lys Val Ser Ala Val Glu Ser Val Ala Arg Gly Val Leu Pro Lys
 100 105 110
 Phe Tyr Leu Glu Asp Met His Ser Ile Ser Tyr Arg Val Phe Thr Ser
 115 120 125
 Cys Gly Lys Arg Tyr Val Val Thr Ala Asn Pro Arg Ile Ile Val Glu
 130 135 140
 Pro Phe Leu Lys Glu Tyr Leu Asp Val Glu Ala Val Met Gly Thr Glu
 145 150 155 160
 Leu Gln Ile Ser Ser Gly Gly Tyr Val Thr Gly Phe Val Thr Gly Pro
 165 170 175
 Gly Val Leu Val Gly Ala Val Lys Glu Arg Ala Val Lys Lys Tyr Phe
 180 185 190
 Gly Pro Asp Pro Pro Asp Val Gly Leu Gly Asp Arg Gln Ser Asp Phe
 195 200 205
 Leu Phe Met Asp Leu Cys Lys Glu Ser Tyr Ile Val His Ser Asp Lys
 210 215 220
 Asp Val Pro Ala Val Ser Lys Glu Asp Phe Leu Lys Pro Leu Ile Phe
 225 230 235 240
 His Asp Gly Arg Leu Val Cys Arg Pro Thr Pro Leu Met Ser Leu Ile
 245 250 255
 Val Met Leu Trp Cys Pro Ile Gly Phe Val Leu Ala Val Val Arg Met
 260 265 270
 Leu Ile Gly Ile Ala Leu Pro Met Trp Trp Ala Leu Pro Leu Glu Ala
 275 280 285
 Met Leu Gly Val Thr Val Arg Val Lys Gly Ile Pro Pro Thr Cys Pro
 290 295 300
 Leu Arg Asn Lys Lys Arg Gly Val Leu Phe Val Cys Ser His Arg Thr
 305 310 315 320
 Leu Leu Asp Pro Ile Phe Leu Ser Ile Ala Val Arg Arg Lys Val Thr
 325 330 335
 Ala Val Thr Tyr Ser Ile Ser Arg Leu Ser Glu Val Leu Ser Pro Ile
 340 345 350
 Arg Thr Val Arg Leu Thr Arg Asp Arg Asn Thr Asp Ala Asn Thr Met
 355 360 365

pctau2012001598-seql

Ala Ser Leu Leu Glu Glu Gly Asp Leu Val Val Cys Pro Glu Gly Thr
370 375 380

Thr Cys Arg Glu Pro Tyr Leu Leu Arg Phe Ser Ser Leu Phe Ala Glu
385 390 395 400

Leu Thr Asp Gln Ile Val Pro Val Thr Met Asn Ile Lys Ile Thr Met
405 410 415

Phe His Gly Thr Ser Ala Arg Gly Trp Lys Gly Met Asp Pro Phe Phe
420 425 430

Phe Phe Met Asn Pro Cys Pro Lys Tyr Glu Val Thr Phe Leu Asp Gln
435 440 445

Leu Pro His Gln Leu Thr Cys Asn Gly Gly Lys Thr Ser His Glu Val
450 455 460

Ala Asn Tyr Ile Gln Arg Val Leu Ala Asn Ser Leu Gly Phe Gln Cys
465 470 475 480

Thr Asn Leu Thr Arg Lys Asp Lys Tyr Thr Val Leu Ala Gly Asn Asp
485 490 495

Gly Ile Val Pro Asp His Thr Leu Arg Gln Arg Ala Gly Arg Met Phe
500 505 510

Gly Leu

<210> 174
<211> 501
<212> PRT
<213> Physcomitrella patens

<400> 174

Met Lys Arg Asp Pro Phe Asp Thr Ile Asp Lys Cys Ser Asp Lys Gly
1 5 10 15

Arg Ser Lys Gln Thr Val Val Ser Asp Leu Asp Gly Thr Leu Leu Arg
20 25 30

Ser Arg Ser Ser Phe Pro Tyr Phe Met Leu Val Ala Phe Glu Ala Gly
35 40 45

Gly Val Ala Arg Ala Phe Val Leu Leu Leu Cys Ser Pro Leu Cys Trp
50 55 60

Leu Leu Tyr His Cys Val Ser Glu Ser Ile Gly Ile Arg Leu Leu Ile
65 70 75 80

Phe Val Thr Phe Ala Gly Leu Lys Ile Gly Gly Ile Glu Ala Ile Ala
Page 210

pctau2012001598-seql

85

90

95

Arg Gly Thr Leu₁₀₀ Pro Lys Phe Tyr Ala₁₀₅ Glu Asp Val His Pro₁₁₀ Asp Thr

Trp Arg Ile₁₁₅ Phe Ser Ser Phe Gly₁₂₀ Glu Arg Tyr Ile Leu₁₂₅ Thr Ala Thr

Pro Arg₁₃₀ Ile Met Ala Glu Thr₁₃₅ Phe Ala Lys Thr Tyr₁₄₀ Leu Gly Val Asp

Gly₁₄₅ Val Leu Gly Thr Glu₁₅₀ Leu His Phe Thr Ser₁₅₅ Gly Gly Ile Ala Thr₁₆₀

Gly Leu Leu Met Asn₁₆₅ Pro Gly Val Leu Thr₁₇₀ Gly Arg Asn Lys Glu₁₇₅ Ile

Val Leu Arg Gln₁₈₀ Glu Phe Gln Gly Leu₁₈₅ Asn Leu Pro Asp Val₁₉₀ Gly Leu

Gly Asp Arg₁₉₅ Pro Ser Asp His Asn₂₀₀ Phe Met Ser Ile Cys₂₀₅ Lys Glu Gly

Tyr Ile₂₁₀ Val Pro Pro Ser Asn₂₁₅ Thr Ile Leu Ala Ala₂₂₀ Ser Lys Glu Ser

Leu₂₂₅ Met Asn Leu Leu Val₂₃₀ Phe His Asp Gly Arg₂₃₅ Leu Ile Gln Arg Pro₂₄₀

Thr Ala Gly Ile Ala₂₄₅ Leu Ile Ile Leu Leu₂₅₀ Trp Tyr Pro Ile Gly₂₅₅ Ala

Val Leu Ala Val₂₆₀ Leu Arg Ile Leu Ala₂₆₅ Gly Ile Leu Leu Pro₂₇₀ Phe His

Leu Leu Lys₂₇₅ Leu Val Tyr Lys Phe₂₈₀ Leu Gly Val Gly₂₈₅ Val Val Val Arg

Gly Thr₂₉₀ Pro Pro Pro Glu Pro₂₉₅ Lys Asp Gly Pro Gly₃₀₀ Arg Gly Gly Tyr

Leu₃₀₅ Tyr Val Cys Ser His₃₁₀ Arg Thr Leu Leu Asp₃₁₅ Pro Val Met Val Gly₃₂₀

Val Ala Leu Lys Arg₃₂₅ Arg Val Thr Ala Val₃₃₀ Thr Tyr Ser Ile Ser₃₃₅ Arg

Leu Ser Glu Val₃₄₀ Leu Ser Pro Ile Lys₃₄₅ Thr Val Ala Leu Lys₃₅₀ Arg Asn

Arg Glu Lys Asp Ala Ala Lys Ile Arg Ser Leu Leu Arg Glu Gly Asp

355

360

365

Leu Ala Ile Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro Tyr Leu Leu
 370 375 380

Arg Phe Ser Ala Leu Phe Ala Glu Leu Ser Asn Gln Leu Val Pro Val
 385 390 395 400

Ala Met Asn Thr Arg Met Ser Met Phe His Gly Thr Thr Ala Gln Gly
 405 410 415

Trp Lys Cys Leu Asp Pro Phe Tyr Phe Phe Met Asn Pro Asn Pro Ile
 420 425 430

Tyr Glu Val Thr Phe Leu Asn Glu Leu Pro Val Glu Leu Thr Cys Ala
 435 440 445

Gly Gly Lys Ser Ser Tyr Glu Val Ala Asn His Ile Gln Gln Leu Leu
 450 455 460

Ser Gln Thr Leu Gly Phe Glu Cys Thr Asn Tyr Thr Arg Lys Asp Lys
 465 470 475 480

Tyr Gly Val Leu Cys Gly Asn Asp Gly Ser Ile Pro Leu Lys Ser Gln
 485 490 495

Asp Ser Phe Gly Ser
 500

<210> 175
 <211> 517
 <212> PRT
 <213> Physcomitrella patens

<400> 175

Met Glu Thr Gln Glu Pro Asp Pro Val Thr Ser Phe His Cys Ser Glu
 1 5 10 15

Asn Tyr Asp Glu Val Glu Asn Cys Arg Val Glu Gly Arg Ala Asn Gln
 20 25 30

Ser Ile Val Ser Asp Leu Asp Gly Thr Leu Leu Arg Ser Arg Ser Ser
 35 40 45

Phe Pro Tyr Phe Met Leu Leu Ala Phe Glu Ala Gly Ser Pro Leu Arg
 50 55 60

Phe Ile Ile Leu Leu Leu Val Ser Pro Ile Val Trp Phe Val Tyr Asn
 65 70 75 80

Phe Ile Ser Glu Val Val Gly Ile Lys Met Leu Ile Phe Ile Ser Leu
 85 90 95

pctau2012001598-seql

Ala Gly Leu Lys Val Ser Ala Ile Glu Ser Val Ala Arg Gly Val Leu
100 105 110

Pro Lys Phe Phe Leu Glu Asp Met His Ser Ile Ser Tyr Arg Val Phe
115 120 125

Ile Ser Cys Glu Lys Arg Tyr Val Val Thr Ala Asn Pro Arg Ile Ile
130 135 140

Val Glu Pro Phe Leu Lys Glu Tyr Leu Asp Val Asp Ala Val Met Gly
145 150 155 160

Thr Glu Leu Gln Ile Ser Ser Gly Gly Tyr Ala Thr Gly Phe Val Thr
165 170 175

Gly Pro Gly Val Leu Val Gly Thr Ala Lys Gln Arg Ala Val Lys Lys
180 185 190

Tyr Phe Gly Ser Asp Gln Pro Asp Leu Gly Leu Gly Asp Arg Thr Ser
195 200 205

Asp Phe Ala Phe Met Asp Leu Cys Lys Glu Ala Tyr Ile Val Pro Ser
210 215 220

Tyr Lys Glu Val Pro Ser Val Thr Lys Lys Asp Tyr Leu Lys Pro Leu
225 230 235 240

Ile Phe His Asp Gly Arg Leu Val Cys Arg Pro Thr Pro Leu Met Ala
245 250 255

Leu Ala Val Thr Leu Trp Cys Pro Ile Gly Phe Val Leu Ala Ile Ile
260 265 270

Arg Met Phe Ile Gly Ile Met Leu Pro Met Trp Trp Ala Leu Pro Leu
275 280 285

Glu Ala Met Leu Gly Val Thr Val Arg Ile Lys Gly Ile Pro Pro Thr
290 295 300

Cys Pro Leu Arg Ser Lys Lys Arg Gly Val Leu Phe Val Cys Ser His
305 310 315 320

Arg Thr Leu Leu Asp Pro Ile Phe Leu Ser Ile Ala Val Arg Arg Lys
325 330 335

Val Thr Ala Val Thr Tyr Ser Ile Ser Arg Leu Ser Glu Val Leu Ala
340 345 350

Pro Ile Arg Thr Val Arg Leu Thr Arg Asp Arg Asn Thr Asp Ala Ser
355 360 365

pctau2012001598-seql

Thr Met Ala Ser Leu Leu Asp Glu Gly Asp Leu Val Val Cys Pro Glu
370 375 380

Gly Thr Thr Cys Arg Glu Pro Tyr Leu Leu Arg Phe Ser Ser Leu Phe
385 390 395 400

Ala Glu Leu Thr Asp Gln Ile Val Pro Val Thr Met Asn Ile Lys Ile
405 410 415

Ser Met Phe His Gly Thr Ser Ala Arg Gly Trp Lys Gly Met Asp Pro
420 425 430

Phe Phe Phe Phe Met Asn Pro Arg Pro Lys Tyr Glu Val Thr Phe Leu
435 440 445

Asp Gln Leu Pro His Gln Leu Thr Cys Asn Gly Gly Lys Thr Ser His
450 455 460

Glu Val Ala Asn Tyr Ile Gln Arg Val Leu Ala Asn Ser Leu Gly Phe
465 470 475 480

Gln Cys Thr Asn Leu Thr Arg Lys Asp Lys Tyr Arg Val Leu Ala Gly
485 490 495

Asn Asp Gly Ile Val Pro Asp Arg Ser Leu Arg Gln Arg Ala Ser Arg
500 505 510

Ile Phe Gly Leu Gln
515

<210> 176
<211> 504
<212> PRT
<213> Physcomitrella patens

<400> 176

Met Glu Val Thr Lys Leu Lys Glu Ala Gly Asn Tyr Ser Phe Ala Glu
1 5 10 15

Val Lys Ser Cys Ser Ser Gly Asn Asp Arg Arg Ser Gln Thr Val Val
20 25 30

Ala Asp Leu Asp Gly Thr Leu Ile Arg Gly Arg Ser Ala Phe Pro Tyr
35 40 45

Phe Phe Leu Val Ala Phe Glu Ala Gly Ser Tyr Phe Arg Ala Leu Leu
50 55 60

Leu Ser Leu Met Ala Pro Leu Ile Gly Phe Leu Tyr Tyr Phe Val Asp
65 70 75 80

pctau2012001598-seql

Glu Ala Ser Gly Ile His Leu Met Ile Phe Leu Ser Phe Ala Gly Leu
85 90 95
Lys Ile Lys Asp Ile Glu Gly Val Ala Arg Ala Val Leu Thr Lys Phe
100 105 110
Tyr Ala Asp Asp Leu His Pro Glu Thr Trp Arg Val Phe Ser Ser Phe
115 120 125
Gly Lys Arg Val Val Leu Thr Ala Asn Pro Arg Val Met Val Glu Pro
130 135 140
Phe Leu Lys Glu Val Leu Gly Ala Asp Glu Val Met Gly Thr Glu Ile
145 150 155 160
Glu Val Asn Lys Lys Gly Arg Ala Thr Gly Arg Leu Leu Lys Pro Gly
165 170 175
Val Leu Val Gly Thr His Lys Glu Glu Ala Leu Lys Arg Cys Asn Ile
180 185 190
Gly Gly Glu Arg Pro His Val Gly Leu Gly Asp Arg Val Thr Asp Phe
195 200 205
Pro Phe Met Ala Tyr Cys Lys Glu Gly Tyr Val Val Pro Lys Thr Lys
210 215 220
Val Pro Ala Val Lys Lys Gln Asp Met Pro Lys Gln Leu Ile Phe His
225 230 235 240
Asp Gly Arg Leu Val Gln Leu Pro Thr Pro Arg Asn Ala Phe Lys Val
245 250 255
Leu Ala Trp Met Pro Ile Gly Leu Leu Leu Ser Ile Val Arg Val Thr
260 265 270
Met Gly Val Trp Val Pro Ile Arg Leu Met Pro Leu Cys Tyr Lys Leu
275 280 285
Thr Gly Ile Asn Leu Val Val Lys Gly Asn Ile Pro Glu Lys Pro Lys
290 295 300
Asn Gly Glu Pro Gly Arg Leu Phe Val Cys Asn His Arg Thr Leu Leu
305 310 315 320
Asp Pro Val Ile Ile Ala Leu Ala Leu Gly Arg Pro Val Pro Ala Val
325 330 335
Thr Tyr Ser Ile Ser Lys Val Ser Glu Phe Leu Ser Pro Met Pro Thr
340 345 350

pctau2012001598-seql

Ile Gly Leu Cys Arg Asp Arg Glu Lys Asp Ser Ala Asn Met Arg Arg
355 360 365

Val Leu Lys Glu Gly Glu Leu Thr Leu Cys Pro Glu Gly Thr Thr Cys
370 375 380

Arg Glu Pro Phe Leu Leu Arg Phe Ser Ala Leu Phe Ala Glu Leu Ser
385 390 395 400

Asp Arg Ile Val Pro Val Ala Val Lys Ile Asn Met Asn Met Phe His
405 410 415

Gly Thr Thr Ala Arg Gly Asn Lys Ala Met Asp Pro Phe Phe Ala Tyr
420 425 430

Met Asn Pro Arg Pro Thr Ile Glu Leu Lys Phe Leu Asp Glu Ile Pro
435 440 445

Lys Asn Met Thr Cys Gly Ser Gly Lys Ser Ser Ile Glu Val Ala Asn
450 455 460

Tyr Ile Gln His Leu Leu Ala Gly Glu Leu Gly Tyr Asp Cys Thr Asp
465 470 475 480

Phe Thr Arg Lys Asp Lys Tyr Arg Leu Leu Ala Gly Asn Asp Gly Ile
485 490 495

Val Pro Val Lys Glu Lys Thr Arg
500

<210> 177
<211> 529
<212> PRT
<213> Physcomitrella patens

<400> 177

Met Glu Ile Pro Gly Val Phe Asp Ala Asn Leu Phe Cys Ala Glu Ser
1 5 10 15

Phe Pro Glu Val Glu Thr Cys Lys Val Glu Gly Arg Glu Lys Gln Thr
20 25 30

Ile Ile Ser Asp Leu Asp Gly Thr Leu Leu Arg Ser Arg Ser Ser Phe
35 40 45

Pro Tyr Phe Met Leu Ile Ala Phe Asp Ala Gly Ser Pro Leu Arg Phe
50 55 60

Ile Val Leu Leu Leu Ala Ser Pro Ile Ile Trp Leu Val Tyr Asn Phe
65 70 75 80

pctau2012001598-seql

Phe Ser Glu Ala Ala Gly Thr Lys Leu Leu Ile Phe Ile Ala Cys Ala
85 90 95

Gly Leu Lys Ala Ser Asp Ile Glu Ser Val Ala Arg Gly Val Leu Pro
100 105 110

Lys Phe Tyr Leu Glu Asp Met His Ser Val Ser Tyr Ser Ile Phe Val
115 120 125

Ser Cys Gly Lys Arg Tyr Val Val Thr Ala Asn Pro Arg Ile Met Val
130 135 140

Glu Ser Phe Leu Lys Glu Tyr Met Gly Val Glu Ala Val Ile Gly Thr
145 150 155 160

Glu Leu His Ile Thr Lys Ser Gly Tyr Ala Thr Gly Leu Val Met Gly
165 170 175

Ala Gly Val Ile Val Gly Ala Asn Lys Thr Arg Ala Val Lys Lys Tyr
180 185 190

Phe Gly Asp Asp Leu Pro Asp Ile Ala Leu Gly Asp Arg Ala Ser Asp
195 200 205

Phe Pro Phe Met Ala Leu Cys Lys Glu Ala Tyr Leu Val Pro Ser Phe
210 215 220

Lys Pro Val Glu Pro Val Ala Lys Ala Asp Tyr Leu Lys Pro Leu Ile
225 230 235 240

Phe His Asp Gly Arg Leu Val Cys Arg Pro Thr Pro Leu Met Ser Leu
245 250 255

Val Ile Ile Leu Trp Ser Pro Ile Gly Leu Ile Leu Ala Leu Ile Arg
260 265 270

Met Val Phe Gly Thr Val Leu Pro Met Trp Tyr Ala Leu Pro Leu Glu
275 280 285

Ala Met Leu Gly Val Ser Ile Arg Val Lys Gly Ile Pro Pro Ser Cys
290 295 300

Pro Gln Thr Ser Gly Arg Arg Gly Val Leu Phe Val Cys Ser His Arg
305 310 315 320

Thr Leu Leu Asp Pro Ile Phe Leu Ser Ile Ala Cys Arg Arg Gln Val
325 330 335

Thr Ala Val Thr Tyr Ser Ile Ser Lys Val Ser Glu Ile Leu Ser Pro
340 345 350

pctau2012001598-seql

Ile Pro Thr Val Arg Leu Thr Arg Cys Arg Lys Thr Asp Ala Glu Thr
355 360 365

Met Thr Lys Leu Leu Asn Glu Gly Asp Leu Ala Val Cys Pro Glu Gly
370 375 380

Thr Thr Cys Arg Glu Pro Tyr Leu Leu Arg Phe Ser Ser Leu Phe Ala
385 390 395 400

Glu Leu Ala Asp Gln Ile Ile Pro Val Thr Met Asn Val Lys Thr Thr
405 410 415

Met Phe His Gly Thr Thr Ala Arg Gly Trp Lys Ala Leu Asp Pro Phe
420 425 430

Ile Phe Leu Met Asn Pro Ser Pro Arg Tyr Glu Leu Glu Phe Leu Asp
435 440 445

Gln Leu Pro His Glu Met Thr Cys Ser Gly Gly Lys Ser Ser His Glu
450 455 460

Val Ala Asn Tyr Ile Gln Arg Val Leu Ala Ala Ser Leu Gly Phe Lys
465 470 475 480

Cys Thr Asn Leu Thr Arg Arg Asp Lys Tyr Arg Ile Leu Ala Gly Asn
485 490 495

Asp Gly Val Val Pro Ala Ser His Phe Leu Lys Gln Arg Ser Arg Ile
500 505 510

Phe Ala Cys Leu Pro Ser Arg Asn Ser Glu Arg Gly Gln Arg Ser Asp
515 520 525

Pro

<210> 178

<211> 503

<212> PRT

<213> Vitis vinifera

<400> 178

Met Ala Val Thr Thr Phe Pro Thr Val Asp Gln Cys Gln Ser Ile Gly
1 5 10 15

Arg Glu Lys His Thr Val Val Ala Asp Met Asp Gly Thr Leu Leu Ile
20 25 30

Gly Arg Ser Ser Phe Pro Tyr Phe Ala Leu Val Ala Phe Glu Gly Gly
35 40 45

Gly Val Leu Arg Leu Leu Phe Leu Leu Leu Ala Ser Pro Leu Ala Gly
Page 218

pctau2012001598-seql

50

55

60

Leu Leu Tyr Tyr Phe Val Ser Glu Ser Ala Gly Ile Gln Ile Leu Ile
 65 70 75 80

Phe Val Thr Phe Ala Gly Met Lys Val Ser Asp Ile Glu Ser Val Ala
 85 90 95

Arg Ala Val Leu Pro Lys Phe Tyr Leu Ser Asp Leu His Pro Glu Ser
 100 105 110

Trp Arg Val Phe Ser Ser Cys Gly Lys Arg Cys Val Leu Thr Ala Asn
 115 120 125

Pro Arg Ile Met Val Glu Ala Phe Leu Lys Asp Phe Leu Gly Ala Asp
 130 135 140

Leu Val Leu Gly Thr Glu Ile Ala Thr Tyr Lys Gly Arg Ala Thr Gly
 145 150 155 160

Phe Val Cys Asp Pro Gly Val Leu Val Gly Lys Asn Lys Ala Tyr Ala
 165 170 175

Leu Asn Lys Thr Phe Gly Glu Thr Gln Pro Glu Ile Gly Leu Gly Asp
 180 185 190

Arg His Thr Asp Phe Pro Phe Met Ala Ser Cys Lys Glu Gly Tyr Ile
 195 200 205

Val Ala Ala Lys Pro Glu Val Lys Ala Val Thr Ser Asp Lys Leu Pro
 210 215 220

Lys Pro Ile Ile Phe His Asp Ser Arg Leu Val Gln Lys Pro Thr Pro
 225 230 235 240

Leu Thr Ala Leu Leu Thr Leu Leu Trp Ile Pro Ile Gly Phe Ala Leu
 245 250 255

Ala Cys Leu Arg Ile Ala Ala Gly Ser Leu Leu Pro Met Arg Ile Val
 260 265 270

Tyr His Ala Phe Trp Ala Leu Gly Val Arg Val Ser Val Lys Gly Ser
 275 280 285

Pro Pro Pro Val Ala Lys Lys Ser Ile Gly Gln Ser Gly Val Leu Phe
 290 295 300

Ile Cys Ser His Arg Thr Leu Leu Asp Pro Ile Phe Leu Ser Ala Ala
 305 310 315 320

Leu Gly Arg Pro Ile Pro Ala Val Thr Tyr Ser Val Ser Arg Leu Ser
 Page 219

pctau2012001598-seql

325

330

335

Glu Phe Ile Ser₃₄₀ Pro Ile Lys Thr Ile₃₄₅ Arg Leu Ser Arg Asp₃₅₀ Arg Val

Lys Asp Ala₃₅₅ Ala Met Ile Lys Lys₃₆₀ Leu Leu Glu Glu Gly₃₆₅ Asp Leu Ala

Ile Cys₃₇₀ Pro Glu Gly Thr Thr₃₇₅ Cys Arg Glu Pro Phe₃₈₀ Leu Leu Arg Phe

Ser₃₈₅ Ala Leu Phe Ala Glu₃₉₀ Leu Thr Asp Gln Leu₃₉₅ Val Pro Val Ala Met₄₀₀

Met Asn Arg Met Ser₄₀₅ Met Phe His Gly Thr₄₁₀ Thr Ala Arg Gly Trp₄₁₅ Lys

Gly Met Asp Pro₄₂₀ Phe Phe Phe Phe Met₄₂₅ Asn Pro Ser Pro Ala₄₃₀ Tyr Glu

Val Thr Phe₄₃₅ Leu Asn Lys Leu Pro₄₄₀ Leu Glu Leu Thr Cys₄₄₅ Ser Ser Gly

Lys Ser₄₅₀ Ser His Glu Val Ala₄₅₅ Asn Tyr Met Gln Arg₄₆₀ Val Ile Ala Ala

Thr₄₆₅ Leu Ser Tyr Glu Cys₄₇₀ Thr Ser Phe Thr Arg₄₇₅ Lys Asp Lys Tyr Arg₄₈₀

Ala Leu Ala Gly Asn₄₈₅ Asp Gly Thr Val Pro₄₉₀ Lys Lys Pro Leu₄₉₅ Leu Tyr

Pro Asp Lys Val₅₀₀ Met Gly Cys

<210> 179

<211> 501

<212> PRT

<213> Vitis vinifera

<400> 179

Met Ser Pro Pro₅ Pro Lys Arg Ala Arg Lys₁₀ Phe Pro Ser Ile Thr Thr₁₅

Tyr Ala Gly Gly₂₀ Asp His Arg Ser Ile₂₅ Ala Ala Asp Leu Asp₃₀ Gly Thr

Leu Leu Val₃₅ Ser Arg Ser Ser Phe₄₀ Pro Tyr Phe Met Leu₄₅ Val Ala Val

Glu Ala Gly Ser Leu Leu Arg₅₅ Gly Leu Phe Leu Leu₆₀ Leu Ser Leu Pro

pctau2012001598-seql

Ile Val Ile Val Ala Tyr Leu Phe Ile Ser Glu Glu Ile Gly Ile Gln
65 70 75 80

Ile Leu Ile Phe Ile Ser Phe Ser Gly Leu Lys Ile Arg Asp Ile Glu
85 90 95

Leu Ala Ser Gln Ala Val Leu Pro Arg Phe Tyr Ala Asn Asp Val Arg
100 105 110

Arg Glu Ser Trp Glu Val Phe Glu Lys Cys Glu Arg Lys Val Val Val
115 120 125

Thr Ala Asn Pro Thr Leu Met Val Glu Pro Phe Val Arg Asp Phe Leu
130 135 140

Gly Gly Thr Lys Val Leu Gly Thr Glu Ile Glu Val Asn Pro Lys Thr
145 150 155 160

Lys Lys Ala Thr Gly Phe Val Lys Lys Pro Gly Val Leu Val Gly Asp
165 170 175

Arg Lys Arg Leu Ala Leu Leu Lys Glu Phe Gly Asp Glu Leu Pro Asp
180 185 190

Ile Gly Ile Gly Asp Arg Glu Ser Asp His Asp Phe Met Ser Ile Cys
195 200 205

Lys Glu Gly Tyr Met Val Leu Pro Ser Lys Ser Ala Thr Pro Val Pro
210 215 220

Pro Asn Arg Leu Lys Thr Pro Ile Ile Phe His Asp Gly Arg Phe Val
225 230 235 240

Gln Pro Pro Thr Pro Leu Thr Ala Leu Ile Ile Tyr Leu Trp Leu Pro
245 250 255

Phe Gly Phe Ala Leu Ser Ile Phe Arg Val Tyr Phe Asn Leu Pro Leu
260 265 270

Pro Glu Arg Ile Val Arg Tyr Thr Tyr Pro Met Leu Gly Ile Asn Leu
275 280 285

Val Ile Arg Gly Asn Pro Pro Pro Pro Ser Pro Gly Ser Pro Gly
290 295 300

Asn Leu Tyr Val Cys Asn His Arg Thr Ala Leu Asp Pro Ile Val Ile
305 310 315 320

Ala Ile Ala Leu Arg Arg Lys Val Ser Cys Val Thr Tyr Ser Val Ser
325 330 335

pctau2012001598-seql

Arg Leu Ser Arg Phe Leu Ser Pro Ile Pro Ala Val Ala Leu Thr Arg
340 345 350

Asp Arg Ala Ala Asp Ala Ala Arg Ile Ser Ser Ile Leu Gln Lys Gly
355 360 365

Asp Leu Val Val Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro Tyr Leu
370 375 380

Leu Arg Phe Ser Ala Leu Phe Ala Glu Leu Ser Asp Arg Ile Val Pro
385 390 395 400

Val Ala Val Asn Val Lys Gln Asn Met Phe His Gly Thr Thr Val Arg
405 410 415

Gly Val Lys Phe Trp Asp Ala Tyr Phe Tyr Phe Met Asn Pro Arg Pro
420 425 430

Thr Tyr Glu Ile Thr Phe Leu Asp Arg Leu Pro Glu Glu Met Thr Cys
435 440 445

Lys Ala Gly Gly Lys Ser Ala Ile Glu Val Ala Asn His Val Gln Lys
450 455 460

Val Leu Gly Gly Val Leu Gly Phe Glu Cys Thr Gly Leu Thr Arg Lys
465 470 475 480

Asp Lys Tyr Met Leu Leu Gly Gly Asn Asp Gly Lys Val Glu Ser Met
485 490 495

Tyr Asn Ala Lys Lys
500

<210> 180
<211> 510
<212> PRT
<213> Vitis vinifera

<400> 180

Met Thr Lys Ser Asn Leu His Ser Pro Pro Thr Ser Leu Thr Ile Asp
1 5 10 15

Gln Cys Thr Ser Ile Gly Arg Glu Lys Glu Thr Val Val Ala Asp Met
20 25 30

Asp Gly Thr Leu Leu Arg Gly Arg Ser Ser Phe Pro Tyr Phe Ala Leu
35 40 45

Val Ala Phe Glu Val Gly Gly Ile Phe Arg Leu Leu Leu Leu Leu
50 55 60

pctau2012001598-seql

Cys Ser Pro Leu Ala Gly Val Leu Tyr Tyr Phe Ile Ser Glu Pro Ala
65 70 75 80
Gly Ile Gln Val Leu Ile Phe Ala Thr Phe Val Gly Met Lys Val Ser
85 90 95
Asp Ile Glu Ser Val Ala Arg Ala Val Leu Pro Lys Phe Tyr Ala Ser
100 105 110
Asp Leu His Pro Glu Ala Trp Arg Val Phe Ser Ser Cys Gly Lys Arg
115 120 125
Cys Val Leu Thr Ala Asn Pro Arg Ile Met Val Glu Ala Phe Leu Lys
130 135 140
Asp Phe Leu Gly Ala Asp Met Val Leu Gly Thr Glu Val Asp Thr Tyr
145 150 155 160
Lys Gly Arg Ala Thr Gly Leu Val Arg Gly Gly Gly Val Leu Val Gly
165 170 175
Lys Arg Lys Ala Glu Ala Leu Arg Asn Ala Phe Ser Glu Thr Ser Leu
180 185 190
Pro Glu Ile Gly Leu Gly Asp Arg Glu Thr Asp Tyr Pro Phe Met Ser
195 200 205
Leu Cys Lys Glu Ser Tyr Ile Val Pro Ala Lys Pro Glu Val Glu Ala
210 215 220
Val Ser His Asp Lys Leu Pro Lys Pro Ile Val Phe His Asp Gly Arg
225 230 235 240
Leu Val Gln Lys Pro Thr Pro Phe Met Ala Leu Leu Thr Ile Leu Trp
245 250 255
Ile Pro Val Gly Phe Leu Leu Ala Cys Leu Arg Ile Ala Ala Gly Ala
260 265 270
Leu Leu Pro Met Pro Leu Val Tyr Tyr Ala Phe Trp Ala Leu Gly Val
275 280 285
Arg Val Tyr Ile Lys Gly Asn Pro Pro Pro Pro Ala Lys Lys Ser Ile
290 295 300
Gly Gln Ser Gly Val Leu Phe Ile Cys Ser His Arg Thr Leu Leu Asp
305 310 315 320
Pro Ile Phe Leu Ser Ala Ala Leu Gly Arg Pro Ile Pro Ala Val Thr
325 330 335

pctau2012001598-seql

Tyr Ser Leu Ser Arg Leu Ser Glu Phe Ile Ser Pro Ile Lys Thr Val
340 345 350

Arg Leu Ser Arg Asp Arg Ala Thr Asp Ala Ala Met Ile Lys Lys Leu
355 360 365

Leu Glu Ala Gly Asp Leu Ala Ile Cys Pro Glu Gly Thr Thr Cys Arg
370 375 380

Glu Pro Phe Leu Leu Arg Phe Ser Ala Leu Phe Ser Glu Leu Thr Asp
385 390 395 400

Glu Leu Val Pro Val Ala Met Ala Asn Arg Met Thr Met Phe His Gly
405 410 415

Thr Thr Ala Arg Gly Trp Lys Gly Met Asp Pro Phe Tyr Tyr Phe Met
420 425 430

Asn Pro Ser Pro Ala Tyr Glu Val Asn Phe Leu Asn Lys Leu Pro His
435 440 445

Glu Leu Thr Cys Gly Ala Gly Lys Ser Ser His Asp Val Ala Asn Tyr
450 455 460

Ile Gln Arg Met Ile Ala Ala Ser Leu Ser Tyr Glu Cys Thr Asn Phe
465 470 475 480

Thr Arg Lys Asp Lys Tyr Arg Ala Leu Ala Gly Asn Asp Gly Thr Val
485 490 495

Val Glu Lys Pro Lys Val Asp Ala Asn Lys Val Met Gly Cys
500 505 510

<210> 181
<211> 499
<212> PRT
<213> Vitis vinifera

<400> 181

Met Gly Ala Gln Arg His Phe Gln Pro Ile Ser Lys Cys Ser Ser Asp
1 5 10 15

Gly Arg Ala Asn Gln Thr Val Ala Ala Asp Leu Asp Gly Thr Leu Leu
20 25 30

Val Ser Arg Ser Ala Phe Pro Tyr Phe Met Leu Val Ala Leu Glu Ala
35 40 45

Gly Ser Leu Val Arg Ala Ile Leu Leu Leu Ala Ser Val Pro Leu Val
50 55 60

pctau2012001598-seql

Tyr Phe Ile Tyr Leu Phe Ile Ser Glu Ser Leu Ala Ile Lys Ile Phe
65 70 75 80
Ile Tyr Met Ala Phe Ala Gly Leu Lys Val Arg Glu Ile Glu Leu Val
85 90 95
Ser Arg Ser Val Leu Pro Lys Phe Tyr Ala Glu Asp Val His Pro Glu
100 105 110
Ser Trp Arg Val Phe Asn Ser Phe Gly Lys Arg Tyr Ile Ile Thr Ala
115 120 125
Asn Pro Arg Ile Met Val Glu Pro Phe Val Lys Thr Tyr Met Gly Ala
130 135 140
Asp Lys Val Leu Gly Thr Glu Leu Glu Val Ser Asn Ser Gly Arg Ala
145 150 155 160
Thr Gly Phe Val Lys Lys Pro Gly Val Leu Val Gly Glu His Lys Lys
165 170 175
Ala Ala Val Gln Gln Glu Phe Gly Met Asn Leu Pro Asp Val Gly Leu
180 185 190
Gly Asp Arg Asp Thr Asp His Asp Phe Met Ser Leu Cys Lys Glu Gly
195 200 205
Tyr Met Val Pro Arg Thr Lys Cys Glu Pro Leu Pro Arg Asn Lys Leu
210 215 220
Leu Ser Pro Val Ile Phe His Glu Gly Arg Leu Val Gln Arg Pro Thr
225 230 235 240
Pro Val Ala Ala Leu Leu Thr Phe Leu Trp Met Pro Ile Gly Ile Ile
245 250 255
Leu Ser Ile Leu Arg Val Tyr Leu Asn Ile Pro Leu Pro Glu Arg Ile
260 265 270
Val Arg Tyr Asn Tyr Lys Ile Leu Gly Ile Lys Leu Ile Val Lys Gly
275 280 285
Thr Pro Pro Pro Pro Pro Lys Gly Gly Gln Ser Gly Val Leu Phe Val
290 295 300
Cys Asn His Arg Thr Val Leu Asp Pro Val Val Thr Ala Val Ala Leu
305 310 315 320
Gly Arg Lys Ile Ser Cys Val Thr Tyr Ser Ile Ser Lys Phe Ser Glu
325 330 335

pctau2012001598-seql

Leu Ile Ser Pro Ile Lys Ala Val Ala Leu Ser Arg Glu Arg Glu Lys
340 345 350

Asp Ala Ala Asn Ile Lys Arg Leu Leu Glu Glu Gly Asp Leu Val Ile
355 360 365

Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro Phe Leu Leu Arg Phe Ser
370 375 380

Ala Leu Phe Ala Glu Leu Thr Asp Arg Ile Val Pro Val Ala Ile Asn
385 390 395 400

Thr Lys Gln Ser Val Phe Tyr Gly Thr Ser Thr Arg Gly Tyr Lys Leu
405 410 415

Leu Asp Pro Tyr Phe Val Phe Met Asn Pro Met Pro Thr Tyr Glu Ile
420 425 430

Thr Phe Leu Asn Gln Leu Pro Ala Glu Leu Thr Cys Lys Gly Gly Lys
435 440 445

Ser Pro Ile Glu Val Ala Asn Tyr Ile Gln Arg Val Leu Ala Gly Thr
450 455 460

Leu Gly Phe Glu Cys Thr Asn Leu Thr Arg Lys Asp Lys Tyr Ala Thr
465 470 475 480

Leu Ala Gly Thr Asp Gly Ser Val Pro Phe Lys Asn Glu Lys Glu Lys
485 490 495

Glu Lys Ala

<210> 182
<211> 497
<212> PRT
<213> Populus trichocarpa

<400> 182

Met Val Met Gly Ala His Arg His Phe Glu Pro Ile Ala Lys Cys Ser
1 5 10 15

Thr Glu Gly Arg Ser Asn Gln Thr Val Ala Ala Asp Leu Asp Gly Thr
20 25 30

Leu Leu Val Ser Arg Ser Ala Phe Pro Tyr Phe Leu Leu Val Ala Ile
35 40 45

Glu Ala Gly Ser Leu Leu Arg Gly Leu Ile Leu Leu Ala Ser Val Pro
50 55 60

Phe Val Tyr Phe Thr Tyr Leu Phe Ile Ser Glu Ala Met Ala Ile Lys
Page 226

pctau2012001598-seql

```

65              70              75              80
Thr Phe Ile Phe Ile Ala Phe Ala Gly Leu Lys Ile Arg Asp Ile Glu
      85              90              95
Leu Val Ser Arg Ser Val Leu Pro Lys Phe Tyr Ala Asp Asp Val His
      100             105             110
Pro Glu Thr Trp Arg Val Phe Asn Ser Phe Gly Lys Arg Tyr Ile Val
      115             120             125
Thr Ala Asn Pro Arg Ile Met Val Glu Pro Phe Val Lys Thr Phe Leu
      130             135             140
Gly Ala Asp Lys Val Leu Gly Thr Glu Leu Glu Ala Thr Lys Ser Gly
      145             150             155             160
Arg Ala Thr Gly Phe Ile Lys Lys Pro Gly Val Leu Val Gly Asp His
      165             170             175
Lys Arg Asp Ala Leu Leu Lys Glu Phe Gly Thr Asn Leu Pro Asp Leu
      180             185             190
Gly Leu Gly Asp Arg Glu Thr Asp His Asp Phe Met Ser Ile Cys Lys
      195             200             205
Glu Gly Tyr Met Val Pro Gly Thr Lys Cys Glu Pro Leu Ala Arg Asn
      210             215             220
Lys Leu Leu Ser Pro Val Ile Phe His Asp Gly Arg Leu Val Gln Arg
      225             230             235             240
Pro Thr Pro Leu Ala Ala Leu Leu Thr Phe Leu Trp Met Pro Ile Gly
      245             250             255
Ile Ile Leu Ser Ile Leu Arg Val Tyr Leu Asn Ile Pro Leu Pro Glu
      260             265             270
Arg Leu Ala Trp Tyr Thr Tyr Lys Leu Leu Gly Ile Arg Val Ile Val
      275             280             285
Lys Gly Thr Pro Pro Pro Pro Pro Gly Lys Gly His Ser Gly Val Leu
      290             295             300
Phe Val Cys Asn His Arg Thr Val Leu Asp Pro Val Val Thr Ala Val
      305             310             315             320
Ala Leu Gly Arg Lys Ile Ser Cys Val Thr Tyr Ser Ile Ser Lys Phe
      325             330             335
Thr Glu Ile Ile Ser Pro Ile Lys Ala Val Ala Leu Ser Arg Glu Arg

```

pctau2012001598-seql

340

345

350

Asp Lys Asp Ala Ala Asn Ile Lys Arg Leu Leu Glu Glu Gly Asp Leu
 355 360 365

Val Ile Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro Phe Leu Leu Arg
 370 375 380

Phe Ser Ala Leu Phe Ala Glu Leu Thr Asp Arg Ile Val Pro Val Ala
 385 390 395 400

Ile Asn Thr Lys Gln Ser Val Phe His Gly Thr Thr Val Arg Gly His
 405 410 415

Lys Leu Leu Asp Pro Tyr Phe Val Phe Met Asn Pro Met Pro Thr Tyr
 420 425 430

Glu Val Thr Phe Leu Asn Gln Leu Pro Lys Glu Leu Thr Cys Lys Gly
 435 440 445

Gly Lys Ser Ala Ile Glu Val Ala Asn Tyr Ile Gln Arg Val Leu Ala
 450 455 460

Gly Thr Leu Gly Phe Glu Cys Thr Asn Leu Thr Arg Lys Asp Lys Tyr
 465 470 475 480

Ala Lys Leu Ala Gly Thr Asp Gly Arg Val Leu Ser Lys Lys Glu Asn
 485 490 495

Ala

<210> 183

<211> 501

<212> PRT

<213> Populus trichocarpa

<400> 183

Met Val Ser Phe Pro Thr Val Asp Lys Cys Ala Ser Ile Gly Arg Glu
 1 5 10 15

Lys His Ser Val Val Ala Asp Met Asp Gly Thr Leu Leu Arg Gly Arg
 20 25 30

Ser Ser Phe Pro Tyr Phe Ala Leu Leu Ala Phe Glu Ala Gly Gly Ile
 35 40 45

Phe Arg Leu Leu Phe Leu Leu Leu Asn Ser Pro Leu Ala Gly Leu Leu
 50 55 60

Tyr Tyr Phe Val Ser Glu Ser Ala Gly Ile Lys Val Leu Ile Phe Ala
 65 70 75 80

pctau2012001598-seql

Thr Cys Ala Gly Met Lys Leu Ser Asp Ile Glu Ser Val Ala Arg Ala
 85 90 95
 Val Leu Pro Lys Phe Tyr Ser Ser Asp Leu His Ser Glu Ser Trp Arg
 100 105 110
 Val Phe Ser Ser Cys Gly Lys Arg Cys Val Leu Thr Ala Asn Pro Arg
 115 120 125
 Ile Met Val Glu Ala Phe Leu Lys Asp Phe Leu Gly Ala Asp Leu Val
 130 135 140
 Leu Gly Thr Glu Ile Leu Thr Tyr Lys Gly Arg Ala Thr Gly Phe Val
 145 150 155 160
 Gln Ser Pro Gly Val Leu Val Gly Lys Asn Lys Ala Asp Ala Leu Lys
 165 170 175
 Lys Ala Phe Gly Lys Thr Gln Pro Glu Ile Gly Leu Gly Asp Arg His
 180 185 190
 Thr Asp Ala Pro Phe Met Ala Leu Cys Lys Glu Gly Tyr Ile Val Pro
 195 200 205
 Pro Lys Pro Glu Val Glu Ala Val Thr Thr Asp Lys Leu Pro Lys Pro
 210 215 220
 Val Ile Phe His Asp Gly Arg Leu Val Gln Lys Pro Thr Pro Leu Ser
 225 230 235 240
 Ala Leu Leu Ile Ile Leu Trp Ile Pro Ile Gly Phe Ile Leu Ala Cys
 245 250 255
 Leu Arg Ile Ala Ala Gly Ser Leu Leu Pro Met Pro Met Val Tyr Tyr
 260 265 270
 Ala Phe Trp Ala Leu Gly Val Arg Val Asn Ile Lys Gly Thr Pro Pro
 275 280 285
 Pro Pro Ala Lys Lys Ser Ile Gly Gln Ser Gly Val Leu Phe Val Cys
 290 295 300
 Ser His Arg Thr Leu Leu Asp Pro Ile Phe Leu Ser Thr Ala Leu Gly
 305 310 315 320
 Arg Pro Ile Pro Ala Val Thr Tyr Ser Leu Ser Arg Leu Ser Glu Ile
 325 330 335
 Ile Ser Pro Ile Lys Thr Val Arg Leu Ser Arg Asp Arg Ala Ala Asp
 340 345 350

pctau2012001598-seql

Ala Ser Met Ile Lys Lys Leu Leu Glu Glu Gly Asp Leu Ala Ile Cys
355 360 365

Pro Glu Arg Thr Thr Cys Arg Glu Pro Phe Leu Leu Arg Phe Ser Ala
370 375 380

Leu Phe Ala Glu Leu Thr Asp Glu Leu Val Pro Val Ala Met Ala Asn
385 390 395 400

Arg Met Ser Met Phe His Gly Thr Thr Ala Arg Gly Trp Lys Gly Met
405 410 415

Asp Pro Phe Tyr Phe Phe Met Asn Pro Ser Pro Ala Tyr Glu Val Thr
420 425 430

Phe Leu Asn Lys Leu Pro Leu Glu Leu Thr Leu Ser Gly Gly Lys Ser
435 440 445

Ser His Glu Val Ala Asn Tyr Ile Gln Arg Val Ile Ala Ala Thr Leu
450 455 460

Ser Tyr Glu Cys Thr Ser Phe Thr Arg Arg Asp Lys Tyr Arg Ala Leu
465 470 475 480

Ala Gly Asn Asp Gly Asn Val Val Glu Lys Thr Lys Val Gln Ala Asn
485 490 495

Lys Val Met Gly Cys
500

<210> 184
<211> 501
<212> PRT
<213> Populus trichocarpa

<400> 184

Met Asp Ser Phe Pro Thr Val Asp Lys Cys Ala Ser Ile Gly Arg Glu
1 5 10 15

Lys His Thr Val Val Ala Asp Met Asp Gly Thr Leu Leu Arg Gly Arg
20 25 30

Ser Ser Phe Pro Tyr Phe Ala Leu Leu Ala Phe Glu Ala Gly Gly Ile
35 40 45

Phe Arg Leu Ile Phe Leu Leu Leu Ala Ser Pro Leu Ala Gly Leu Leu
50 55 60

Tyr Tyr Phe Val Ser Glu Ser Ala Gly Ile Lys Val Leu Ile Phe Ala
65 70 75 80

pctau2012001598-seql

Thr Cys Ala Gly Met Lys Val Ser Asp Ile Glu Ser Val Ala Arg Ala
 85 90 95
 Val Leu Pro Lys Phe Tyr Ser Ser Asp Leu His Ser Glu Ser Trp Arg
 100 105 110
 Val Phe Ser Ser Cys Gly Lys Arg Cys Val Leu Thr Ala Asn Pro Arg
 115 120 125
 Ile Met Val Glu Ala Phe Leu Lys Asp Phe Leu Gly Ala Asp Leu Val
 130 135 140
 Leu Gly Thr Glu Met Ser Thr Tyr Lys Gly Arg Ala Thr Gly Phe Val
 145 150 155 160
 Gln Ser Pro Gly Val Leu Val Gly Lys Asn Lys Ala Asp Ala Leu Lys
 165 170 175
 Lys Ala Phe Gly Met Thr Gln Pro Glu Ile Gly Leu Gly Asp Arg His
 180 185 190
 Thr Asp Ala Pro Phe Met Ala Leu Cys Lys Glu Gly Tyr Ile Val Pro
 195 200 205
 Pro Lys Pro Glu Val Glu Ala Val Thr Thr Asp Lys Leu Pro Lys Pro
 210 215 220
 Val Ile Phe His Asp Gly Arg Leu Val Gln Lys Pro Thr Pro Leu Ser
 225 230 235 240
 Ala Leu Leu Thr Ile Leu Trp Ile Pro Ile Gly Phe Ile Leu Ala Cys
 245 250 255
 Leu Arg Ile Ala Ala Gly Ser Leu Leu Pro Met Pro Met Val Tyr Tyr
 260 265 270
 Ala Phe Trp Ala Leu Gly Val Arg Val Ile Ile Lys Gly Thr Pro Pro
 275 280 285
 Pro Pro Ala Lys Lys Ser Ile Gly Gln Ser Gly Val Leu Phe Ile Cys
 290 295 300
 Ser His Arg Thr Leu Leu Asp Pro Ile Phe Leu Ser Thr Ala Leu Gly
 305 310 315 320
 Arg Pro Ile Pro Ala Val Thr Tyr Ser Val Ser Arg Leu Ser Glu Ile
 325 330 335
 Ile Ser Pro Ile Lys Thr Val Arg Leu Ser Arg Asp Arg Ala Thr Asp
 340 345 350

pctau2012001598-seql

Ala Ser Met Ile Lys Lys Leu Leu Glu Glu Gly Asp Leu Ala Ile Cys
355 360 365

Pro Glu Gly Thr Thr Cys Arg Glu Pro Phe Leu Leu Arg Phe Ser Ala
370 375 380

Leu Phe Ala Glu Leu Thr Asp Gln Leu Val Pro Val Ala Val Val Asn
385 390 395 400

Arg Met Ser Met Phe His Gly Thr Thr Ala Arg Gly Trp Lys Gly Met
405 410 415

Asp Pro Phe Tyr Phe Phe Met Asn Pro Ser Pro Ala Tyr Glu Val Thr
420 425 430

Phe Leu Asn Arg Leu Pro Gln Glu Leu Thr Cys Thr Gly Gly Lys Ser
435 440 445

Ser His Glu Val Ala Asn Tyr Ile Gln Arg Val Ile Ala Ala Thr Leu
450 455 460

Ser Tyr Glu Cys Thr Ser Phe Thr Arg Arg Asp Lys Tyr Arg Ala Leu
465 470 475 480

Ala Gly Asn Asp Gly Thr Val Val Glu Lys Thr Lys Leu Gln Ala Asn
485 490 495

Lys Val Met Gly Cys
500

<210> 185
<211> 500
<212> PRT
<213> Populus trichocarpa

<400> 185

Met Val Met Gly Ala His Arg His Phe Glu Pro Ile Ala Lys Cys Ser
1 5 10 15

Thr Glu Gly Arg Ser Asn Gln Thr Val Ala Ala Asp Leu Asp Gly Thr
20 25 30

Leu Leu Val Ser Arg Ser Ala Phe Pro Tyr Phe Leu Leu Val Ala Ile
35 40 45

Glu Ala Gly Ser Leu Leu Arg Gly Leu Ile Leu Leu Ser Ser Val Pro
50 55 60

Phe Val Tyr Phe Thr Tyr Leu Ile Ile Ser Glu Glu Met Ala Ile Lys
65 70 75 80

pctau2012001598-seql

Thr Phe Ile Phe Ile Ala Phe Ala Gly Leu Lys Ile Arg Asp Ile Glu
 85 90 95
 Leu Val Ser Arg Ser Val Leu Pro Lys Phe Tyr Ala Asp Asp Val His
 100 105 110
 Pro Glu Thr Trp Arg Val Phe Asn Ser Phe Gly Lys Arg Tyr Ile Val
 115 120 125
 Thr Ala Asn Pro Arg Ile Met Val Glu Pro Phe Ala Lys Asn Phe Leu
 130 135 140
 Gly Ala Asp Lys Val Leu Gly Thr Glu Ile Gln Ala Thr Lys Ser Gly
 145 150 155 160
 Arg Ala Thr Gly Phe Val Lys Lys Pro Gly Val Leu Val Gly Asp His
 165 170 175
 Lys Arg Asn Ala Leu Leu Lys Glu Phe Gly Thr Ser Leu Pro Asp Leu
 180 185 190
 Gly Leu Gly Asp Arg Glu Thr Asp His Asp Phe Met Ser Ile Cys Lys
 195 200 205
 Glu Gly Tyr Ile Val Pro Arg Ser Lys Cys Glu Pro Leu Pro Arg Asn
 210 215 220
 Lys Leu Leu Ser Pro Val Ile Phe His Glu Gly Arg Leu Val Gln Arg
 225 230 235 240
 Pro Thr Pro Leu Val Ala Leu Leu Thr Phe Leu Trp Met Pro Ile Gly
 245 250 255
 Ile Ile Leu Ser Val Leu Arg Val Tyr Leu Asn Ile Pro Leu Pro Glu
 260 265 270
 Arg Ile Ala Trp Tyr Asn Tyr Lys Leu Leu Gly Ile Arg Val Val Val
 275 280 285
 Lys Gly Asn Pro Pro Pro Pro Arg Lys Gly Gln Ser Gly Val Leu
 290 295 300
 Phe Val Cys Asn His Arg Thr Val Leu Asp Pro Val Val Thr Ala Val
 305 310 315 320
 Ala Leu Arg Arg Lys Ile Ser Cys Val Thr Tyr Ser Ile Ser Lys Phe
 325 330 335
 Thr Glu Ile Ile Ser Pro Ile Lys Ala Val Ala Leu Ser Arg Glu Arg
 340 345 350

pctau2012001598-seql

Asp Arg Asp Ala Ala Asn Ile Lys Arg Leu Leu Glu Glu Gly Asp Leu
355 360 365

Val Ile Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro Phe Leu Leu Arg
370 375 380

Phe Ser Ala Leu Phe Ala Glu Leu Thr Asp Arg Ile Val Pro Val Ala
385 390 395 400

Ile Asn Thr Lys Gln Ser Val Phe Asn Gly Thr Thr Val Arg Gly His
405 410 415

Lys Leu Leu Asp Pro Tyr Phe Val Phe Met Asn Pro Met Pro Thr Tyr
420 425 430

Glu Ile Thr Phe Leu Asn Gln Leu Pro Thr Glu Leu Thr Cys Lys Gly
435 440 445

Gly Lys Ser Ser Ile Glu Val Ala Asn Tyr Ile Gln Arg Val Leu Ala
450 455 460

Gly Thr Leu Gly Phe Glu Cys Thr Asn Leu Thr Arg Lys Asp Lys Tyr
465 470 475 480

Ala Ile Leu Ala Gly Thr Asp Gly Arg Val Ala Thr Lys Lys Glu Lys
485 490 495

Ala Ser Glu Asn
500

<210> 186
<211> 524
<212> PRT
<213> Sorghum bicolor

<400> 186

Met Glu Arg Pro Ser Pro Thr Cys Leu Ala Ala Asn Gly Pro Arg Arg
1 5 10 15

Lys Ala Arg Thr Ile Pro Phe Pro Pro Pro Ala Asp Leu Arg Thr
20 25 30

Gly Arg Leu Arg Ser Arg Gln Thr Val Ala Ala Asp Leu Asp Gly Thr
35 40 45

Leu Leu Leu Ser Arg Ser Ala Phe Pro Tyr Tyr Leu Leu Val Ala Leu
50 55 60

Glu Ala Gly Ser Leu Leu Arg Ala Val Ala Leu Leu Met Ser Val Pro
65 70 75 80

Phe Val Tyr Leu Thr Tyr Val Thr Val Ser Glu Thr Leu Ala Val Arg
Page 234

pctau2012001598-seql

85

90

95

Ala Phe Leu Tyr Val Ala Val Ala Gly Leu Asp Val Ser Asp Ile Glu
 100 105 110

Ala Val Ala Arg Thr Val Leu Pro Arg Phe Tyr Ala Gly Asp Val His
 115 120 125

Pro Glu Gly Trp Arg Val Phe Arg Ser Phe Gly Arg Arg Cys Val Val
 130 135 140

Thr Ala Ser Pro Arg Val Met Val Glu Pro Phe Ala Arg Ala Phe Leu
 145 150 155 160

Gly Ala Asp Val Val Ile Gly Thr Glu Met Glu Val Gly Ala Ser Gly
 165 170 175

Lys Ala Thr Gly Phe Val Ala Gly Pro Gly Val Leu Val Gly Glu His
 180 185 190

Lys Arg Arg Ala Val Val Arg Glu Phe Gly Asp Ala Leu Pro Asp Val
 195 200 205

Gly Met Gly Asp Arg Glu Ser Asp Phe Asp Phe Met Ala Ile Cys Lys
 210 215 220

Glu Ala Tyr Met Val Thr Arg Gln Lys Tyr Arg Ala Leu Pro Arg Glu
 225 230 235 240

Gln Leu Gln Ser Arg Val Ile Leu His Asp Gly Arg Leu Ala Arg Arg
 245 250 255

Pro Thr Ala Thr Asn Thr Leu Leu Thr Leu Leu Trp Met Pro Leu Gly
 260 265 270

Phe Ala Leu Ala Leu Ala Arg Val His Leu His Leu Leu Leu Pro Ala
 275 280 285

Arg Ala Leu Ser Tyr Ala Tyr Lys Leu Met Gly Val Lys Leu Val Val
 290 295 300

Arg Gly Asn Arg Pro Pro Pro Pro Pro Ser Lys Lys Gly Gly Gly
 305 310 315 320

Pro Gly Val Leu Phe Val Cys Asn His Arg Thr Thr Leu Asp Pro Val
 325 330 335

Ala Val Ala Val Ala Leu Gly Arg Lys Val Arg Trp Val Val Thr Asp
 340 345 350

Gly Ala Ser Ser Arg Phe Ser Glu Ala Val Ser Pro Val Met Thr Gly
 Page 235

355

360

365

Val Ala Leu Pro Val Pro Val Pro Ser Arg Glu Ser Asp Asp Asp Ala
 370 375 380

Asp Ala Ala Ala Arg Ile Arg Arg Leu Leu Glu Glu Gly Asp Asp Val
 385 390 395 400

Val Ile Phe Pro Glu Gly Thr Ile Cys Arg Glu Pro Phe Leu Leu Arg
 405 410 415

Phe Gly Ala Leu Phe Ala Glu Val Thr Asp Arg Ile Val Pro Val Ala
 420 425 430

Ile Gly Ala Arg Glu Gly Met Phe His Gly Ser Thr Ala Arg Gly Leu
 435 440 445

Arg Arg Met Asp Pro Tyr Phe Phe Phe Met Asn Pro Arg Pro Thr Tyr
 450 455 460

Glu Val Thr Phe Leu Asn Gln Leu Pro Arg Glu Leu Thr Cys Gly Gly
 465 470 475 480

Gly Arg Ser Pro Val Glu Val Ala Asn Tyr Val Gln Glu Val Leu Ala
 485 490 495

Ala Gln Leu Gly Phe Asp Cys Thr Cys Thr Ser Lys Gln Ala Lys Thr
 500 505 510

Gly Trp Ser Pro Ala Ala Pro Thr Val Ala Trp Arg
 515 520

<210> 187
 <211> 503
 <212> PRT
 <213> Sorghum bicolor

<400> 187

Met Val Ala Ser Tyr Arg Ser Arg Arg Phe Lys Pro Ile Glu Glu Cys
 1 5 10 15

Cys Ser Glu Gly Arg Ser Glu Gln Thr Val Ala Ala Asp Leu Asp Gly
 20 25 30

Thr Leu Leu Ile Ser Arg Ser Ala Phe Pro Tyr Tyr Leu Leu Val Ala
 35 40 45

Leu Glu Ala Gly Ser Val Leu Arg Ala Val Leu Leu Leu Ser Val
 50 55 60

Pro Phe Val Tyr Val Thr Tyr Ile Phe Phe Ser Glu Ser Leu Ala Ile
 65 70 75 80

pctau2012001598-seql

Ser Thr Leu Val Tyr Ile Ser Val Ala Gly Leu Lys Val Arg Ser Ile
85 90 95

Glu Met Val Ala Arg Ser Val Leu Pro Arg Phe Tyr Ala Gly Asp Val
100 105 110

His Pro Glu Ser Trp Arg Val Phe Asn Ser Phe Gly Lys Arg Tyr Ile
115 120 125

Ile Thr Ala Ser Pro Arg Ile Met Val Glu Pro Phe Ala Arg Ala Phe
130 135 140

Leu Gly Ala Asp Lys Val Val Gly Thr Glu Leu Glu Val Gly Lys Asn
145 150 155 160

Gly Lys Ala Thr Gly Phe Met Val Lys Pro Gly Val Leu Val Gly Asp
165 170 175

His Lys Lys Lys Ala Val Val Lys Glu Leu Gly Asp Ala Val Pro Asp
180 185 190

Val Gly Met Gly Asp Arg Glu Thr Asp Phe Asp Phe Met Ser Ile Cys
195 200 205

Lys Glu Ala Tyr Leu Val Thr Ser Arg Lys Tyr Ser Pro Val Gly Lys
210 215 220

Asn Gln Leu Leu Ser Pro Leu Ile Leu His Asp Gly Arg Leu Val Gln
225 230 235 240

Arg Pro Thr Pro Leu Val Ala Leu Val Thr Phe Leu Trp Met Pro Phe
245 250 255

Gly Phe Ala Leu Ala Leu Met Arg Val Tyr Ile Asn Leu Pro Leu Pro
260 265 270

Glu Arg Ile Val Tyr Tyr Thr Tyr Lys Leu Met Gly Ile Arg Leu Ile
275 280 285

Val Lys Gly Asn Pro Pro Pro Pro Lys Lys Gly His Pro Gly Val
290 295 300

Leu Phe Val Cys Asn His Arg Thr Val Leu Asp Pro Val Glu Val Ala
305 310 315 320

Val Ala Leu Arg Arg Lys Val Ser Cys Val Thr Tyr Ser Ile Ser Lys
325 330 335

Phe Ser Glu Leu Ile Ser Pro Ile Lys Ala Val Ala Leu Ser Arg Glu
340 345 350

pctau2012001598-seql

Arg Glu Lys Asp Ala Glu Asn Ile Arg Arg Leu Leu Glu Glu Gly Asp
355 360 365

Leu Val Ile Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro Phe Leu Leu
370 375 380

Arg Phe Ser Ala Leu Phe Ala Glu Leu Thr Asp Arg Ile Val Pro Val
385 390 395 400

Ala Ile Asn Thr Lys Glu Ser Met Phe His Gly Ser Thr Val Arg Gly
405 410 415

Phe Lys Leu Met Asp Pro Tyr Phe Phe Phe Met Asn Pro Arg Pro Thr
420 425 430

Tyr Glu Ile Thr Phe Leu Thr Gln Leu Pro Lys Asp Leu Thr Cys Ser
435 440 445

Gly Gly Lys Ser Pro Ile Glu Val Ala Asn Tyr Ile Gln Lys Thr Leu
450 455 460

Ser Gly Gln Leu Gly Phe Glu Cys Thr Thr Ile Thr Arg Lys Glu Lys
465 470 475 480

Tyr Gly Ile Leu Ala Gly Thr Asp Gly Arg Val Pro Ser Lys Asn Lys
485 490 495

Asp Lys Glu Lys Glu Lys Asn
500

<210> 188
<211> 505
<212> PRT
<213> Sorghum bicolor

<400> 188

Met Ala Ala Ser Pro Phe Pro Thr Val Glu Lys Cys Ser Ser Thr Asp
1 5 10 15

Arg Ser Gly Asp Thr Val Val Ala Asp Leu Asp Gly Thr Leu Leu Cys
20 25 30

Gly Arg Ser Ser Phe Pro Tyr Phe Ala His Met Ala Phe Glu Thr Gly
35 40 45

Gly Val Leu Arg Leu Leu Leu Leu Ile Ala Leu Ala Pro Leu Ala Gly
50 55 60

Leu Leu Tyr Tyr Phe Val Ser Glu Pro Ala Gly Ile Gln Val Leu Ile
65 70 75 80

pctau2012001598-seql

Phe Ala Ser Met Ala Gly Ala Arg Val Ala Asp Ile Glu Ala Val Ala
85 90 95
Arg Ala Val Leu Pro Lys Phe Tyr Cys Ala Asp Leu His Pro Glu Ser
100 105 110
Trp Arg Val Phe Ser Ala Cys Gly Arg Arg Cys Val Leu Thr Ala Asn
115 120 125
Pro Arg Ile Met Val Glu Ala Phe Leu Lys Glu Tyr Val Gly Thr Asp
130 135 140
Val Val Val Gly Thr Glu Leu Val Val Trp Arg Gly Arg Ala Thr Gly
145 150 155 160
Leu Val Arg Ser Pro Gly Val Leu Val Gly Glu Gln Lys Ala Asp Ala
165 170 175
Leu Arg Arg Thr Phe Gly Asp Asp Val Ala Pro Glu Val Gly Leu Gly
180 185 190
Asp Arg Lys Thr Asp Tyr Pro Phe Met Arg Leu Cys Lys Glu Gly Tyr
195 200 205
Val Val Pro Ala Thr Pro Lys Leu Lys Pro Val Pro Arg Glu Asn Leu
210 215 220
Pro Lys Pro Val Val Phe His Asp Gly Arg Leu Val Gln Lys Pro Ser
225 230 235 240
Pro Ala Leu Ala Leu Leu Thr Val Leu Trp Ile Pro Ile Gly Phe Leu
245 250 255
Leu Ala Cys Leu Arg Ile Ala Ala Gly Ala Leu Leu Pro Met Arg Met
260 265 270
Val Tyr His Ala Phe Arg Ala Leu Gly Val Arg Val Thr Ile Arg Gly
275 280 285
Asn Pro Pro Pro Pro Ala Ser Arg Glu Thr Gly Gln Thr Gly Val Leu
290 295 300
Phe Ile Cys Ser His Arg Thr Leu Leu Asp Pro Ile Phe Leu Ser Thr
305 310 315 320
Ala Leu Gly Arg Pro Ile Thr Ala Val Thr Tyr Ser Val Ser Arg Leu
325 330 335
Ser Glu Ile Leu Ser Pro Ile Arg Thr Val Arg Leu Thr Arg Asp Arg
340 345 350

pctau2012001598-seql

Ala Ala Asp Ala Ala Met Ile Arg Arg Leu Leu Thr Glu Gly Asp Leu
355 360 365

Val Ile Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro Phe Leu Leu Arg
370 375 380

Phe Ser Ala Leu Phe Ala Glu Leu Thr Asp Glu Ile Val Pro Val Ala
385 390 395 400

Met Glu Asn Gln Met Ser Met Phe His Gly Thr Thr Ala Arg Gly Trp
405 410 415

Lys Gly Leu Asp Pro Phe Tyr Phe Phe Met Asn Pro Ser Pro Gly Tyr
420 425 430

Val Val Thr Phe Leu Asn Lys Leu Pro Ala Glu Leu Thr Cys Asn Gly
435 440 445

Gly Gly Lys Ser Ser His Glu Val Ala Asn Tyr Ile Gln Arg Leu Ile
450 455 460

Ala Ser Thr Leu Ser Tyr Glu Cys Thr Asn Phe Thr Arg Lys Asp Lys
465 470 475 480

Tyr Lys Ala Leu Ala Gly Asn Asp Gly Thr Val Val Ser Lys Pro Asn
485 490 495

Ile Asp Lys Asn Lys Val Met Gly Cys
500 505

<210> 189
<211> 518
<212> PRT
<213> Sorghum bicolor

<400> 189

Met Glu Glu Ala Val Glu Ala Val Val Ala Ala Gly Val Glu Pro Phe
1 5 10 15

Pro Thr Val Asp Lys Cys Asp Ala Ser Gly Leu Gly Ala His Ser Ser
20 25 30

Val Val Ala Asp Leu Asp Gly Thr Leu Leu Arg Ser Arg Ser Ala Phe
35 40 45

Pro Tyr Tyr Ala Leu Val Ala Phe Glu Thr Gly Gly Val Pro Arg Leu
50 55 60

Leu Leu Leu Leu Leu Leu Ser Pro Leu Ala Ala Leu Leu Arg Phe Leu
65 70 75 80

pctau2012001598-seql

Ala Ser Glu Ser Ala Cys Leu Arg Val Leu Val Phe Gly Ala Thr Ala
85 90 95

Gly Ala Arg Val Arg Asp Val Glu Ser Ala Ala Arg Ala Val Leu Pro
100 105 110

Arg Phe Tyr Ala Ala Asp Val His Pro Ala Ala Trp Arg Val Phe Ser
115 120 125

Ala Cys Ser Arg Arg Arg Val Val Leu Thr Ala Thr Pro Arg Leu Phe
130 135 140

Ala Glu Pro Phe Leu Arg Asp Cys Leu Gly Ala Asp Ala Val Ala Gly
145 150 155 160

Thr Glu Leu Ala Thr Trp Arg Gly Arg Ala Thr Gly Leu Val Asp Ala
165 170 175

Arg Arg Gly Gly Val Leu Val Gly Arg Ser Lys Ala Gln Ala Leu Arg
180 185 190

Glu Ile Leu Ala Ala Asp Gly Gly Asp Ala Pro Asp Val Gly Leu Gly
195 200 205

Asp Ser Arg Ser Asp Tyr Pro Phe Met Ser Met Cys Lys Glu Ala Tyr
210 215 220

Ile Val Pro Arg Ala Pro Val Glu Pro Ala Leu Pro Met Asp Gln Leu
225 230 235 240

Pro Arg Pro Val Ile Phe His Asp Gly Arg Leu Val Arg Arg Pro Thr
245 250 255

Pro Leu Val Ala Leu Val Val Val Leu Trp Phe Pro Val Gly Leu Ala
260 265 270

Leu Ala Cys Leu Arg Ile Ala Ala Gly Ala Leu Leu Pro Met Pro Leu
275 280 285

Val Tyr Tyr Ala Phe Trp Ala Leu Gly Val Arg Val Val Val Arg Gly
290 295 300

Ala Pro Pro Pro Pro Arg Ala Glu Arg Ala Thr Gly Arg Arg Gly Val
305 310 315 320

Leu Phe Ala Cys Ser His Arg Thr Leu Leu Asp Pro Ile Phe Leu Ser
325 330 335

Thr Ala Leu Gly Arg Pro Val Ala Ala Val Thr Tyr Ser Leu Ser Arg
340 345 350

pctau2012001598-seql

Leu Ser Glu Phe Leu Ser Pro Ile Arg Thr Val Arg Leu Ser Arg Asp
355 360 365

Arg Ala Thr Asp Ala Ala Met Ile Arg Asp Leu Leu Ala Glu Gly Asp
370 375 380

Leu Val Ile Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro Phe Leu Leu
385 390 395 400

Arg Phe Ser Ala Leu Phe Ala Glu Leu Thr Ser Glu Val Val Pro Val
405 410 415

Ala Met Glu Asn Arg Met Ser Met Phe His Gly Thr Thr Ala Arg Gly
420 425 430

Trp Lys Gly Met Asp Pro Phe Tyr Phe Phe Met Asn Pro Ser Pro Ala
435 440 445

Tyr Val Val Thr Phe Leu Asn Lys Leu Pro Pro Glu Leu Thr Cys Ala
450 455 460

Gly Gly Arg Thr Ser His Glu Val Ala Asn Tyr Ile Gln Arg Leu Ile
465 470 475 480

Ala Ala Thr Leu Ser Tyr Glu Cys Thr Ser Leu Thr Arg Lys Asp Lys
485 490 495

Tyr Arg Ala Leu Ala Gly Asn Asp Gly Val Val Ala Thr Pro Lys Pro
500 505 510

Pro Pro Thr Val Ala Ser
515

<210> 190
<211> 501
<212> PRT
<213> Ricinus communis

<400> 190

Met Pro Ala Met Lys Gln Ala Gln Lys Phe Pro Ser Ile Thr Ser Cys
1 5 10 15

Ser Gly Thr Ser Tyr Glu Ser Ile Ala Ala Asp Leu Asp Gly Thr Leu
20 25 30

Leu Val Ser Ser Ser Ser Phe Pro Tyr Phe Met Leu Val Ala Val Glu
35 40 45

Ala Gly Ser Leu Leu Arg Gly Leu Val Leu Leu Leu Ser Leu Pro Phe
50 55 60

Ile Ile Ile Ser Tyr Phe Phe Ile Ser Glu Ala Ile Gly Ile Gln Ile
Page 242

pctau2012001598-seql

```

65          70          75          80
Leu Ile Phe Ile Ser Phe Ala Gly Leu Lys Ile Arg Asp Ile Glu Leu
      85          90          95
Val Ser Arg Ala Val Leu Pro Arg Phe Tyr Ala Ala Asp Val Arg Lys
      100         105         110
Glu Ser Tyr Glu Val Phe Asp Arg Cys Lys Arg Lys Val Val Val Thr
      115         120         125
Ala Asn Pro Thr Ile Met Val Glu Pro Phe Val Lys Asp Phe Leu Gly
      130         135         140
Gly Asp Lys Val Leu Gly Thr Glu Ile Glu Val Asn Pro Lys Thr Lys
      145         150         155         160
Arg Ala Thr Gly Phe Val Lys Lys Pro Gly Val Leu Val Gly Lys Trp
      165         170         175
Lys Lys Leu Ala Ile Leu Lys Glu Phe Gly Glu Asp Ala Pro Asp Leu
      180         185         190
Gly Ile Gly Asp Arg Lys Thr Asp His Asp Phe Met Ser Ile Cys Lys
      195         200         205
Glu Gly Tyr Met Val Tyr His Ser Lys Lys Ala Ala Thr Pro Leu Pro
      210         215         220
Arg Asp Arg Leu Lys Ser Pro Ile Ile Phe His Asp Gly Arg Phe Val
      225         230         235         240
Gln Arg Pro Asp Pro Leu Asn Ala Leu Ala Thr Tyr Leu Trp Leu Pro
      245         250         255
Phe Gly Phe Met Leu Ser Ile Phe Arg Val Tyr Phe Asn Leu Pro Leu
      260         265         270
Pro Glu Arg Ile Val Arg Tyr Thr Tyr Glu Met Leu Gly Ile His Leu
      275         280         285
Val Ile Arg Gly Tyr Pro Pro Pro Ala Pro Ser Arg Gly Thr Pro Gly
      290         295         300
Asn Leu Tyr Val Cys Asn His Arg Thr Ala Leu Asp Pro Ile Val Ile
      305         310         315         320
Ala Ile Ala Leu Gly Arg Lys Val Ser Cys Val Thr Tyr Ser Val Ser
      325         330         335
Arg Leu Ser Arg Phe Leu Ser Pro Ile Pro Ala Ile Ala Leu Thr Arg

```

pctau2012001598-seql

340

345

350

Asp Arg Ala Ala Asp Ala Glu Arg Ile Thr Ala Leu Leu Gln Lys Gly
 355 360 365

Asp Leu Val Val Cys Pro Glu Gly Thr Thr Cys Arg Glu Gln Phe Leu
 370 375 380

Leu Arg Phe Ser Ala Leu Phe Ala Glu Met Ser Asp Arg Ile Val Pro
 385 390 395 400

Val Ala Val Asn Cys Lys Gln Ser Met Phe Tyr Gly Thr Thr Val Arg
 405 410 415

Gly Val Lys Phe Trp Asp Pro Tyr Phe Phe Phe Met Asn Pro Arg Pro
 420 425 430

Thr Tyr Glu Val Thr Phe Leu Asp Arg Leu Pro Glu Glu Met Thr Val
 435 440 445

Lys Ala Gly Gly Lys Ser Ser Ile Glu Val Ala Asn Tyr Val Gln Lys
 450 455 460

Val Leu Gly Asp Val Leu Gly Phe Gln Cys Thr Gly Leu Thr Arg Lys
 465 470 475 480

Asp Lys Tyr Leu Leu Leu Gly Gly Asn Asp Gly Lys Val Glu Ser Met
 485 490 495

Tyr Asn Ser Lys Lys
 500

<210> 191
 <211> 512
 <212> PRT
 <213> Ricinus communis

<400> 191

Met Ala Asn Pro Lys Glu Thr Tyr Ala Pro Phe Ser Ser Phe Pro Asn
 1 5 10 15

Ile Glu Lys Cys Thr Ser Val Gly Arg Gln Asn Asp Thr Val Val Thr
 20 25 30

Asp Met Asp Gly Thr Leu Leu Cys Gly Arg Ser Pro Phe Pro Tyr Phe
 35 40 45

Ala Leu Val Ala Tyr Glu Ala Ser Gly Ile Leu Arg Leu Leu Phe Leu
 50 55 60

Leu Leu Gly Ser Pro Ile Ala Gly Ile Leu Arg Tyr Phe Ile Ser Glu
 65 70 75 80

pctau2012001598-seql

Ser Ala Gly Ile Arg Val Leu Val Phe Ala Thr Phe Ala Gly Met Arg
85 90 95
Val Ser Asp Ile Glu Ser Val Ala Arg Ala Val Leu Pro Lys Phe Tyr
100 105 110
Ser Ser Asp Leu His Pro Glu Thr Trp Arg Val Phe Ser Ser Cys Gly
115 120 125
Lys Lys Cys Val Ile Thr Glu Asn Pro Arg Ile Met Val Glu Ala Phe
130 135 140
Leu Lys Glu Phe Val Gly Ala Asp Met Val Ile Gly Thr Glu Ile Val
145 150 155 160
Val His Lys Gly Arg Ala Thr Gly Phe Ile Arg Ser Pro Gly Val Leu
165 170 175
Val Gly Gln Asn Lys Ala Asp Ala Leu Lys Met Ala Phe Tyr Asp Ala
180 185 190
Pro Val Pro Asn Ile Gly Ile Gly Asp Arg Arg Thr Asp Arg Pro Phe
195 200 205
Met Lys Leu Cys Lys Glu Ser Tyr Ile Val Pro Ser Asn Ser Gln Val
210 215 220
Glu Pro Val Gly Pro Lys Lys Leu Pro Lys Pro Ile Val Phe His Asp
225 230 235 240
Gly Arg Leu Val Gln Lys Pro Thr His Leu Ile Ala Leu Ala Thr Leu
245 250 255
Leu Trp Ile Pro Val Gly Phe Ile Leu Ala Cys Leu Arg Leu Ala Ala
260 265 270
Gly Ala Leu Leu Pro Met Pro Leu Val Tyr Tyr Ala Leu Trp Ala Leu
275 280 285
Gly Val Arg Ile His Ile Lys Gly Thr Pro Pro Pro Pro Ala Lys Lys
290 295 300
Ser Thr Gly Gln Ala Gly Val Leu Phe Val Cys Ser His Arg Thr Leu
305 310 315 320
Leu Asp Pro Val Phe Leu Ser Ile Ala Leu Gly Arg Pro Ile Pro Ala
325 330 335
Val Thr Tyr Ser Leu Ser Arg Leu Ser Glu Phe Ile Ser Pro Ile Lys
340 345 350

pctau2012001598-seql

Thr Val Arg Leu Thr Arg Asp Arg Val Thr Asp Ala Asn Met Ile Lys
355 360 365

Glu Leu Leu Gln Glu Gly Asp Leu Val Ile Cys Pro Glu Gly Thr Thr
370 375 380

Cys Arg Glu Pro Phe Leu Leu Arg Phe Ser Ala Leu Phe Ala Glu Leu
385 390 395 400

Thr Asp Glu Leu Val Pro Val Ala Met Ala Asn Lys Met Thr Met Phe
405 410 415

His Gly Thr Thr Ala Arg Gly Trp Lys Gly Met Asp Pro Phe Tyr Phe
420 425 430

Leu Met Asn Pro Ser Pro Ala Tyr Glu Val Thr Phe Leu Asn Lys Leu
435 440 445

Pro Tyr Asn Leu Thr Cys Gly Ala Gly Lys Ser Ser His Ser Val Ala
450 455 460

Asn Tyr Ile Gln Arg Thr Ile Ala Ala Thr Leu Ser Tyr Gln Cys Thr
465 470 475 480

Ser Phe Thr Arg Lys Asp Lys Tyr Arg Ala Leu Ala Gly Asn Asp Gly
485 490 495

Thr Val Val Glu Glu Pro Lys Gln Ala Arg Asp Glu Val Met Asp Ser
500 505 510

<210> 192
<211> 504
<212> PRT
<213> Ricinus communis

<400> 192

Met Val Met Val Asp Ser Phe Pro Glu Val Ser Lys Cys Ala Ser Lys
1 5 10 15

Gly Arg Glu Lys His Thr Val Val Ala Asp Met Asp Gly Thr Leu Phe
20 25 30

Ile Gly Arg Ser Ser Phe Pro Tyr Phe Ala Leu Ile Ala Phe Glu Ala
35 40 45

Gly Gly Ile Phe Arg Leu Leu Phe Leu Leu Leu Ala Ser Pro Leu Ala
50 55 60

Ala Leu Leu Tyr Tyr Phe Val Ser Glu Ser Ala Gly Ile Gln Val Leu
65 70 75 80

pctau2012001598-seql

Ile Phe Ala Ser Phe Ala Gly Met Lys Val Ser Asp Ile Glu Ser Val
85 90 95
Ala Arg Ala Val Leu Pro Lys Phe Tyr Ser Ser Asp Leu His Pro Glu
100 105 110
Ser Trp Arg Val Phe Ser Ser Cys Gly Lys Arg Cys Val Leu Thr Ala
115 120 125
Asn Pro Arg Ile Met Val Glu Ala Phe Leu Lys Asp Phe Leu Gly Ala
130 135 140
Asp Leu Val Phe Gly Thr Glu Ile Ser Thr Tyr Lys Gly Arg Ala Thr
145 150 155 160
Gly Phe Val Arg Thr Pro Gly Val Leu Val Gly Lys Asn Lys Ala Asn
165 170 175
Ala Leu Lys Lys Ala Phe Gly Asp Ala Arg Pro Glu Val Gly Leu Gly
180 185 190
Asp Arg Arg Thr Asp Ala Pro Phe Met Ala Leu Cys Lys Glu Gly Tyr
195 200 205
Leu Val Pro Pro Lys Pro Glu Val Arg Ala Val Thr Gly Asp Lys Leu
210 215 220
Pro Lys Pro Ile Val Phe His Asp Gly Arg Leu Val Gln Lys Pro Thr
225 230 235 240
Pro Leu Met Ala Leu Leu Ile Ile Leu Trp Ile Pro Ile Gly Phe Ile
245 250 255
Leu Ala Cys Leu Arg Ile Ala Ala Gly Ser Leu Leu Pro Met Pro Met
260 265 270
Val Tyr Tyr Ala Phe Leu Ala Leu Gly Val Arg Val Thr Val Lys Gly
275 280 285
Asn Pro Pro Pro Pro Ala Lys Lys Ser Thr Gly Gln Ser Gly Val Leu
290 295 300
Phe Val Cys Ser His Arg Thr Leu Leu Asp Pro Ile Phe Leu Ser Thr
305 310 315 320
Ala Leu Gly Arg Pro Ile Ala Ala Val Thr Tyr Ser Val Ser Arg Leu
325 330 335
Ser Glu Ile Ile Ser Pro Ile Lys Thr Val Arg Leu Ser Arg Asp Arg
340 345 350

pctau2012001598-seql

Ala Thr Asp Ala Ala Met Ile Lys Lys Leu Leu Val Glu Gly Asp Leu
355 360 365

Ala Ile Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro Phe Leu Leu Arg
370 375 380

Phe Ser Ala Leu Phe Ala Glu Leu Thr Asp Gln Ile Val Pro Val Ala
385 390 395 400

Met Val Asn Arg Met Ser Met Phe His Gly Thr Thr Ala Arg Gly Trp
405 410 415

Lys Gly Met Asp Pro Phe Tyr Phe Phe Met Asn Pro Ser Pro Ala Tyr
420 425 430

Glu Val Thr Phe Leu Ser Lys Leu Pro Gln Glu Leu Thr Cys Ser Ser
435 440 445

Gly Lys Ser Ser His Glu Val Ala Asn Tyr Ile Gln Arg Val Ile Ala
450 455 460

Ala Thr Leu Ser Tyr Glu Cys Thr Gly Phe Thr Arg Lys Asp Lys Tyr
465 470 475 480

Arg Ala Leu Ala Gly Asn Asp Gly Ser Val Val Glu Lys Pro Lys Leu
485 490 495

Glu Ala Asn Lys Val Met Gly Cys
500

<210> 193
<211> 498
<212> PRT
<213> Arabidopsis lyrata

<400> 193

Met Ser Pro Ala Lys Lys Ser Gly Asn Phe Pro Pro Ile Thr Glu Cys
1 5 10 15

Arg Asp Gly Glu Tyr Asp Ser Ile Ala Ala Asp Leu Asp Gly Thr Leu
20 25 30

Leu Leu Ser Arg Ser Ser Phe Pro Tyr Phe Met Leu Val Ala Val Glu
35 40 45

Ala Gly Ser Leu Phe Arg Gly Leu Ile Leu Leu Leu Ser Leu Pro Ile
50 55 60

Val Ile Ile Ser Tyr Leu Phe Val Ser Glu Ser Leu Gly Ile Gln Ile
65 70 75 80

pctau2012001598-seql

Leu Ile Phe Ile Ser Phe Ala Gly Leu Lys Ile Arg Asp Ile Glu Leu
85 90 95

Val Ser Arg Ala Val Leu Pro Arg Phe Tyr Ala Ala Asp Val Arg Lys
100 105 110

Asp Ser Phe Asp Val Phe Asp Lys Cys Lys Arg Lys Val Val Val Thr
115 120 125

Ala Asn Pro Ile Val Met Val Glu Ala Phe Val Lys Asp Tyr Leu Gly
130 135 140

Gly Asp Lys Val Leu Gly Thr Glu Ile Glu Val Asn Pro Lys Ser Asn
145 150 155 160

Arg Ala Thr Gly Phe Val Lys Lys Pro Gly Val Leu Val Gly Asp Leu
165 170 175

Lys Arg Leu Ala Ile Leu Lys Glu Phe Gly Asp Glu Ser Pro Asp Leu
180 185 190

Gly Leu Gly Asp Arg Thr Ser Asp His Asp Phe Met Ser Leu Cys Lys
195 200 205

Lys Gly Tyr Met Val His Ala Cys Lys Ser Ala Thr Thr Ile Pro Lys
210 215 220

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

Arg Pro Thr Pro Leu Asn Ala Ile Val Thr Tyr Leu Trp Leu Pro Phe
245 250 255

Gly Phe Ile Leu Ser Ile Ile Arg Val Tyr Phe Asn Leu Pro Leu Pro
260 265 270

Glu Arg Phe Val Arg Tyr Thr Tyr Glu Ile Leu Gly Ile His Leu Thr
275 280 285

Ile Arg Gly His Arg Pro Pro Pro Ser Pro Gly Thr Leu Gly Asn
290 295 300

Leu Tyr Val Leu Asn His Arg Thr Ala Leu Asp Pro Ile Ile Val Ala
305 310 315 320

Ile Ala Leu Gly Arg Lys Ile Cys Cys Val Thr Tyr Ser Val Ser Arg
325 330 335

Leu Ser Leu Met Leu Ser Pro Ile Pro Ala Val Ala Leu Thr Arg Asp
340 345 350

pctau2012001598-seql

Arg Val Thr Asp Ala Ala Asn Met Arg Lys Leu Leu Glu Lys Gly Asp
355 360 365

Leu Val Ile Cys Pro Glu Gly Thr Thr Cys Arg Glu Glu Tyr Leu Leu
370 375 380

Arg Phe Ser Ala Leu Phe Ser Glu Leu Ser Asp Arg Ile Val Pro Val
385 390 395 400

Ala Met Asn Cys Lys Gln Gly Met Phe Asn Gly Thr Thr Val Arg Gly
405 410 415

Val Lys Phe Trp Asp Pro Tyr Phe Phe Phe Met Asn Pro Arg Pro Ser
420 425 430

Tyr Glu Ala Thr Phe Leu Asp Arg Leu Pro Glu Glu Met Thr Val Asn
435 440 445

Gly Gly Gly Lys Thr Pro Ile Glu Val Ala Asn Tyr Val Gln Lys Val
450 455 460

Ile Gly Ala Val Leu Gly Phe Glu Cys Thr Glu Leu Thr Arg Lys Asp
465 470 475 480

Lys Tyr Leu Leu Leu Gly Gly Asn Asp Gly Lys Val Glu Ser Ile Met
485 490 495

Leu Gly

<210> 194
<211> 501
<212> PRT
<213> Arabidopsis lyrata

<400> 194

Met Gly Ala Gln Glu Lys Arg Arg Arg Phe Glu Pro Ile Ser Lys Cys
1 5 10 15

Asp Val Lys Asp Arg Ser Asn His Thr Val Ala Ala Asp Leu Asp Gly
20 25 30

Thr Leu Leu Ile Ser Arg Ser Ala Phe Pro Tyr Tyr Phe Leu Val Ala
35 40 45

Leu Glu Ala Gly Gly Leu Leu Arg Ala Leu Ile Leu Leu Val Ser Val
50 55 60

Pro Phe Val Tyr Leu Thr Tyr Leu Thr Val Ser Glu Thr Leu Ala Ile
65 70 75 80

Asn Val Phe Val Phe Ile Thr Phe Ala Gly Leu Lys Ile Arg Asp Val
Page 250

pctau2012001598-seql

85

90

95

Glu Leu Val Val Arg Ser Val Leu Pro Arg Phe Tyr Ala Glu Asp Val
 100 105 110

Arg Pro Asp Thr Trp Arg Ile Phe Asn Thr Phe Gly Lys Arg Tyr Ile
 115 120 125

Val Thr Ala Ser Pro Arg Ile Met Val Glu Pro Phe Val Lys Thr Tyr
 130 135 140

Leu Gly Val Asp Lys Val Leu Gly Thr Glu Leu Glu Val Ser Lys Ser
 145 150 155 160

Gly Arg Ala Thr Gly Phe Thr Arg Lys Pro Gly Ile Leu Val Gly Gln
 165 170 175

His Lys Arg Glu Val Val Leu Arg Glu Phe Gly Ser Leu Ala Ser Asp
 180 185 190

Leu Pro Asp Leu Gly Leu Gly Asp Ser Lys Thr Asp Tyr Asp Phe Met
 195 200 205

Ser Ile Cys Lys Glu Gly Tyr Met Val Pro Arg Thr Lys Cys Glu Pro
 210 215 220

Leu Pro Arg Asn Lys Leu Leu Ser Pro Ile Ile Phe His Glu Gly Arg
 225 230 235 240

Leu Val Gln Arg Pro Thr Pro Leu Val Ala Leu Leu Thr Leu Leu Trp
 245 250 255

Leu Pro Ile Gly Phe Leu Leu Ser Leu Ile Arg Val Tyr Thr Asn Ile
 260 265 270

Pro Leu Pro Glu Arg Ile Ala Arg Tyr Asn Tyr Lys Leu Thr Gly Ile
 275 280 285

Lys Leu Ile Val Asn Gly His Pro Pro Pro Pro Pro Lys Pro Gly Gln
 290 295 300

Pro Gly His Leu Leu Val Cys Asn His Arg Thr Val Leu Asp Pro Val
 305 310 315 320

Val Thr Ala Val Ala Leu Gly Arg Lys Ile Ser Cys Val Thr Tyr Ser
 325 330 335

Ile Ser Lys Phe Ser Glu Leu Ile Ser Pro Ile Lys Ala Val Ala Leu
 340 345 350

Thr Arg Gln Arg Glu Lys Asp Ala Ala Asn Ile Lys Arg Leu Leu Glu
 Page 251

355

360

365

Glu Gly Asp Leu Val Ile Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro
 370 375 380

Phe Leu Leu Arg Phe Ser Ala Leu Phe Ala Glu Leu Thr Asp Arg Ile
 385 390 395 400

Val Pro Val Ala Ile Asn Thr Lys Gln Ser Met Phe Asn Gly Thr Thr
 405 410 415

Thr Arg Gly Tyr Lys Leu Leu Asp Pro Tyr Phe Ala Phe Met Asn Pro
 420 425 430

Arg Pro Thr Tyr Glu Ile Thr Phe Leu Lys Gln Ile Pro Ala Glu Leu
 435 440 445

Thr Cys Lys Gly Gly Lys Ser Pro Ile Glu Val Ala Asn Tyr Ile Gln
 450 455 460

Arg Val Leu Gly Gly Thr Leu Gly Phe Glu Cys Thr Asn Phe Thr Arg
 465 470 475 480

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

Lys Lys Glu Lys Thr
 500

<210> 195
 <211> 502
 <212> PRT
 <213> Vernicia fordii

<400> 195

Met Ser Gln Thr Lys Pro Ala Pro Lys Phe Pro Ser Ile Thr Ser Cys
 1 5 10 15

Thr Gly Ser Ala Tyr Gln Ser Ile Ala Ala Asp Leu Asp Gly Thr Leu
 20 25 30

Leu Val Ser Ser Ser Ser Phe Pro Tyr Phe Met Leu Val Ala Val Glu
 35 40 45

Ala Gly Ser Leu Leu Arg Gly Leu Val Ser Leu Leu Ser Leu Pro Leu
 50 55 60

Val Ile Ile Ser Tyr Phe Phe Ile Ser Glu Ala Val Gly Ile Gln Ile
 65 70 75 80

Leu Ile Tyr Ile Ser Phe Ala Gly Leu Lys Ile Arg Asp Ile Glu Leu
 85 90 95

pctau2012001598-seql

Val Ser Arg Ala Val Leu Pro Arg Phe Tyr Ala Ala Asp Val Arg Lys
100 105 110
Glu Ser Phe Glu Val Phe Asp Lys Cys Lys Arg Lys Val Val Val Thr
115 120 125
Ala Asn Pro Thr Ile Met Val Glu Pro Phe Val Lys Asp Phe Leu Gly
130 135 140
Gly Asp Lys Val Leu Gly Thr Glu Ile Glu Val Asn Pro Lys Thr Lys
145 150 155 160
Arg Ala Thr Gly Phe Val Lys Asn Pro Gly Val Leu Val Gly Lys Trp
165 170 175
Lys Lys Leu Ser Ile Leu Lys Glu Phe Gly Glu Glu Ser Pro Asp Leu
180 185 190
Gly Ile Gly Asp Arg Lys Thr Asp His Asp Phe Met Ser Ile Cys Lys
195 200 205
Glu Gly Tyr Met Val Gln Arg Ser Lys Ser Ala Thr Pro Ile Pro Leu
210 215 220
Asp Arg Leu Lys Ser Arg Ile Ile Phe His Asp Gly Arg Phe Val Gln
225 230 235 240
Arg Pro Asp Pro Leu Asn Ala Leu Val Thr Tyr Leu Trp Leu Pro Phe
245 250 255
Gly Phe Ile Leu Ser Ile Ile Arg Val Tyr Phe Asn Leu Pro Leu Pro
260 265 270
Glu Arg Ile Val Arg Tyr Thr Tyr Glu Met Leu Gly Ile His Leu Val
275 280 285
Ile Arg Gly Asn Pro Pro Pro Ala Pro Ser Pro Gly Thr Pro Gly Asn
290 295 300
Leu Tyr Val Cys Asn His Arg Ser Ala Leu Asp Pro Ile Val Ile Ala
305 310 315 320
Ile Ala Leu Gly Arg Lys Val Ser Cys Val Thr Tyr Ser Val Ser Arg
325 330 335
Leu Ser Arg Phe Leu Ser Pro Ile Pro Ala Ile Ala Leu Thr Arg Asp
340 345 350
Arg Ala Ala Asp Ala Ala Arg Ile Ser Glu Leu Leu Gln Lys Gly Asp
355 360 365

pctau2012001598-seql

Leu Val Val Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro Phe Leu Leu
370 375 380

Arg Phe Ser Ala Leu Phe Ala Glu Met Ser Asp Arg Ile Val Pro Val
385 390 395 400

Ala Val Asn Cys Lys Gln Asn Met Phe Tyr Gly Thr Thr Val Arg Gly
405 410 415

Val Lys Phe Trp Asp Pro Tyr Tyr Phe Phe Met Asn Pro Arg Pro Thr
420 425 430

Tyr Glu Val Thr Phe Leu Asp Arg Leu Pro Glu Glu Met Thr Ala Lys
435 440 445

Ala Gly Gly Lys Ser Ser Ile Glu Val Ala Asn Tyr Val Gln Lys Val
450 455 460

Leu Gly Asp Val Leu Gly Phe Glu Cys Thr Gly Leu Thr Arg Lys Asp
465 470 475 480

Lys Tyr Met Leu Leu Gly Gly Asn Asp Gly Lys Val Glu Ser Met Tyr
485 490 495

Asn Thr Lys Lys Ala Gly
500

<210> 196
<211> 514
<212> PRT
<213> Oryza sativa

<400> 196

Met Gly Asp Ala Thr Met Ile Ala Ser Pro Phe Pro Ala Val Glu Lys
1 5 10 15

Cys Ser Ser Lys Asp Arg Gly Gly Asp Thr Ile Val Ala Asp Leu Asp
20 25 30

Gly Thr Leu Leu Cys Gly Arg Ser Ser Phe Pro Tyr Phe Ala His Met
35 40 45

Ala Phe Glu Thr Gly Gly Val Leu Arg Leu Leu Ala Leu Ile Leu Leu
50 55 60

Ala Pro Leu Ala Gly Val Leu Tyr Tyr Leu Val Ser Glu Ser Ala Gly
65 70 75 80

Ile Gln Val Leu Ile Phe Ala Ser Met Ala Gly Ala Arg Val Ala Asp
85 90 95

pctau2012001598-seql

Val Glu Ala Val Ala Arg Ala Val Leu Pro Lys Phe Tyr Cys Ser Asp
 100 105 110
 Ile His Pro Glu Ser Trp Arg Val Phe Ser Ala Cys Gly Arg Arg Cys
 115 120 125
 Val Leu Thr Ala Asn Pro Arg Met Met Val Glu Ala Phe Leu Lys Glu
 130 135 140
 Tyr Leu Gly Ala Asp Ile Val Val Gly Thr Glu Leu Val Val Trp Arg
 145 150 155 160
 Gly Arg Ala Thr Gly Leu Val Arg Ser Pro Gly Val Leu Val Gly Glu
 165 170 175
 Asn Lys Ala Ala Ala Leu Arg Glu Ala Phe Gly Asp Ala Ser Pro Glu
 180 185 190
 Ile Gly Ile Gly Asp Arg Arg Thr Asp Tyr Pro Phe Met Arg Leu Cys
 195 200 205
 Lys Glu Gly Tyr Val Val Pro Val Ser Pro Thr Pro Arg Pro Val Pro
 210 215 220
 Arg Glu Glu Leu Pro Lys Pro Val Val Phe His Asp Gly Arg Leu Val
 225 230 235 240
 Gln Lys Pro Ser Pro Ala Leu Ala Leu Leu Thr Val Leu Trp Ile Pro
 245 250 255
 Ile Gly Phe Val Leu Ala Cys Leu Arg Ile Ala Ala Gly Ala Leu Leu
 260 265 270
 Pro Met Arg Val Val Tyr His Ala Phe Arg Ala Leu Gly Val Arg Val
 275 280 285
 Thr Val Lys Gly Thr Pro Pro Pro Pro Ala Ser Arg Glu Thr Gly Gln
 290 295 300
 Ser Gly Val Leu Phe Ile Cys Ser His Arg Thr Leu Leu Asp Pro Ile
 305 310 315 320
 Phe Leu Ser Thr Ala Leu Gly Arg Pro Ile Thr Ala Val Thr Tyr Ser
 325 330 335
 Val Ser Arg Leu Ser Glu Ile Leu Ser Pro Ile Arg Thr Val Arg Leu
 340 345 350
 Thr Arg Asp Arg Ala Ala Asp Ala Ala Met Ile Arg Arg Leu Leu Ala
 355 360 365

pctau2012001598-seql

Glu Gly Asp Leu Val Ile Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro
370 375 380

Phe Leu Leu Arg Phe Ser Ala Leu Phe Ala Glu Leu Thr Asp Glu Ile
385 390 395 400

Val Pro Val Ala Met Glu Asn Gln Met Ser Met Phe His Gly Thr Thr
405 410 415

Ala Arg Gly Trp Lys Ala Leu Asp Pro Phe Tyr Phe Phe Met Asn Pro
420 425 430

Ser Pro Gly Tyr Val Val Thr Phe Leu Ser Lys Leu Pro Gly Glu Leu
435 440 445

Thr Cys Asn Gly Gly Gly Gly Gly Gly Gly Gly Arg Ser Ser His Glu
450 455 460

Val Ala Asn Tyr Ile Gln Arg Leu Ile Ala Ser Thr Leu Ser Tyr Glu
465 470 475 480

Cys Thr Ser Phe Thr Arg Lys Asp Lys Tyr Arg Ala Leu Ala Gly Asn
485 490 495

Asp Gly Thr Val Val Ser Lys Pro Asn Ile Asp Glu Lys Lys Ala Met
500 505 510

Gly Cys

<210> 197
<211> 503
<212> PRT
<213> Brassica napus

<400> 197

Met Ser Pro Ala Lys Lys Lys Ser Arg Ser Phe Pro Pro Ile Ser Glu
1 5 10 15

Cys Lys Arg Gly Glu Tyr Asp Ser Ile Ala Ala Asp Leu Asp Gly Thr
20 25 30

Met Leu Leu Ser Arg Ser Ser Phe Pro Tyr Phe Met Leu Val Ala Ile
35 40 45

Glu Ala Gly Ser Leu Phe Arg Gly Leu Ile Leu Leu Leu Ser Leu Pro
50 55 60

Ile Val Ile Ile Ala Tyr Leu Phe Val Ser Glu Ala Leu Gly Ile Gln
65 70 75 80

pctau2012001598-seql

Ile Leu Ile Tyr Ile Ser Phe Ala Gly Ile Lys Ile Arg Asp Ile Glu
85 90 95
Leu Val Ser Arg Ala Val Leu Pro Arg Phe Tyr Ala Ala Asp Val Arg
100 105 110
Lys Asp Ser Phe Glu Val Phe Asp Lys Cys Lys Arg Lys Val Val Val
115 120 125
Thr Ala Asn Pro Ile Val Met Val Glu Pro Phe Val Lys Asp Tyr Leu
130 135 140
Gly Gly Asp Lys Val Leu Gly Thr Glu Ile Glu Val Asn Pro Lys Thr
145 150 155 160
Met Lys Ala Thr Gly Phe Val Lys Lys Pro Gly Val Leu Val Gly Asp
165 170 175
Leu Lys Arg Leu Ala Ile Leu Lys Glu Phe Gly Glu Glu Ser Pro Asp
180 185 190
Leu Gly Leu Gly Asp Arg Thr Ser Asp His Asp Phe Met Ser Ile Cys
195 200 205
Lys Glu Gly Tyr Met Val His Glu Ser Lys Ser Ala Thr Thr Val Pro
210 215 220
Ile Glu Arg Leu Lys Asn Arg Ile Ile Phe His Asp Gly Arg Leu Val
225 230 235 240
Gln Arg Pro Thr Pro Leu Asn Ala Ile Ile Ile Tyr Leu Trp Leu Pro
245 250 255
Phe Gly Phe Met Leu Ser Ile Phe Arg Val Tyr Phe Asn Leu Pro Leu
260 265 270
Pro Glu Arg Phe Val Arg Tyr Thr Tyr Glu Ile Leu Gly Ile His Leu
275 280 285
Thr Ile Arg Gly His Arg Pro Pro Pro Pro Ser Pro Gly Thr Pro Gly
290 295 300
Asn Leu Tyr Val Leu Asn His Arg Thr Ala Leu Asp Pro Ile Ile Ile
305 310 315 320
Ala Ile Ala Leu Gly Arg Lys Ile Ser Cys Val Thr Tyr Ser Val Ser
325 330 335
Arg Leu Ser Arg Met Leu Ser Pro Ile Pro Ala Val Ala Leu Thr Arg
340 345 350

pctau2012001598-seql

Asp Arg Ala Ala Asp Ala Ala Arg Met Arg Lys Leu Leu Glu Lys Gly
355 360 365

Asp Leu Val Ile Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro Tyr Leu
370 375 380

Leu Arg Phe Ser Ala Leu Phe Ala Glu Leu Ser Asp Arg Ile Val Pro
385 390 395 400

Val Ala Met Asn Cys Lys Gln Gly Met Phe Asn Gly Thr Thr Val Arg
405 410 415

Gly Val Lys Phe Trp Asp Pro Tyr Phe Phe Phe Met Asn Pro Arg Pro
420 425 430

Ser Tyr Glu Ala Thr Phe Leu Asp Arg Leu Pro Glu Glu Trp Thr Val
435 440 445

Asn Gly Gly Gly Lys Thr Pro Ile Glu Val Ala Asn Tyr Val Gln Lys
450 455 460

Val Ile Gly Gly Val Leu Gly Phe Glu Cys Thr Glu Leu Thr Arg Lys
465 470 475 480

Asp Lys Tyr Leu Leu Leu Gly Gly Asn Asp Gly Lys Val Glu Ser Ile
485 490 495

Asn Lys Ser Lys Ala Lys Glu
500

<210> 198
<211> 500
<212> PRT
<213> Populus trichocarpa

<400> 198

Met Ser Pro Thr Lys Pro Ala Lys Lys Phe Pro Pro Ile Thr Ala Cys
1 5 10 15

Asn Gly Thr Thr His Gln Ser Ile Ala Ala Asp Leu Asp Gly Thr Leu
20 25 30

Leu Val Ser Ser Ser Ser Phe Pro Tyr Phe Met Ile Val Ala Val Glu
35 40 45

Ala Gly Ser Leu Phe Arg Gly Leu Val Leu Leu Leu Ser Leu Pro Ile
50 55 60

Val Ile Val Ser Tyr Leu Phe Ile Ser Glu Ala Leu Gly Ile Gln Met
65 70 75 80

Leu Ile Phe Ile Ser Met Ser Gly Leu Lys Ile Arg Asp Ile Glu Leu
Page 258

Val Ser Arg Ala Val Leu Pro Arg Phe Tyr Ala Ala Asp Val Arg Ser
 100 105 110

Glu Ser Phe Glu Val Phe Asp Arg Cys Lys Arg Lys Val Val Val Thr
 115 120 125

Ala Asn Pro Thr Ile Met Val Glu Pro Phe Val Lys Asp Phe Leu Gly
 130 135 140

Gly Asp Lys Val Leu Gly Thr Glu Ile Glu Val Asn Pro Lys Thr Lys
 145 150 155 160

Arg Ala Thr Gly Phe Val Lys Lys Pro Gly Val Leu Val Gly Lys Trp
 165 170 175

Lys Glu Leu Ala Val Leu Lys Glu Phe Gly Glu Glu Ala Pro Asp Leu
 180 185 190

Gly Ile Gly Asp Arg Lys Thr Asp His Asp Phe Met Ser Leu Cys Lys
 195 200 205

Glu Gly Tyr Met Val His Arg Ser Lys Ser Ala Thr Pro Leu Pro Arg
 210 215 220

Asp Arg Leu Lys Asn Arg Ile Ile Phe His Asp Gly Arg Leu Val Gln
 225 230 235 240

Arg Pro Asp Pro Leu Asn Ala Leu Ile Thr Tyr Ile Trp Leu Pro Phe
 245 250 255

Gly Phe Ile Leu Ser Ile Ile Arg Val Tyr Phe Asn Leu Pro Leu Pro
 260 265 270

Glu Arg Ile Val Arg Tyr Thr Tyr Glu Met Leu Gly Ile His Leu Val
 275 280 285

Ile Arg Gly Thr Pro Pro Pro Ala Pro Ser Pro Gly Thr Pro Gly Asn
 290 295 300

Leu Tyr Val Cys Asn His Arg Thr Ala Leu Asp Pro Ile Val Ile Ala
 305 310 315 320

Ile Ala Leu Gly Arg Lys Val Ser Cys Val Thr Tyr Ser Val Ser Arg
 325 330 335

Leu Ser Arg Phe Leu Ser Pro Ile Pro Ala Ile Ala Leu Thr Arg Asp
 340 345 350

Arg Ala Ala Asp Ala Ala Arg Ile Ser Ser Ile Leu Gln Lys Gly Asp
 Page 259

355

360

365

Leu Val Val Cys Pro Glu Gly Thr Thr Cys Arg Glu Glu Phe Leu Leu
 370 375 380

Arg Phe Ser Ala Leu Phe Ala Glu Leu Ser Asp Arg Ile Val Pro Val
 385 390 395 400

Ala Val Asn Cys Lys Gln Asn Met Phe Tyr Gly Thr Thr Val Arg Gly
 405 410 415

Val Lys Phe Trp Asp Pro Tyr Phe Phe Phe Met Asn Pro Arg Pro Thr
 420 425 430

Tyr Glu Val Thr Phe Leu Asp Arg Leu Pro Glu Glu Met Thr Val Lys
 435 440 445

Ala Gly Gly Lys Ser Ser Ile Glu Val Ala Asn Tyr Val Gln Lys Val
 450 455 460

Leu Gly Glu Val Leu Gly Phe Glu Asn Thr Gly Leu Thr Arg Lys Asp
 465 470 475 480

Lys Tyr Leu Leu Leu Gly Gly Asn Asp Gly Lys Val Glu Ser Met His
 485 490 495

Asn Ala Lys Lys
 500

<210> 199
 <211> 518
 <212> PRT
 <213> Sorghum bicolor

<400> 199

Met Ala Pro Ala Ala Ser Ala Arg Phe Pro Pro Val Ser Ser Tyr Asp
 1 5 10 15

Ala Ser Ala Arg Thr Arg Arg Thr Ala Ala Ala Asp Leu Asp Gly Thr
 20 25 30

Leu Leu Ala Ser Ser Ser Ala Phe Pro Tyr Tyr Phe Leu Val Ala Leu
 35 40 45

Glu Ala Gly Ser Cys Leu Arg Ala Ala Ala Leu Leu Leu Val Ala Pro
 50 55 60

Leu Leu Leu Ala Leu Tyr Thr Leu Val Ser Glu Ala Ala Ala Ile Ala
 65 70 75 80

Leu Leu Ala Phe Val Thr Phe Ala Gly Leu Arg Val Arg Asp Val Glu
 85 90 95

pctau2012001598-seql

Ala Val Ala Arg Gly Val Leu Pro Arg His Tyr Ala Ala Gly Val Arg
100 105 110

Ala Asp Thr Trp Thr Val Phe Arg Gly Cys Gly Glu Gly Arg Arg Val
115 120 125

Val Val Thr Ala Ser Pro Ala Val Met Val Gly Glu Phe Val Arg Glu
130 135 140

Phe Leu Gly Ala Glu Val Ala Gly Thr Glu Leu Glu Thr Phe Ser Ala
145 150 155 160

Phe Gly Ala Ala Arg Phe Thr Gly Arg Ile Lys Ala Val Leu Val Gly
165 170 175

Glu Arg Lys Ala Glu Val Val Arg Arg Leu Phe Ala Gly Gly Glu Met
180 185 190

Pro Asp Val Gly Leu Gly Asp Arg Glu Ser Asp His Asp Phe Met Ala
195 200 205

Ile Cys Lys Glu Ala Tyr Met Val Pro Pro Asp Lys Arg Ala Ala Arg
210 215 220

Ala Ala Ala Asp Thr Leu Leu Ser Arg Ser Val Phe His Asp Gly Arg
225 230 235 240

Leu Val Gln Arg Pro Asp Pro Ala Gln Ala Leu Phe Ala Leu Ala Tyr
245 250 255

Leu Pro Leu Gly Phe Leu Leu Ala Leu Phe Arg Val Phe Phe Asn Leu
260 265 270

Met Met Pro Leu Arg Leu Val Arg His Thr Tyr Arg Leu Thr Gly Ile
275 280 285

Arg Leu Arg Val Arg Gly Thr Pro Pro Pro Pro Pro Ala Pro Gly Ala
290 295 300

Pro Gly Ser Leu Leu Val Cys Asn His Arg Thr Ala Leu Asp Pro Ile
305 310 315 320

Ile Leu Ser Val Ala Leu Gly Arg Pro Val Ser Cys Val Thr Tyr Ser
325 330 335

Ala Ser Arg Leu Ser Thr Ala Ile Ser Pro Ile Arg Ala Val Ala Leu
340 345 350

Ser Arg Asp Arg Ala Thr Asp Ala Ala Arg Met Ala Ala Leu Leu Ala
355 360 365

pctau2012001598-seql

Glu Gly Asp Val Val Val Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro
370 375 380

Cys Leu Leu Arg Phe Ser Ala Leu Phe Ala Glu Leu Thr Asp Arg Ile
385 390 395 400

Val Pro Val Ala Met Glu Ala Arg Gln Gly Thr Tyr Tyr Gly Ser Thr
405 410 415

Ala Arg Gly Trp Lys Trp Leu Asp Pro Tyr Phe Phe Tyr Met Asn Pro
420 425 430

Arg Pro Gly Tyr Asp Val Thr Phe Leu Pro Pro Leu Arg Pro Glu Glu
435 440 445

Thr Cys Gly Ala Gly Gly Arg Ser Ala Val Asp Val Ala Asn His Val
450 455 460

Gln Arg Val Ile Ala Lys Glu Leu Gly Phe Gln Cys Thr Thr Leu Thr
465 470 475 480

Arg Lys Asp Lys Tyr Met Lys Leu Ala Gly Asn Asp Gly Ser Val Ala
485 490 495

Ala Arg Ala Lys Lys Asp Ala Ala Asp Asp Asn Ala Ala Thr Val Thr
500 505 510

Thr Thr Lys Lys Phe Val
515

<210> 200
<211> 293
<212> PRT
<213> Ricinus communis

<400> 200

Met Ser Ile Cys Lys Glu Gly Tyr Met Val Pro Arg Thr Lys Cys Glu
1 5 10 15

Pro Leu Pro Arg Asn Lys Leu Leu Ser Pro Val Ile Phe His Glu Gly
20 25 30

Arg Leu Val Gln Arg Pro Thr Pro Leu Val Ala Leu Leu Thr Phe Leu
35 40 45

Trp Met Pro Ile Gly Ile Ile Leu Ser Ile Leu Arg Val Tyr Leu Asn
50 55 60

Ile Pro Leu Pro Glu Arg Ile Ala Tyr Tyr Asn Tyr Lys Ile Leu Gly
65 70 75 80

pctau2012001598-seql

Ile Lys Val Ile Val Lys Gly Thr Pro Pro Pro Ala Pro Arg Lys Gly
85 90 95

Gln Ser Gly Val Leu Phe Val Cys Asn His Arg Thr Val Leu Asp Pro
100 105 110

Val Val Thr Ala Val Ala Leu Gly Arg Lys Ile Ser Cys Val Thr Tyr
115 120 125

Ser Ile Ser Lys Phe Thr Glu Ile Ile Ser Pro Ile Lys Ala Val Ala
130 135 140

Leu Ser Arg Glu Arg Glu Lys Asp Ala Ala Asn Ile Lys Arg Leu Leu
145 150 155 160

Glu Glu Gly Asp Leu Val Ile Cys Pro Glu Gly Thr Thr Cys Arg Glu
165 170 175

Pro Phe Leu Leu Arg Phe Ser Ala Leu Phe Ala Glu Leu Thr Asp Arg
180 185 190

Ile Val Pro Val Ala Ile Asn Thr Lys Gln Thr Val Phe His Gly Thr
195 200 205

Thr Val Arg Gly His Lys Leu Leu Asp Pro Tyr Phe Val Phe Met Asn
210 215 220

Pro Met Pro Thr Tyr Glu Ile Thr Phe Leu Asn Gln Leu Pro Ile Glu
225 230 235 240

Leu Thr Cys Lys Gly Gly Lys Ser Ser Ile Glu Val Ala Asn Tyr Ile
245 250 255

Gln Arg Val Leu Ala Gly Thr Leu Gly Phe Glu Cys Thr Asn Leu Thr
260 265 270

Arg Lys Asp Lys Tyr Ala Ile Leu Ala Gly Thr Asp Gly Arg Val Pro
275 280 285

Ser Lys Lys Glu Lys
290

<210> 201
<211> 502
<212> PRT
<213> Arabidopsis lyrata

<400> 201

Met Ser Pro Ala Lys Lys Ser Arg Ser Phe Pro Pro Ile Ser Glu Cys
1 5 10 15

pctau2012001598-seql

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

pctau2012001598-seql

Ile Arg Gly His Arg Pro Pro Pro Pro Ser Pro Gly Lys Pro Gly Asn
 290 295 300

Leu Tyr Val Leu Asn His Arg Thr Ala Leu Asp Pro Ile Ile Ile Ala
 305 310 315 320

Ile Ala Leu Gly Arg Lys Ile Thr Cys Val Thr Tyr Ser Val Ser Arg
 325 330 335

Leu Ser Leu Met Leu Ser Pro Ile Pro Ala Val Ala Leu Thr Arg Asp
 340 345 350

Arg Val Ala Asp Ala Ala Arg Met Arg Gln Leu Leu Glu Lys Gly Asp
 355 360 365

Leu Val Ile Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro Tyr Leu Leu
 370 375 380

Arg Phe Ser Ala Leu Phe Ala Glu Leu Ser Asp Arg Ile Val Pro Val
 385 390 395 400

Ala Met Asn Cys Lys Gln Gly Met Phe Asn Gly Thr Thr Val Arg Gly
 405 410 415

Val Lys Phe Trp Asp Pro Tyr Phe Phe Phe Met Asn Pro Arg Pro Ser
 420 425 430

Tyr Glu Ala Thr Phe Leu Asp Arg Leu Pro Glu Glu Met Thr Val Asn
 435 440 445

Gly Gly Gly Lys Thr Pro Phe Glu Val Ala Asn Tyr Val Gln Lys Val
 450 455 460

Ile Gly Gly Val Leu Gly Phe Glu Cys Thr Glu Leu Thr Arg Lys Asp
 465 470 475 480

Lys Tyr Leu Leu Leu Gly Gly Asn Asp Gly Lys Val Glu Ser Ile Asn
 485 490 495

Lys Thr Lys Ser Met Asp
 500

<210> 202
 <211> 585
 <212> PRT
 <213> Arabidopsis thaliana

<400> 202

Met Val Leu Pro Glu Leu Leu Val Ile Leu Ala Glu Trp Val Leu Tyr
 1 5 10 15

Arg Leu Leu Ala Lys Ser Cys Tyr Arg Ala Ala Arg Lys Leu Arg Gly
 Page 265

Tyr Gly Phe Gln Leu Lys Asn Leu Leu Ser Leu Ser Lys Thr Gln Ser
 35 40 45

Leu His Asn Asn Ser Gln His His Leu His Asn His His Gln Gln Asn
 50 55 60

His Pro Asn Gln Thr Leu Gln Asp Ser Leu Asp Pro Leu Phe Pro Ser
 65 70 75 80

Leu Thr Lys Tyr Gln Glu Leu Leu Leu Asp Lys Asn Arg Ala Cys Ser
 85 90 95

Val Ser Ser Asp His Tyr Arg Asp Thr Phe Phe Cys Asp Ile Asp Gly
 100 105 110

Val Leu Leu Arg Gln His Ser Ser Lys His Phe His Thr Phe Phe Pro
 115 120 125

Tyr Phe Met Leu Val Ala Phe Glu Gly Gly Ser Ile Ile Arg Ala Ile
 130 135 140

Leu Leu Leu Leu Ser Cys Ser Phe Leu Trp Thr Leu Gln Gln Glu Thr
 145 150 155 160

Lys Leu Arg Val Leu Ser Phe Ile Thr Phe Ser Gly Leu Arg Val Lys
 165 170 175

Asp Met Asp Asn Val Ser Arg Ser Val Leu Pro Lys Phe Phe Leu Glu
 180 185 190

Asn Leu Asn Ile Gln Val Tyr Asp Ile Trp Ala Arg Thr Glu Tyr Ser
 195 200 205

Lys Val Val Phe Thr Ser Leu Pro Gln Val Leu Val Glu Arg Phe Leu
 210 215 220

Arg Glu His Leu Asn Ala Asp Asp Val Ile Gly Thr Lys Leu Gln Glu
 225 230 235 240

Ile Lys Val Met Gly Arg Lys Phe Tyr Thr Gly Leu Ala Ser Gly Ser
 245 250 255

Gly Phe Val Leu Lys His Lys Ser Ala Glu Asp Tyr Phe Phe Asp Ser
 260 265 270

Lys Lys Lys Pro Ala Leu Gly Ile Gly Ser Ser Ser Ser Pro Gln Asp
 275 280 285

His Ile Phe Ile Ser Ile Cys Lys Glu Ala Tyr Phe Trp Asn Glu Glu
 Page 266

pctau2012001598-seql

290

295

300

Glu Ser Met Ser Lys Asn Asn Ala Leu Pro Arg Glu Arg Tyr Pro Lys
 305 310 315 320

Pro Leu Ile Phe His Asp Gly Arg Leu Ala Phe Leu Pro Thr Pro Leu
 325 330 335

Ala Thr Leu Ala Met Phe Ile Trp Leu Pro Ile Gly Phe Leu Leu Ala
 340 345 350

Val Phe Arg Ile Ser Val Gly Val Phe Leu Pro Tyr His Val Ala Asn
 355 360 365

Phe Leu Ala Ser Met Ser Gly Val Arg Ile Thr Phe Lys Thr His Asn
 370 375 380

Leu Asn Asn Gly Arg Pro Glu Lys Gly Asn Ser Gly Val Leu Tyr Val
 385 390 395 400

Cys Asn His Arg Thr Leu Leu Asp Pro Val Phe Leu Thr Thr Ser Leu
 405 410 415

Gly Lys Pro Leu Thr Ala Val Thr Tyr Ser Leu Ser Lys Phe Ser Glu
 420 425 430

Phe Ile Ala Pro Leu Lys Thr Val Ser Leu Lys Arg Asp Arg Lys Lys
 435 440 445

Asp Gly Glu Ala Met Gln Arg Leu Leu Ser Lys Gly Asp Leu Val Val
 450 455 460

Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro Tyr Leu Leu Arg Phe Ser
 465 470 475 480

Pro Leu Phe Ala Glu Leu Thr Glu Asp Ile Val Pro Val Ala Val Asp
 485 490 495

Ala Arg Val Ser Met Phe Tyr Gly Thr Thr Ala Ser Gly Leu Lys Cys
 500 505 510

Leu Asp Pro Ile Phe Phe Leu Met Asn Pro Arg Pro Val Tyr Cys Leu
 515 520 525

Glu Ile Leu Lys Lys Leu Pro Lys Glu Met Thr Cys Ala Gly Gly Lys
 530 535 540

Ser Ser Phe Glu Val Ala Asn Phe Ile Gln Gly Glu Leu Ala Arg Val
 545 550 555 560

Leu Gly Phe Glu Cys Thr Asn Leu Thr Arg Arg Asp Lys Tyr Leu Val
 Page 267

565

570

575

Leu Ala Gly Asn Glu Gly Ile Val Arg
 580 585

<210> 203
 <211> 520
 <212> PRT
 <213> Arabidopsis thaliana

<400> 203

Met Ser Ala Lys Ile Ser Ile Phe Gln Ala Leu Val Phe Leu Phe Tyr
 1 5 10 15

Arg Phe Ile Leu Arg Arg Tyr Arg Asn Ser Lys Pro Lys Tyr Gln Asn
 20 25 30

Gly Pro Ser Ser Leu Leu Gln Ser Asp Leu Ser Arg His Thr Leu Ile
 35 40 45

Phe Asn Val Glu Gly Ala Leu Leu Lys Ser Asp Ser Leu Phe Pro Tyr
 50 55 60

Phe Met Leu Val Ala Phe Glu Ala Gly Gly Val Ile Arg Ser Phe Leu
 65 70 75 80

Leu Phe Ile Leu Tyr Pro Leu Ile Ser Leu Met Ser His Glu Met Gly
 85 90 95

Val Lys Val Met Val Met Val Ser Phe Phe Gly Ile Lys Lys Glu Gly
 100 105 110

Phe Arg Ala Gly Arg Ala Val Leu Pro Lys Tyr Phe Leu Glu Asp Val
 115 120 125

Gly Leu Glu Met Phe Glu Val Leu Lys Arg Gly Gly Lys Lys Ile Gly
 130 135 140

Val Ser Asp Asp Leu Pro Gln Val Met Ile Glu Gly Phe Leu Arg Asp
 145 150 155 160

Tyr Leu Glu Ile Asp Val Val Val Gly Arg Glu Met Lys Val Val Gly
 165 170 175

Gly Tyr Tyr Leu Gly Ile Met Glu Asp Lys Thr Lys His Asp Leu Val
 180 185 190

Phe Asp Glu Leu Val Arg Lys Glu Arg Leu Asn Thr Gly Arg Val Ile
 195 200 205

Gly Ile Thr Ser Phe Asn Thr Ser Leu His Arg Tyr Leu Phe Ser Gln
 210 215 220

pctau2012001598-seql

```

Phe Cys Gln Glu Ile Tyr Phe Val Lys Lys Ser Asp Lys Arg Ser Trp
225      230      235      240

Gln Thr Leu Pro Arg Ser Gln Tyr Pro Lys Pro Leu Ile Phe His Asp
      245      250      255

Gly Arg Leu Ala Ile Lys Pro Thr Leu Met Asn Thr Leu Val Leu Phe
      260      265      270

Met Trp Gly Pro Phe Ala Ala Ala Ala Ala Ala Arg Leu Phe Val
      275      280      285

Ser Leu Cys Ile Pro Tyr Ser Leu Ser Ile Pro Ile Leu Ala Phe Ser
      290      295      300

Gly Cys Arg Leu Thr Val Thr Asn Asp Tyr Val Ser Ser Gln Lys Gln
305      310      315      320

Lys Pro Ser Gln Arg Lys Gly Cys Leu Phe Val Cys Asn His Arg Thr
      325      330      335

Leu Leu Asp Pro Leu Tyr Val Ala Phe Ala Leu Arg Lys Lys Asn Ile
      340      345      350

Lys Thr Val Thr Tyr Ser Leu Ser Arg Val Ser Glu Ile Leu Ala Pro
      355      360      365

Ile Lys Thr Val Arg Leu Thr Arg Asp Arg Val Ser Asp Gly Gln Ala
      370      375      380

Met Glu Lys Leu Leu Thr Glu Gly Asp Leu Val Val Cys Pro Glu Gly
385      390      395      400

Thr Thr Cys Arg Glu Pro Tyr Leu Leu Arg Phe Ser Pro Leu Phe Thr
      405      410      415

Glu Val Ser Asp Val Ile Val Pro Val Ala Val Thr Val His Val Thr
      420      425      430

Phe Phe Tyr Gly Thr Thr Ala Ser Gly Leu Lys Ala Leu Asp Pro Leu
      435      440      445

Phe Phe Leu Leu Asp Pro Tyr Pro Thr Tyr Thr Ile Gln Phe Leu Asp
      450      455      460

Pro Val Ser Gly Ala Thr Cys Gln Asp Pro Asp Gly Lys Leu Lys Phe
465      470      475      480

Glu Val Ala Asn Asn Val Gln Ser Asp Ile Gly Lys Ala Leu Asp Phe
      485      490      495

```

pctau2012001598-seql

Glu Cys Thr Ser Leu Thr Arg Lys Asp Lys Tyr Leu Ile Leu Ala Gly
500 505 510

Asn Asn Gly Val Val Lys Lys Asn
515 520

<210> 204
<211> 1330
<212> DNA
<213> Arabidopsis thaliana

<400> 204
agttaaaaga ttggttattt gggctctgca ctcaagtgag agagaagata gatagatctg 60
agtagaatct tcgattcatt attcgttgtc gtcgttcac tgtgagaagc ggacaaacca 120
aagaatccac cggagctagt gatatgggtg gttccagaga gttccgagct gaggaacatt 180
caaatcaatt ccactctatc atcgccatgg ccatctggct tggcgccatt cacttcaacg 240
tcgctcttgt tctctgttct ctcatcttcc ttctctcttc tctatctctc atggtcttgg 300
gcttgctctc tctgtttatc tttatcccaa tcgatcatcg tagcaaatat ggtcgtaagc 360
tcgctaggta catatgcaag cacgcgtgta attatttccc cgtctctctg tacgtcgagg 420
attacgaagc tttccagcct aatcgtgcct atgtctttgg ttatgaacca cattcgggtgc 480
taccgattgg agttgttgc ttttgtgatc tcacagggtt tatgcctatt cctaacatta 540
aagttcttgc aagtagtgct atattctaca ctccctttct aaggcatata tggacatggt 600
tagggctcac cgctgcttct aggaagaatt tcacttcctt tttggattct ggctacagtt 660
gtgttcttgt acctgggtgt gtgcaggaga cttttcatat gcaacatgat gctgagaatg 720
tcttcctttc aaggagaaga ggatttgtgc gcatagccat ggaacagggg agccctctgg 780
ttccagtatt ctgctttggt caggcacgcg tgtacaaatg gtggaagccg gattgtgatc 840
tctatcttaa actatctaga gcaatcagat tcaccccgat ctgcttctgg ggagtttttg 900
gatcaccatt accgtgtcga cagcctatgc atgtggctgt tggtaaacca atagaagtca 960
caaaaactct gaagccaact gacgaagaga ttgctaagtt tcatggccag tatgtggaag 1020
cgcttaggga tctgtttgag aggcacaagt cccgagtcgg ctatgatctt gagctgaaaa 1080
ttctttgaac aaaatctcca atggaaataa ttacttgtgt gtatccttca ttaattgtta 1140
ccttggagct ggatttggac ttaatataaa tgactacatc atgtagtcta catgtattgc 1200
atgtctttag catcgactgt tgaagtaatg gaatacgttt ataaagcctg taaattacat 1260
gtcgtcttgc acaagagtat gtggtaataa taacatttga cccaaaaata atactagtta 1320
aatttttcct 1330

<210> 205
<211> 1381
<212> DNA
<213> Ricinus communis

pctau2012001598-seql

<220>

<221> misc_feature

<222> (1347)..(1347)

<223> n is a, c, g, or t

<400> 205

```

actggagcac gaggacactg acatggactg aaggagtaga aaagggacgc aaaaaaagat      60
tttacatggc ggagaggtcg atgctgcttt attttcatta atacagaaag agtttgagca      120
ctgagcactg aatagcaaga agaagaagaa gaagaaatgg gggaagaagc gaatcataat      180
aataataata ataatatcaa tagtaatgat gagaagaatg aagagaaatc aaattataca      240
gttgtaaatt cgagagaact atacccaacg aacatatattc acgcactgtt agcggtgagc      300
atatggattg gttcaatcca tttcaatctc ttcttactct tcatctctta tctcttcctt      360
tcttttccca cattcctcct gattgttgga ttttttgtgg tgtaaatgtt cattccgatac      420
gacgaacaca gtaagttggg ccgtcgtttg tgcaggtatg tatgcagaca tgcgtgcagt      480
cattttccgg taactctcca tgttgaagac atgaatgctt ttcattctga tcgtgcttac      540
gtttttgggt atgagccaca ttcagtattt ccccttggtg tttctgtact atcagatcac      600
tttgctgtcc tgcccccttc taaaatgaag gtccttgcaa gtaacgctgt gtttcggaca      660
ccagttttta ggcatatatg gacatgggtg ggtcttacat cagcaacaaa gaaaaatttc      720
actgccctcc tagcatctgg ttatagttgc attgtgattc ccggtggagt tcaagagaca      780
ttttatatga agcatggctc tgagattgct ttccttaagg cgagaagagg gtttgtccga      840
gtagctatgg agatgggtaa acccttggtt ccagttttct gctttggtca atcgaacgtg      900
tacaagtggg ggaaacctga tggcgagtta tttatgaaaa ttgctagagc tattaagttc      960
agcccaattg tcttttgggg agttctcggt tctcatctac cgctacaacg tccaatgcat     1020
gttgctcgtc gtaaaccgat tgaggtgaag caaaatccac agcctacagt ggaagaggtc     1080
tcagaagtac agggtcagtt tgttgcgga cttaaagatc tttttgaaag gcataaagca     1140
cgggttggtc atgcagacct tacacttgaa attctttgat agaagtcagt acttgtttta     1200
gttcagttgg atgcctcttt cagtttttgg cgataatgtt tgtattcatt agataattag     1260
gaaaatgaaa tgtaaaacgg tgaataaagc ttttcaatat gcgggcatca caaagttggt     1320
ctagatatcc aaaaaaaaaa aaaaacnctg tatgccgtta cgtaccgtat cgttgacagc     1380
a                                                                    1381

```

<210> 206

<211> 1257

<212> DNA

<213> Vernicia fordii

<400> 206

```

agaagggcga tgcacatgat actctcctct taataagagt attgataccc agagaagcca      60
gaagcaagaa acaacaact tttgcaagaa ggagattaaa atggggatgg tggaagttaa      120
gaatgaagaa gaggttacca tattcaaatc tggagaaata tacccaacaa acatttttca      180

```


pctau2012001598-seq1

atcagtgtta gctttggcta tatggcttgg ctctttccat ttcattcctct tcttagtctc	240
ttcttcaatc ttccttcctt tctccaaatt tctcctcgtt attgggttgc ttttgttttt	300
tatggtgatt cctatcaacg atcgagtaa gcttggtcaa tgcttggtca gctatatatc	360
taggcatgtt tgcagctatt ttcccatcac tcttcatgtt gaagatataa atgcctttcg	420
ttctgatcgt gcttacgtgt ttggttatga gccacattca gtttttccca ttggtgtgat	480
gattctttct ttgggtttta ttcctcttcc taatataaaa ttccttgcaa gtagtgctgt	540
gttctacact ccatttttaa gacatatatg gtcattggtt ggtcttacac cagcaacaag	600
gaaaaatttt gtttcctcc tgtcttctgg ttatagttgc attctggtgc ctggtggagt	660
ccaagagaca ttttatatga agcaggactc tgagattgct ttccttaagg caagaagagg	720
atttattcga atagctatgc agacgggcac acccttggtt cctgttttct gctttggtca	780
aatgcacaca ttcaaattgt ggaaacctga tggagagttg ttcattgaaa ttgctagagc	840
tattaaattc accccaacta tcttttgggg agttctcgga actcctttac cttcaaaaa	900
tccaatgcat gtcgtggtgg gtagacccat tgaggtgaag caaaatccac agcctaccgc	960
tgaagagggt gcggaagtac agagagagtt tattgcatca cttaaaaatc tctttgaaag	1020
gcacaaagca cgggtcggct attcagacct taaacttgaa attttttgac cagaaatcag	1080
aacttaaaaag tagttttgca tgctttctca ttttcattta tttagttttt ttttaaacct	1140
ttttttttcc cttttgttaa tttgtgttct tagataatga tgaaaatgag aaaaaaaaaa	1200
aatttgtaaa acagctaata agccttaatt ttgtggcaaa aaaaaaaaaa aaaaaaa	1257

<210> 207
 <211> 1280
 <212> DNA
 <213> *Mortierella ramanniana*

<220>
 <221> misc_feature
 <222> (1265)..(1266)
 <223> n is a, c, g, or t

<220>
 <221> misc_feature
 <222> (1269)..(1269)
 <223> n is a, c, g, or t

<220>
 <221> misc_feature
 <222> (1275)..(1275)
 <223> n is a, c, g, or t

<220>
 <221> misc_feature
 <222> (1277)..(1277)
 <223> n is a, c, g, or t

<220>
 <221> misc_feature
 <222> (1279)..(1279)
 <223> n is a, c, g, or t

pctau2012001598-seql

<400> 207
 tgcaaagtca cactttcagt cagccatggc cagcaaggat caacatttac agcagaaggt 60
 caagcatacg ctagaagcta tcccatcccc tcgctatgct ccattgcgag tgccattaag 120
 acggagatta caaacattgg cagttttatt atgggtgttc atgatgtcaa tatgcatggt 180
 catattcttc tttttatgct ccattcctgt tctcctttgg ttccccatta tcctttattt 240
 gacctggatc ttgggtgtggg ataaggcgcc agagaacggt ggaagaccta ttcgctggct 300
 gcggaatgct gcttgggtga agctgtttgc agggatatttt cccgcacatg tcatcaagga 360
 agccgattta gatccatcca agaactacat ctttggttat ccccccatg gaatcatatc 420
 catgggctcg ttctgtactt ttagtaccaa tgctactggc ttgatgact tgttcccagg 480
 catccggcca tcgcttttga cattaacatc taattttaat atcccacttt atcgtgatta 540
 tttgatggcg tgcggacttt gctccgtctc caaaacatcc tgtcaaaata ttttaaccaa 600
 aggtgggtccg ggccgttcca ttgccattgt cgtgggaggt gcttccgagt ctctcaatgc 660
 tagaccgggt gtcattggacc ttgtgttgaa gagacgcttt ggttttatca agattgctgt 720
 tcaaaccgggt gcaagtctag tgccactat cagttttggt gaaaatgagc tgtacgaaca 780
 gattgaaagc aatgaaaact caaagttgca tagatggcaa aagaagattc aacatgccct 840
 tggttttact atgccgctct ttcattggac cggtgtattc aattatgact ttggtttgct 900
 ccccatcgc catcctatct acacgattgt tggaaagccc atcccgtcc ctagcatcaa 960
 gtatggacag acaaaggatg agattataag agaactacat gactcgtaca tgcattgccgt 1020
 gcaggatctc tatgatcgtt acaaggatat ctatgcaaag gatcgggtaa aagaactaga 1080
 attcgtcgaa tagttgggtt gttaccattt aacgttttct tcaaattttt cattgattat 1140
 accgcttcat cccgacctgg ttttttattg aatgggtttc tcattctccag ttcaattccc 1200
 actcctgcat gtatctatta tcattcatcct ttctaataat gaaaaccttt tgaacgtttc 1260
 ataannaana gaaananana 1280

<210> 208
 <211> 2485
 <212> DNA
 <213> Homo sapiens

<400> 208
 tgccccgttg tgaggtgata aagtgttgcg ctccgggacg ccagcgccgc ggctgccgcc 60
 tctgctgggg tctaggctgt ttctctcgcg ccaccactgg ccgccggccg cagctccagg 120
 tgcctagcc gccagcctc gacgccgtcc cgggacctct gtgctctgcg cgaagccctg 180
 gccccggggg ccggggcatg ggccaggggc gcggggtgaa gcggcttccc gcggggccgt 240
 gactgggcgg gcttcagcca tgaagaccct catagccgcc tactccgggg tcctgcgcgg 300
 cgagcgtcag gccgaggctg accggagcca gcgctctcac ggaggacctg cgctgtcgcg 360
 cgaggggtct gggagatggg gcactggatc cagcatcctc tccgccctcc aggacctctt 420
 ctctgtcacc tggctcaata ggtccaaggt ggaaaagcag ctacagggtca tctcagtgtc 480

pctau2012001598-seq1

ccagtgggtc	ctgtccttcc	ttgtactggg	agtggcctgc	agtgccatcc	tcatgtacat	540
attctgcact	gattgctggc	tcatcgctgt	gctctacttc	acttggctgg	tgtttgactg	600
gaacacaccc	aagaaaggtg	gcaggaggtc	acagtgggtc	cgaaactggg	ctgtgtggcg	660
ctactttcga	gactactttc	ccatccagct	ggtgaagaca	cacaacctgc	tgaccaccag	720
gaactatatc	tttggatacc	accccatgg	tatcatgggc	ctgggtgcct	tctgcaactt	780
cagcacagag	gccacagaag	tgagcaagaa	gttcccaggc	atacggcctt	acctggctac	840
actggcaggc	aacttccgaa	tgcctgtgtt	gagggagtac	ctgatgtctg	gaggtatctg	900
ccctgtcagc	cgggacacca	tagactatth	gctttcaaag	aatgggagtg	gcaatgctat	960
catcatcgtg	gtcgggggtg	cggctgagtc	tctgagctcc	atgcctggca	agaatgcagt	1020
caccctgcgg	aaccgcaagg	gctttgtgaa	actggccctg	cgtcatggag	ctgacctggt	1080
tcccatctac	tcctttggag	agaatgaagt	gtacaagcag	gtgatcttcg	aggagggctc	1140
ctggggccga	tgggtccaga	agaagttcca	gaaatacatt	ggtttcgccc	catgcatctt	1200
ccatggtcga	ggcctcttct	cctccgacac	ctgggggctg	gtgccctact	ccaagcccat	1260
caccactgtt	gtgggagagc	ccatcaccat	ccccaagctg	gagcacccaa	cccagcaaga	1320
catcgacctg	taccacacca	tgtacatgga	ggccctgggt	aagctcttcg	acaagcacia	1380
gaccaagttc	ggcctcccgg	agactgaggt	cctggaggtg	aactgagcca	gccttcgggg	1440
ccaattccct	ggaggaacca	gctgcaaadc	acttttttgc	tctgtaaatt	tggaagtgtc	1500
atgggtgtct	gtgggttatt	taaaagaaat	tataacaatt	ttgctaaacc	attacaatgt	1560
taggtctttt	ttaagaagga	aaaagtcagt	atttcaagtt	ctttcacttc	cagcttgccc	1620
tgttctaggt	ggtggctaaa	tctgggccta	atctgggtgg	ctcagctaac	ctctcttctt	1680
cccttcctga	agtgacaaag	gaaactcagt	cttcttgggg	aagaaggatt	gccattagtg	1740
acttggacca	gtagatgat	tcactttttg	cccctagggg	tgagaggcga	aagccacttc	1800
tcatacaagc	ccctttattg	ccactacccc	acgctcgtct	agtcctgaaa	ctgcaggacc	1860
agtttctctg	ccaaggggag	gagttggaga	gcacagttgc	cccgttgtgt	gagggcagta	1920
gtaggcatct	ggaatgctcc	agtttgatct	cccttctgcc	accctacct	caccctagt	1980
cactcatatc	ggagcctgga	ctggcctcca	ggatgaggat	gggggtggca	atgacaccct	2040
gcaggggaaa	ggactgcccc	ccatgcacca	ttgcagggag	gatgccgcca	ccatgagcta	2100
ggtggagtaa	ctggtttttc	ttgggtggct	gatgacatgg	atgcagcaca	gactcagcct	2160
tggcctggag	cacatgctta	ctggtggcct	cagtttacct	tcccagatc	ctagattctg	2220
gatgtgagga	agagatccct	cttcagaagg	ggcctggcct	tctgagcagc	agattagttc	2280
caaagcaggt	ggcccccgaa	cccaagcctc	acttttctgt	gccttcctga	gggggttggg	2340
ccggggagga	aaccacccc	tctcctgtgt	gttctgttat	ctcttgatga	gatcattgca	2400
ccatgtcaga	cttttgtata	tgccttgaaa	ataaatgaaa	gtgagaatcc	tcaaaaaaaaa	2460
aaaaaaaaaa	aaaaaaaaaa	aaaaa				2485

pctau2012001598-seq1

<210> 209
 <211> 1411
 <212> DNA
 <213> Homo sapiens

<400> 209
 actgttctga gatctttgcc tccctcaggc tcccgagaat catgggtcat tccaagcagc 60
 ctagtcaactt ccagagtctg atgcttctgc agtggccttt gagctacctt gccatctttt 120
 ggatcttgca gccattgttc gtctacctgc tgtttacatc cttgtggccg ctaccagtgc 180
 ttacttttgc ctggttggtc ctggactgga agaccccaga gcgaggtggc aggcgttcgg 240
 cctgggtaag gaactggtgt gtctggaccc acatcagga ctatttcccc attacgatcc 300
 tgaagacaaa ggacctatca cctgagcaca actacctcat ggggggttcac ccccatggcc 360
 tcctgacctt tggcgccttc tgcaacttct gactgaggc cacaggcttc tcgaagacct 420
 tcccaggcat cactcctcac ttggccacgc tgtcctgggt cttcaagatc ccctttgtta 480
 gggagtacct catggccaaa ggtgtgtgct ctgtgagcca gccagccatc aactatctgc 540
 tgagccatgg cactggcaac ctcgtgggca ttgtagtggg aggtgtgggt gaggccctgc 600
 aaagtgtgcc caacaccacc accctcatcc tccagaagcg caaggggttc gtgcgcacag 660
 ccctccagca tgggggtcat ctggtcccca ctttactttt tggggaaact gaggtgtatg 720
 atcaggtgct gttccataag gatagcagga tgtacaagtt ccagagctgc ttccgccgta 780
 tctttggttt ctactgttgt gtcttctatg gacaaagctt ctgtcaaggc tccactgggc 840
 tcctgccata ctccaggcct attgtcactg tggttgggga gcctctgcca ctgccccaaa 900
 ttgaaaagcc aagccaggag atggtggaca aataccatgc actttatatg gatgctctgc 960
 aaaaactgtt cgaccagcat aagaccact atggctgctc agagacccaa aagctgtttt 1020
 tcctgtgaat gaaggtactg catgcccagg agcacaggag tgcttgcctt gaagaagaga 1080
 ctcatctgcc actaaccaaa gacaggcagg agatgagga ggttatatgt ggtaggggag 1140
 ggcatgagga attccttctt tgccttcttg ccacagggtc cttacaggaa ttctttctga 1200
 agagctgcac acagtcattc ctcaaaggag ggcaccttag tgccctcat gctggggcct 1260
 gatgcctggt catcatttga gtcctctggg acacattagc agtcaactgca tattccccca 1320
 gccctgggc aactatctac tttctgtctc tgtggatttg ccaattctga acattttata 1380
 taaatggaat aatacaataa aaaaaaaaaa a 1411

<210> 210
 <211> 2488
 <212> DNA
 <213> Bos taurus

<400> 210
 gccgtggccg tggccaaggc tccgctttcc gccggtgctc cggcccgtct ggcggcccgt 60
 ccctccgcac cctcggccgt agcgggtccc gcgctcccct gctcggcgcg aagccctggc 120
 cccggcggcc ggggcatggg ccaggggcgc ggggtctggc ggcttcccgc ggggccccga 180

pctau2012001598-seq1

cctgtactgg	cttcgtcatg	aagaccctca	tagccgccta	ctccggggtc	ctgcgaggca	240
ctgggtccag	catcctctct	gccctccagg	acctgttttc	tgtcacttgg	ctcaataggt	300
ccaaggtaga	gaagcagctc	caagtcatct	cggtgctaca	atgggtcctg	tctttcctcg	360
tgctgggagt	ggcctgcagc	gtcatcctca	tgtacacatt	ctgcaccgat	tgctgggtca	420
ttgccgtgct	ctacttcacc	tggctgggtg	ttgactggaa	cacaccaag	aaaggtggca	480
ggaggtcaca	gtgggtccga	aactgggctg	tgtggcgcta	ctttcgagac	tactttccca	540
ttcagctggg	gaagacacac	aacttactga	ccagcaggaa	ctacatcttt	gggtaccatc	600
cccatggcat	catgggcctg	ggtgccttct	gcaacttcag	cacagaggcc	acagaagtaa	660
gcaagaagtt	ccctggcata	aggccctacc	tggccacgct	ggccggcaac	ttccggatgc	720
cagtgtgcg	ggagtacctg	atgtctggag	gcatctgccc	agtgaaccgg	gacaccatag	780
actacttgct	ttcaaagaat	gggagtggca	atgccatcat	catcgtgggtg	gggggcgcgg	840
ctgaatccct	gagctccatg	cccggcaaga	atgcagtcac	cctgcgcaat	cgcaagggtc	900
ttgtgaaact	ggccctgcgc	catggagccg	acctggttcc	cacctactcc	tttggggaga	960
atgaggtgta	caagcagggtg	atctttgagg	agggtcctg	gggccgggtg	gtgcagaaga	1020
agttccagaa	gtacattggc	tttgcccat	gcatcttcca	tggtcgaggc	ctcttctcct	1080
ctgacacctg	ggggctgggtg	ccctactcca	agcccatcac	cactgtcgtg	ggcgagccca	1140
ttaccatccc	caggctggag	cgcccgacgc	agcaggacat	cgacctgtac	cacgccatgt	1200
acgtgcaagc	cctggtgaag	ctcttcgacc	agcataagac	caagttcggc	ctcccggaga	1260
ccgaggtcct	ggaggtgaac	tgagcctgtc	tgcaggggcc	agcttctggg	aggaaccatc	1320
tgcaaatcgt	tttctaccaa	gttctcaagt	gctttttgtt	ctgtaaattt	ggaagcatca	1380
tgggtgtctg	tgggttatatt	aaaagaaatt	ataataattt	tgttaaaccg	ttacaatggt	1440
aggtcttttt	taagaaggag	aaagggaata	tttcaagctc	tctcagttcc	agtttgtcct	1500
gttccagggtg	ggcgtgccca	catctggggc	tttatggttt	cacaaccctt	cttctctcct	1560
cccccaaatt	acagagaaaa	ttcagtcctg	gttaactggg	gaagaaggac	agtcattagt	1620
gactcggggc	agttagatta	tttgcccttt	ggccctgagg	gatgaggggc	agaagcctcc	1680
tctagtacaa	acatctctat	gccagctac	ccaaggcttg	actgcaggac	taaaccctct	1740
tgccaaggaa	agacacagag	gaaacagggtg	ttcccatgct	gaagagtgca	gggataagca	1800
cctggaatat	tcaagcttga	ttctcctttc	tgccaccccc	atggccccag	tctctgaaat	1860
ctgagcctgg	actggcctcc	aaagagagga	tcagggtggc	agtgttcctg	tgctggagaa	1920
tggagctgtc	cagatgaacc	atctcttggg	ttatcaagcc	ccatcatgtt	ctggttggag	1980
tgactggttt	tcctcagggg	gctgatgaca	cagaccaga	cccagcacag	accctgcctt	2040
ggcctgccag	tggcctcgct	ttgcttaacc	cacatcctca	gcagtggatg	cgaggaaggg	2100
gtcccccttc	ctgtttacag	aggggcctga	gcagtagact	ggtaccagaa	cagaagttcc	2160
ccgaacccaa	acctcatgta	tttgtgcctt	ttctgagagg	ggggccaggg	gggacagcta	2220

pctau2012001598-seq1

gtactctgtg tactctgctt tctcttgatg agatcattat accatgtcag acttttcgtat	2280
attcctaaat ggataaatga aaacaagtgt cctctatgag ttattgttag ggtcacatta	2340
gcatcttgct gctgatgatt ggggagacca ctgccccagc tgggctgcct gctaagccct	2400
ctcccacttt ctcttcagc ccaaattggcc ctttgtatat tagttatcaa ttgctgtgta	2460
ataaattact ccaaaacttg gtggctgc	2488

<210> 211
 <211> 1330
 <212> DNA
 <213> Mus musculus

<400> 211	
ggtcaggggc gcggcgtgag gcggctttct gcacggccgt gacgtgcatt ggcttcagca	60
tgaagaccct catcgccgcc tactccgggg tcctgcgggg tgagcgtcgg gcggaagctg	120
cccgcagcga aaacaagaat aaaggatctg ccctgtcacg cgaggggtct gggcgatggg	180
gcactggctc cagcatcctc tcagccctcc aagacatctt ctctgtcacc tggctcaaca	240
gatctaaggt ggaaaaacag ctgcaggcca tctcagtact acaatgggtc ctatccttcc	300
tggtgctagg agtggcctgc agtgtcatcc tcatgtacac cttctgcaca gactgctggc	360
tgatagctgt gctctacttc acctggctgg catttgactg gaacacgccc aagaaagggtg	420
gcaggagatc gcagtgggtg cgaaactggg ccgtgtggcg ctacttccga gactactttc	480
ccatccagct ggtgaagaca cacaacctgc tgaccaccag gaactatatc tttggatacc	540
accccatgg catcatgggc ctgggtgcct tctgtaactt cagcacagag gctactgaag	600
tcagcaagaa gtttcctggc ataaggccct atttggctac gttggctggt aacttccgga	660
tgctgtgct tcgcgagtac ctgatgtctg gaggcattct ccctgtcaac cgagacacca	720
tagactactt gctctccaag aatgggagtg gcaatgctat catcatcgtg gtgggaggtg	780
cagctgagtc cctgagctcc atgcctggca agaacgcagt caccctgaag aaccgcaaag	840
gctttgtgaa gctggccctg cgccatggag ctgatctggt tcccacttat tcctttggag	900
agaatgaggt atacaagcag gtgatctttg aggagggttc ctggggccga tgggtccaga	960
agaagttcca gaagtatat ggtttcgccc cctgcatctt ccatggccga ggcctcttct	1020
cctctgacac ctgggggctg gtgccctact ccaagcccat caccaccgtc gtgggggagc	1080
ccatcactgt cccaagctg gagcaccga cccagaaaga catcgacctg taccatgcca	1140
tgtacatgga ggccctgggtg aagctctttg acaatcacia gaccaaattt ggccttccag	1200
agactgaggt gctggaggtg aactgacca gccgcgcgt gccagctcct gggagggagc	1260
actgcagatc cttttctacc gagttctcga gtgcattttg ttctgtaaat ttggaagcgt	1320
catgggtgtc	1330

<210> 212
 <211> 314
 <212> PRT
 <213> Arabidopsis thaliana

pctau2012001598-seql

<400> 212

Met Gly Gly Ser Arg Glu Phe Arg Ala Glu Glu His Ser Asn Gln Phe
1 5 10 15

His Ser Ile Ile Ala Met Ala Ile Trp Leu Gly Ala Ile His Phe Asn
20 25 30

Val Ala Leu Val Leu Cys Ser Leu Ile Phe Leu Pro Pro Ser Leu Ser
35 40 45

Leu Met Val Leu Gly Leu Leu Ser Leu Phe Ile Phe Ile Pro Ile Asp
50 55 60

His Arg Ser Lys Tyr Gly Arg Lys Leu Ala Arg Tyr Ile Cys Lys His
65 70 75 80

Ala Cys Asn Tyr Phe Pro Val Ser Leu Tyr Val Glu Asp Tyr Glu Ala
85 90 95

Phe Gln Pro Asn Arg Ala Tyr Val Phe Gly Tyr Glu Pro His Ser Val
100 105 110

Leu Pro Ile Gly Val Val Ala Leu Cys Asp Leu Thr Gly Phe Met Pro
115 120 125

Ile Pro Asn Ile Lys Val Leu Ala Ser Ser Ala Ile Phe Tyr Thr Pro
130 135 140

Phe Leu Arg His Ile Trp Thr Trp Leu Gly Leu Thr Ala Ala Ser Arg
145 150 155 160

Lys Asn Phe Thr Ser Leu Leu Asp Ser Gly Tyr Ser Cys Val Leu Val
165 170 175

Pro Gly Gly Val Gln Glu Thr Phe His Met Gln His Asp Ala Glu Asn
180 185 190

Val Phe Leu Ser Arg Arg Arg Gly Phe Val Arg Ile Ala Met Glu Gln
195 200 205

Gly Ser Pro Leu Val Pro Val Phe Cys Phe Gly Gln Ala Arg Val Tyr
210 215 220

Lys Trp Trp Lys Pro Asp Cys Asp Leu Tyr Leu Lys Leu Ser Arg Ala
225 230 235 240

Ile Arg Phe Thr Pro Ile Cys Phe Trp Gly Val Phe Gly Ser Pro Leu
245 250 255

Pro Cys Arg Gln Pro Met His Val Val Val Gly Lys Pro Ile Glu Val
Page 278

pctau2012001598-seql

260

265

270

Thr Lys Thr Leu Lys Pro Thr Asp Glu Glu Ile Ala Lys Phe His Gly
 275 280 285

Gln Tyr Val Glu Ala Leu Arg Asp Leu Phe Glu Arg His Lys Ser Arg
 290 295 300

Val Gly Tyr Asp Leu Glu Leu Lys Ile Leu
 305 310

<210> 213

<211> 340

<212> PRT

<213> Ricinus communis

<400> 213

Met Gly Glu Glu Ala Asn His Asn Asn Asn Asn Asn Ile Asn Ser
 1 5 10 15

Asn Asp Glu Lys Asn Glu Glu Lys Ser Asn Tyr Thr Val Val Asn Ser
 20 25 30

Arg Glu Leu Tyr Pro Thr Asn Ile Phe His Ala Leu Leu Ala Leu Ser
 35 40 45

Ile Trp Ile Gly Ser Ile His Phe Asn Leu Phe Leu Leu Phe Ile Ser
 50 55 60

Tyr Leu Phe Leu Ser Phe Pro Thr Phe Leu Leu Ile Val Gly Phe Phe
 65 70 75 80

Val Val Leu Met Phe Ile Pro Ile Asp Glu His Ser Lys Leu Gly Arg
 85 90 95

Arg Leu Cys Arg Tyr Val Cys Arg His Ala Cys Ser His Phe Pro Val
 100 105 110

Thr Leu His Val Glu Asp Met Asn Ala Phe His Ser Asp Arg Ala Tyr
 115 120 125

Val Phe Gly Tyr Glu Pro His Ser Val Phe Pro Leu Gly Val Ser Val
 130 135 140

Leu Ser Asp His Phe Ala Val Leu Pro Leu Pro Lys Met Lys Val Leu
 145 150 155 160

Ala Ser Asn Ala Val Phe Arg Thr Pro Val Leu Arg His Ile Trp Thr
 165 170 175

Trp Cys Gly Leu Thr Ser Ala Thr Lys Lys Asn Phe Thr Ala Leu Leu
 180 185 190

pctau2012001598-seql

Ala Ser Gly Tyr Ser Cys Ile Val Ile Pro Gly Gly Val Gln Glu Thr
195 200 205

Phe Tyr Met Lys His Gly Ser Glu Ile Ala Phe Leu Lys Ala Arg Arg
210 215 220

Gly Phe Val Arg Val Ala Met Glu Met Gly Lys Pro Leu Val Pro Val
225 230 235 240

Phe Cys Phe Gly Gln Ser Asn Val Tyr Lys Trp Trp Lys Pro Asp Gly
245 250 255

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

Phe Trp Gly Val Leu Gly Ser His Leu Pro Leu Gln Arg Pro Met His
275 280 285

Val Val Val Gly Lys Pro Ile Glu Val Lys Gln Asn Pro Gln Pro Thr
290 295 300

Val Glu Glu Val Ser Glu Val Gln Gly Gln Phe Val Ala Ala Leu Lys
305 310 315 320

Asp Leu Phe Glu Arg His Lys Ala Arg Val Gly Tyr Ala Asp Leu Thr
325 330 335

Leu Glu Ile Leu
340

<210> 214
<211> 322
<212> PRT
<213> Vernicia fordii

<400> 214

Met Gly Met Val Glu Val Lys Asn Glu Glu Glu Val Thr Ile Phe Lys
1 5 10 15

Ser Gly Glu Ile Tyr Pro Thr Asn Ile Phe Gln Ser Val Leu Ala Leu
20 25 30

Ala Ile Trp Leu Gly Ser Phe His Phe Ile Leu Phe Leu Val Ser Ser
35 40 45

Ser Ile Phe Leu Pro Phe Ser Lys Phe Leu Leu Val Ile Gly Leu Leu
50 55 60

Leu Phe Phe Met Val Ile Pro Ile Asn Asp Arg Ser Lys Leu Gly Gln
65 70 75 80

pctau2012001598-seql

Cys Leu Phe Ser Tyr Ile Ser Arg His Val Cys Ser Tyr Phe Pro Ile
85 90 95

Thr Leu His Val Glu Asp Ile Asn Ala Phe Arg Ser Asp Arg Ala Tyr
100 105 110

Val Phe Gly Tyr Glu Pro His Ser Val Phe Pro Ile Gly Val Met Ile
115 120 125

Leu Ser Leu Gly Leu Ile Pro Leu Pro Asn Ile Lys Phe Leu Ala Ser
130 135 140

Ser Ala Val Phe Tyr Thr Pro Phe Leu Arg His Ile Trp Ser Trp Cys
145 150 155 160

Gly Leu Thr Pro Ala Thr Arg Lys Asn Phe Val Ser Leu Leu Ser Ser
165 170 175

Gly Tyr Ser Cys Ile Leu Val Pro Gly Gly Val Gln Glu Thr Phe Tyr
180 185 190

Met Lys Gln Asp Ser Glu Ile Ala Phe Leu Lys Ala Arg Arg Gly Phe
195 200 205

Ile Arg Ile Ala Met Gln Thr Gly Thr Pro Leu Val Pro Val Phe Cys
210 215 220

Phe Gly Gln Met His Thr Phe Lys Trp Trp Lys Pro Asp Gly Glu Leu
225 230 235 240

Phe Met Lys Ile Ala Arg Ala Ile Lys Phe Thr Pro Thr Ile Phe Trp
245 250 255

Gly Val Leu Gly Thr Pro Leu Pro Phe Lys Asn Pro Met His Val Val
260 265 270

Val Gly Arg Pro Ile Glu Val Lys Gln Asn Pro Gln Pro Thr Ala Glu
275 280 285

Glu Val Ala Glu Val Gln Arg Glu Phe Ile Ala Ser Leu Lys Asn Leu
290 295 300

Phe Glu Arg His Lys Ala Arg Val Gly Tyr Ser Asp Leu Lys Leu Glu
305 310 315 320

Ile Phe

<210> 215
<211> 355
<212> PRT

<213> Mortierella ramanniana

<400> 215

Met Ala Ser Lys Asp Gln His Leu Gln Gln Lys Val Lys His Thr Leu
1 5 10 15

Glu Ala Ile Pro Ser Pro Arg Tyr Ala Pro Leu Arg Val Pro Leu Arg
20 25 30

Arg Arg Leu Gln Thr Leu Ala Val Leu Leu Trp Cys Ser Met Met Ser
35 40 45

Ile Cys Met Phe Ile Phe Phe Phe Leu Cys Ser Ile Pro Val Leu Leu
50 55 60

Trp Phe Pro Ile Ile Leu Tyr Leu Thr Trp Ile Leu Val Trp Asp Lys
65 70 75 80

Ala Pro Glu Asn Gly Gly Arg Pro Ile Arg Trp Leu Arg Asn Ala Ala
85 90 95

Trp Trp Lys Leu Phe Ala Gly Tyr Phe Pro Ala His Val Ile Lys Glu
100 105 110

Ala Asp Leu Asp Pro Ser Lys Asn Tyr Ile Phe Gly Tyr His Pro His
115 120 125

Gly Ile Ile Ser Met Gly Ser Phe Cys Thr Phe Ser Thr Asn Ala Thr
130 135 140

Gly Phe Asp Asp Leu Phe Pro Gly Ile Arg Pro Ser Leu Leu Thr Leu
145 150 155 160

Thr Ser Asn Phe Asn Ile Pro Leu Tyr Arg Asp Tyr Leu Met Ala Cys
165 170 175

Gly Leu Cys Ser Val Ser Lys Thr Ser Cys Gln Asn Ile Leu Thr Lys
180 185 190

Gly Gly Pro Gly Arg Ser Ile Ala Ile Val Val Gly Gly Ala Ser Glu
195 200 205

Ser Leu Asn Ala Arg Pro Gly Val Met Asp Leu Val Leu Lys Arg Arg
210 215 220

Phe Gly Phe Ile Lys Ile Ala Val Gln Thr Gly Ala Ser Leu Val Pro
225 230 235 240

Thr Ile Ser Phe Gly Glu Asn Glu Leu Tyr Glu Gln Ile Glu Ser Asn
245 250 255

pctau2012001598-seql

Glu Asn Ser Lys₂₆₀ Leu His Arg Trp Gln₂₆₅ Lys Lys Ile Gln His₂₇₀ Ala Leu

Gly Phe Thr₂₇₅ Met Pro Leu Phe His₂₈₀ Gly Arg Gly Val Phe₂₈₅ Asn Tyr Asp

Phe Gly₂₉₀ Leu Leu Pro His Arg₂₉₅ His Pro Ile Tyr Thr₃₀₀ Ile Val Gly Lys

Pro Ile Pro Val Pro Ser₃₁₀ Ile Lys Tyr Gly Gln₃₁₅ Thr Lys Asp Glu Ile₃₂₀

Ile Arg Glu Leu His₃₂₅ Asp Ser Tyr Met His₃₃₀ Ala Val Gln Asp Leu₃₃₅ Tyr

Asp Arg Tyr Lys₃₄₀ Asp Ile Tyr Ala Lys₃₄₅ Asp Arg Val Lys Glu₃₅₀ Leu Glu

Phe Val Glu₃₅₅

<210> 216
<211> 388
<212> PRT
<213> Homo sapiens
<400> 216

Met Lys Thr Leu₅ Ile Ala Ala Tyr Ser Gly₁₀ Val Leu Arg Gly Glu₁₅ Arg

Gln Ala Glu Ala₂₀ Asp Arg Ser Gln Arg₂₅ Ser His Gly Gly Pro₃₀ Ala Leu

Ser Arg Glu₃₅ Gly Ser Gly Arg Trp₄₀ Gly Thr Gly Ser Ser₄₅ Ile Leu Ser

Ala Leu₅₀ Gln Asp Leu Phe Ser₅₅ Val Thr Trp Leu Asn₆₀ Arg Ser Lys Val

Glu Lys Gln Leu Gln₇₀ Val Ile Ser Val Leu Gln₇₅ Trp Val Leu Ser Phe₈₀

Leu Val Leu Gly₈₅ Val Ala Cys Ser Ala Ile₉₀ Leu Met Tyr Ile Phe₉₅ Cys

Thr Asp Cys Trp₁₀₀ Leu Ile Ala Val Leu₁₀₅ Tyr Phe Thr Trp Leu₁₁₀ Val Phe

Asp Trp Asn₁₁₅ Thr Pro Lys Lys Gly₁₂₀ Gly Arg Arg Ser Gln₁₂₅ Trp Val Arg

Asn Trp Ala Val Trp Arg Tyr Phe Arg Asp Tyr Phe Pro Ile Gln Leu
Page 283

pctau2012001598-seql

130

135

140

Val Lys Thr His Asn Leu Leu Thr Thr Arg Asn Tyr Ile Phe Gly Tyr
 145 150 155 160

His Pro His Gly Ile Met Gly Leu Gly Ala Phe Cys Asn Phe Ser Thr
 165 170 175

Glu Ala Thr Glu Val Ser Lys Lys Phe Pro Gly Ile Arg Pro Tyr Leu
 180 185 190

Ala Thr Leu Ala Gly Asn Phe Arg Met Pro Val Leu Arg Glu Tyr Leu
 195 200 205

Met Ser Gly Gly Ile Cys Pro Val Ser Arg Asp Thr Ile Asp Tyr Leu
 210 215 220

Leu Ser Lys Asn Gly Ser Gly Asn Ala Ile Ile Ile Val Val Gly Gly
 225 230 235 240

Ala Ala Glu Ser Leu Ser Ser Met Pro Gly Lys Asn Ala Val Thr Leu
 245 250 255

Arg Asn Arg Lys Gly Phe Val Lys Leu Ala Leu Arg His Gly Ala Asp
 260 265 270

Leu Val Pro Ile Tyr Ser Phe Gly Glu Asn Glu Val Tyr Lys Gln Val
 275 280 285

Ile Phe Glu Glu Gly Ser Trp Gly Arg Trp Val Gln Lys Lys Phe Gln
 290 295 300

Lys Tyr Ile Gly Phe Ala Pro Cys Ile Phe His Gly Arg Gly Leu Phe
 305 310 315 320

Ser Ser Asp Thr Trp Gly Leu Val Pro Tyr Ser Lys Pro Ile Thr Thr
 325 330 335

Val Val Gly Glu Pro Ile Thr Ile Pro Lys Leu Glu His Pro Thr Gln
 340 345 350

Gln Asp Ile Asp Leu Tyr His Thr Met Tyr Met Glu Ala Leu Val Lys
 355 360 365

Leu Phe Asp Lys His Lys Thr Lys Phe Gly Leu Pro Glu Thr Glu Val
 370 375 380

Leu Glu Val Asn
 385

<210> 217

pctau2012001598-seql

<211> 328
 <212> PRT
 <213> Homo sapiens

<400> 217

Met Ala His Ser Lys Gln Pro Ser His Phe Gln Ser Leu Met Leu Leu
 1 5 10 15

Gln Trp Pro Leu Ser Tyr Leu Ala Ile Phe Trp Ile Leu Gln Pro Leu
 20 25 30

Phe Val Tyr Leu Leu Phe Thr Ser Leu Trp Pro Leu Pro Val Leu Tyr
 35 40 45

Phe Ala Trp Leu Phe Leu Asp Trp Lys Thr Pro Glu Arg Gly Gly Arg
 50 55 60

Arg Ser Ala Trp Val Arg Asn Trp Cys Val Trp Thr His Ile Arg Asp
 65 70 75 80

Tyr Phe Pro Ile Thr Ile Leu Lys Thr Lys Asp Leu Ser Pro Glu His
 85 90 95

Asn Tyr Leu Met Gly Val His Pro His Gly Leu Leu Thr Phe Gly Ala
 100 105 110

Phe Cys Asn Phe Cys Thr Glu Ala Thr Gly Phe Ser Lys Thr Phe Pro
 115 120 125

Gly Ile Thr Pro His Leu Ala Thr Leu Ser Trp Phe Phe Lys Ile Pro
 130 135 140

Phe Val Arg Glu Tyr Leu Met Ala Lys Gly Val Cys Ser Val Ser Gln
 145 150 155 160

Pro Ala Ile Asn Tyr Leu Leu Ser His Gly Thr Gly Asn Leu Val Gly
 165 170 175

Ile Val Val Gly Gly Val Gly Glu Ala Leu Gln Ser Val Pro Asn Thr
 180 185 190

Thr Thr Leu Ile Leu Gln Lys Arg Lys Gly Phe Val Arg Thr Ala Leu
 195 200 205

Gln His Gly Ala His Leu Val Pro Thr Phe Thr Phe Gly Glu Thr Glu
 210 215 220

Val Tyr Asp Gln Val Leu Phe His Lys Asp Ser Arg Met Tyr Lys Phe
 225 230 235 240

Gln Ser Cys Phe Arg Arg Ile Phe Gly Phe Tyr Cys Cys Val Phe Tyr
 245 250 255

pctau2012001598-seql

Gly Gln Ser Phe Cys Gln Gly Ser Thr Gly Leu Leu Pro Tyr Ser Arg
260 265 270

Pro Ile Val Thr Val Val Gly Glu Pro Leu Pro Leu Pro Gln Ile Glu
275 280 285

Lys Pro Ser Gln Glu Met Val Asp Lys Tyr His Ala Leu Tyr Met Asp
290 295 300

Ala Leu His Lys Leu Phe Asp Gln His Lys Thr His Tyr Gly Cys Ser
305 310 315 320

Glu Thr Gln Lys Leu Phe Phe Leu
325

<210> 218
<211> 361
<212> PRT
<213> Bos taurus

<400> 218

Met Lys Thr Leu Ile Ala Ala Tyr Ser Gly Val Leu Arg Gly Thr Gly
1 5 10 15

Ser Ser Ile Leu Ser Ala Leu Gln Asp Leu Phe Ser Val Thr Trp Leu
20 25 30

Asn Arg Ala Lys Val Glu Lys Gln Leu Gln Val Ile Ser Val Leu Gln
35 40 45

Trp Val Leu Ser Phe Leu Val Leu Gly Val Ala Cys Ser Val Ile Leu
50 55 60

Met Tyr Thr Phe Cys Thr Asp Cys Trp Leu Ile Ala Val Leu Tyr Phe
65 70 75 80

Thr Trp Leu Val Phe Asp Trp Asn Thr Pro Lys Lys Gly Gly Arg Arg
85 90 95

Ser Gln Trp Val Arg Asn Trp Ala Val Trp Arg Tyr Phe Arg Asp Tyr
100 105 110

Phe Pro Ile Gln Leu Val Lys Thr His Asn Leu Leu Thr Ser Arg Asn
115 120 125

Tyr Ile Phe Gly Tyr His Pro His Gly Ile Met Gly Leu Gly Ala Phe
130 135 140

Cys Asn Phe Ser Thr Glu Ala Thr Glu Val Ser Lys Lys Phe Pro Gly
145 150 155 160

pctau2012001598-seql

Ile Arg Pro Tyr Leu Ala Thr Leu Ala Gly Asn Phe Arg Met Pro Val
165 170 175

Leu Arg Glu Tyr Leu Met Ser Gly Gly Ile Cys Pro Val Asn Arg Asp
180 185 190

Thr Ile Asp Tyr Leu Leu Ser Lys Asn Gly Ser Gly Asn Ala Ile Ile
195 200 205

Ile Val Val Gly Gly Ala Ala Glu Ser Leu Ser Ser Met Pro Gly Lys
210 215 220

Asn Ala Val Thr Leu Arg Asn Arg Lys Gly Phe Val Lys Leu Ala Leu
225 230 235 240

Arg His Gly Ala Asp Leu Val Pro Thr Tyr Ser Phe Gly Glu Asn Glu
245 250 255

Val Tyr Lys Gln Val Ile Phe Glu Glu Gly Ser Trp Gly Arg Trp Val
260 265 270

Gln Lys Lys Phe Gln Lys Tyr Ile Gly Phe Ala Pro Cys Ile Phe His
275 280 285

Gly Arg Gly Leu Phe Ser Ser Asp Thr Trp Gly Leu Val Pro Tyr Ser
290 295 300

Lys Pro Ile Thr Thr Val Val Gly Glu Pro Ile Thr Ile Pro Arg Leu
305 310 315 320

Glu Arg Pro Thr Gln Gln Asp Ile Asp Leu Tyr His Ala Met Tyr Val
325 330 335

Gln Ala Leu Val Lys Leu Phe Asp Gln His Lys Thr Lys Phe Gly Leu
340 345 350

Pro Glu Thr Glu Val Leu Glu Val Asn
355 360

<210> 219
<211> 388
<212> PRT
<213> Mus musculus

<400> 219

Met Lys Thr Leu Ile Ala Ala Tyr Ser Gly Val Leu Arg Gly Glu Arg
1 5 10 15

Arg Ala Glu Ala Ala Arg Ser Glu Asn Lys Asn Lys Gly Ser Ala Leu
20 25 30

pctau2012001598-seql

Ser Arg Glu Gly Ser Gly Arg Trp Gly Thr Gly Ser Ser Ile Leu Ser
35 40 45

Ala Leu Gln Asp Ile Phe Ser Val Thr Trp Leu Asn Arg Ser Lys Val
50 55 60

Glu Lys Gln Leu Gln Val Ile Ser Val Leu Gln Trp Val Leu Ser Phe
65 70 75 80

Leu Val Leu Gly Val Ala Cys Ser Val Ile Leu Met Tyr Thr Phe Cys
85 90 95

Thr Asp Cys Trp Leu Ile Ala Val Leu Tyr Phe Thr Trp Leu Ala Phe
100 105 110

Asp Trp Asn Thr Pro Lys Lys Gly Gly Arg Arg Ser Gln Trp Val Arg
115 120 125

Asn Trp Ala Val Trp Arg Tyr Phe Arg Asp Tyr Phe Pro Ile Gln Leu
130 135 140

Val Lys Thr His Asn Leu Leu Thr Thr Arg Asn Tyr Ile Phe Gly Tyr
145 150 155 160

His Pro His Gly Ile Met Gly Leu Gly Ala Phe Cys Asn Phe Ser Thr
165 170 175

Glu Ala Thr Glu Val Ser Lys Lys Phe Pro Gly Ile Arg Pro Tyr Leu
180 185 190

Ala Thr Leu Ala Gly Asn Phe Arg Met Pro Val Leu Arg Glu Tyr Leu
195 200 205

Met Ser Gly Gly Ile Cys Pro Val Asn Arg Asp Thr Ile Asp Tyr Leu
210 215 220

Leu Ser Lys Asn Gly Ser Gly Asn Ala Ile Ile Ile Val Val Gly Gly
225 230 235 240

Ala Ala Glu Ser Leu Ser Ser Met Pro Gly Lys Asn Ala Val Thr Leu
245 250 255

Lys Asn Arg Lys Gly Phe Val Lys Leu Ala Leu Arg His Gly Ala Asp
260 265 270

Leu Val Pro Thr Tyr Ser Phe Gly Glu Asn Glu Val Tyr Lys Gln Val
275 280 285

Ile Phe Glu Glu Gly Ser Trp Gly Arg Trp Val Gln Lys Lys Phe Gln
290 295 300

pctau2012001598-seql

Lys Tyr Ile Gly Phe Ala Pro Cys Ile Phe His Gly Arg Gly Leu Phe
305 310 315 320

Ser Ser Asp Thr Trp Gly Leu Val Pro Tyr Ser Lys Pro Ile Thr Thr
325 330 335

Val Val Gly Glu Pro Ile Thr Val Pro Lys Leu Glu His Pro Thr Gln
340 345 350

Lys Asp Ile Asp Leu Tyr His Ala Met Tyr Met Glu Ala Leu Val Lys
355 360 365

Leu Phe Asp Asn His Lys Thr Lys Phe Gly Leu Pro Glu Thr Glu Val
370 375 380

Leu Glu Val Asn
385

<210> 220
<211> 3
<212> PRT
<213> Artificial Sequence

<220>
<223> conserved sequence

<400> 220

Tyr Phe Pro
1

<210> 221
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> conserved sequence

<400> 221

His Pro His Gly
1

<210> 222
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> conserved sequence

<400> 222

Glu Pro His Ser
1

<210> 223
<211> 24

<212> PRT
 <213> Artificial Sequence

<220>
 <223> conserved sequence

<220>
 <221> X
 <222> (2)..(2)
 <223> any amino acid

<220>
 <221> X
 <222> (5)..(5)
 <223> any amino acid

<220>
 <221> X
 <222> (6)..(6)
 <223> Lysine (K) or Arginine (R)

<220>
 <221> X
 <222> (7)..(7)
 <223> any amino acid

<220>
 <221> X
 <222> (9)..(11)
 <223> any amino acid

<220>
 <221> X
 <222> (13)..(15)
 <223> any amino acid

<220>
 <221> X
 <222> (16)..(16)
 <223> Leucine (L) or Valine (V)

<220>
 <221> X
 <222> (19)..(21)
 <223> any amino acid

<220>
 <221> X
 <222> (24)..(24)
 <223> Glutamic Acid (E) or Glutamine (Q)

<400> 223

Arg Xaa Gly Phe Xaa Xaa Xaa Ala Xaa Xaa Xaa Gly Xaa Xaa Xaa Xaa
 1 5 10 15

Val Pro Xaa Xaa Xaa Phe Gly Xaa
 20

<210> 224
 <211> 8
 <212> PRT
 <213> Artificial Sequence

<220>

<223> conserved sequence

<220>

<221> X

<222> (3)..(3)

<223> any amino acid

<220>

<221> X

<222> (5)..(7)

<223> any amino acid

<400> 224

Phe Leu Xaa Leu Xaa Xaa Xaa Asn
1 5

<210> 225

<211> 118

<212> PRT

<213> Artificial Sequence

<220>

<223> conserved sequence

<400> 225

Ala Leu Val Val Ala Asn His Gln Ser Phe Leu Asp Pro Leu Val Leu
1 5 10 15

Ser Ala Leu Leu Pro Arg Lys Gly Gly Arg Val Arg Phe Val Ala Lys
20 25 30

Lys Glu Leu Phe Tyr Val Pro Leu Leu Gly Trp Leu Leu Arg Leu Leu
35 40 45

Gly Ala Ile Phe Ile Asp Arg Glu Asn Gly Arg Leu Ala Arg Ala Ala
50 55 60

Leu Arg Glu Ala Val Arg Leu Leu Arg Asp Gly Gly Trp Leu Leu Ile
65 70 75 80

Phe Pro Glu Gly Thr Arg Ser Arg Pro Gly Lys Leu Leu Pro Phe Lys
85 90 95

Lys Gly Ala Ala Arg Leu Ala Leu Glu Ala Gly Val Pro Ile Val Pro
100 105 110

Val Ala Ile Arg Gly Thr
115

<210> 226

<211> 187

<212> PRT

<213> Artificial Sequence

<220>

<223> conserved sequence

<220>
<221> X
<222> (15)..(15)
<223> any amino acid

<220>
<221> X
<222> (18)..(18)
<223> any amino acid

<220>
<221> X
<222> (23)..(23)
<223> any amino acid

<220>
<221> X
<222> (25)..(26)
<223> any amino acid

<220>
<221> X
<222> (28)..(30)
<223> any amino acid

<220>
<221> X
<222> (32)..(33)
<223> any amino acid

<220>
<221> X
<222> (35)..(38)
<223> any amino acid

<220>
<221> X
<222> (41)..(41)
<223> any amino acid

<220>
<221> X
<222> (46)..(48)
<223> any amino acid

<220>
<221> X
<222> (53)..(53)
<223> any amino acid

<220>
<221> X
<222> (55)..(57)
<223> any amino acid

<220>
<221> X
<222> (61)..(61)
<223> any amino acid

<220>
<221> X
<222> (67)..(67)
<223> any amino acid

<220>

pctau2012001598-seql

<221> X
<222> (72)..(72)
<223> any amino acid

<220>
<221> X
<222> (74)..(77)
<223> any amino acid

<220>
<221> X
<222> (79)..(79)
<223> any amino acid

<220>
<221> X
<222> (114)..(114)
<223> any amino acid

<220>
<221> X
<222> (127)..(128)
<223> any amino acid

<220>
<221> X
<222> (136)..(136)
<223> any amino acid

<220>
<221> X
<222> (139)..(142)
<223> any amino acid

<220>
<221> X
<222> (144)..(144)
<223> any amino acid

<220>
<221> X
<222> (150)..(150)
<223> any amino acid

<220>
<221> X
<222> (164)..(165)
<223> any amino acid

<220>
<221> X
<222> (167)..(172)
<223> any amino acid

<400> 226

Ala Val Phe Asp Lys Asp Gly Thr Leu Thr Glu Asp Asp Thr Xaa Phe
1 5 10 15

Leu Xaa Tyr Leu Leu Lys Xaa Leu Xaa Xaa Leu Xaa Xaa Xaa Leu Xaa
20 25 30

Xaa Asp Xaa Xaa Xaa Xaa Gly Ser Xaa Leu Thr Leu Ser Xaa Xaa Xaa
35 40 45

pctau2012001598-seql

Asp Leu Leu Glu Xaa Leu Xaa Xaa Xaa Gly Gly Ile Xaa Val Ile Gly
50 55 60

Leu Ala Xaa Arg Tyr Leu Glu Xaa Leu Xaa Xaa Xaa Xaa Glu Xaa Ala
65 70 75 80

Lys Leu Phe Glu Gly Phe Ile Lys Pro Asp Ala Ala Glu Leu Leu Lys
85 90 95

Glu Leu His Glu Ala Gly Leu Arg Val Val Val Leu Thr Gly Asp Pro
100 105 110

Arg Xaa Ile Ala Lys Pro Val Ala Lys Glu Leu Gly Ile Asp Xaa Xaa
115 120 125

Asn Val Leu Ala Thr Glu Leu Xaa Asp Glu Xaa Xaa Xaa Xaa Val Xaa
130 135 140

Gly Arg Ile Thr Gly Xaa Leu Asp Lys Ala Arg Ala Val Glu Arg Leu
145 150 155 160

Val Val Leu Xaa Xaa Lys Xaa Xaa Xaa Xaa Xaa Xaa Val Val Ala Ile
165 170 175

Gly Asp Ser Ala Asn Asp Leu Pro Ala Leu Lys
180 185

<210> 227
<211> 190
<212> PRT
<213> Artificial Sequence

<220>
<223> conserved sequence

<400> 227

Ile Lys Ala Val Val Phe Asp Lys Asp Gly Thr Leu Thr Asp Gly Glu
1 5 10 15

Pro Pro Ile Ala Glu Ala Ile Val Glu Ala Ala Ala Glu Leu Gly Leu
20 25 30

Pro Leu Leu Leu Pro Leu Glu Glu Val Glu Lys Leu Leu Gly Arg Gly
35 40 45

Val Glu Gly Ile Glu Arg Ile Leu Leu Glu Gly Gly Leu Thr Ala Glu
50 55 60

Leu Leu Leu Glu Leu Glu Gly Glu Leu Ala Ala Gly Lys Thr Ala Val
65 70 75 80

Leu Val Ala Leu Asp Gly Glu Val Leu Gly Leu Ile Ala Leu Ala Asp
85 90 95

pctau2012001598-seql

Lys Leu Tyr Pro Gly Ala Arg Glu Ala Leu Lys Ala Leu Lys Glu Arg
100 105 110

Gly Ile Lys Val Ala Ile Leu Thr Asn Gly Asp Arg Ala Asn Ala Glu
115 120 125

Ala Val Leu Glu Ala Leu Gly Leu Ala Asp Leu Phe Asp Val Ile Val
130 135 140

Asp Ser Asp Asp Val Gly Pro Val Lys Pro Lys Pro Glu Ile Phe Leu
145 150 155 160

Lys Ala Leu Glu Arg Leu Gly Val Lys Pro Glu Glu Val Leu Met Val
165 170 175

Gly Asp Gly Val Asn Asp Ala Pro Ala Leu Ala Ala Ala Gly
180 185 190

<210> 228
<211> 15
<212> PRT
<213> Artificial Sequence

<220>
<223> conserved sequence

<400> 228

Gly Asp Leu Val Ile Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro
1 5 10 15

<210> 229
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> conserved sequence

<220>
<221> X
<222> (2)..(2)
<223> any amino acid

<220>
<221> X
<222> (4)..(4)
<223> any amino acid

<220>
<221> X
<222> (5)..(5)
<223> Threonine (T) or Valine (V)

<220>
<221> X
<222> (6)..(6)
<223> Leucine (L) or Valine (V)

pctau2012001598-seql

<400> 229

Asp Xaa Asp Xaa Xaa Xaa
1 5

<210> 230
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> conserved sequence

<220>
<221> X
<222> (2)..(2)
<223> 17 to 20 amino acids; where the amino acids can be any amino acids

<220>
<221> X
<222> (3)..(3)
<223> Glycine (G) or Serine (S)

<220>
<221> X
<222> (4)..(4)
<223> Aspartic Acid (D) or Serine (S)

<220>
<221> X
<222> (5)..(7)
<223> any amino acid

<220>
<221> X
<222> (8)..(8)
<223> Aspartic Acid (D) or Asparagine (N)

<400> 230

Lys Xaa Xaa Xaa Xaa Xaa Xaa Xaa
1 5

<210> 231
<211> 1580
<212> DNA
<213> Arabidopsis thaliana

<400> 231
cagggttat ttaacttgcc ctttctcggt tcctcctttt tttcttaaacc cactctgctt 60
cctcttcctc tgagaaatca aatcactcac actccaaaaa aaaatctaaa ctttctcaga 120
gtttaatgaa gaagcgctta accacttcca cttgttcttc ttctccatct tcctctgttt 180
cttcttctac tactacttcc tctcctattc agtcggaggc tccaaggcct aaacgagcca 240
aaagggctaa gaaatcttct ctttctggtg ataaatctca taacccgaca agccctgctt 300
ctacccgacg cagctctatc tacagaggag tcactagaca tagatggact gggagattcg 360
aggctcatct ttgggacaaa agctcttgga attcgattca gaacaagaaa ggcaaacaag 420

pctau2012001598-seq1

gtttcgagca tatgacagtg aagaagcagc agcacatacg tacgatctgg ctgctctcaa	480
gtactgggga cccgacacca tcttgaattt tccggcagag acgtacacaa aggaattgga	540
agaaatgcag agagtgcacaa aggaagaata tttggcttct ctccgccgcc agagcagtg	600
tttctccaga ggcgtctcta aatatcgcg cgtcgctagg catcaccaca acggaagatg	660
ggaggctcgg atcggaagag tgtttgggaa caagtacttg tacctcggca cctataatac	720
gcaggaggaa gctgctgcag catatgacat ggctgcgatt gagtatcgag gcgcaaacgc	780
ggttactaat ttcgacatta gtaattacat tgaccgggta aagaagaaag gtgttttccc	840
gttccctgtg aaccaagcta accatcaaga gggatttctt gttgaagcca aacaagaagt	900
tgaacgaga gaagcgaagg aagagcctag agaagaagt aaacaacagt acgtggaaga	960
accaccgcaa gaagaagaag agaaggaaga agagaaagca gagcaacaag aagcagagat	1020
tgtaggatat tcagaagaag cagcagtggt caattgctgc atagactctt caaccataat	1080
ggaaatggat cgttgtgggg acaacaatga gctggcttgg aacttctgta tgatggatac	1140
agggttttct ccgtttttga ctgatcagaa tctcgcaat gagaatccca tagagtatcc	1200
ggagctattc aatgagttag catttgagga caacatcgac ttcattgttcg atgatgggaa	1260
gcacgagtg tgaacttgg aaaatctgga ttgttgctg gtgggaagag agagcccacc	1320
ctcttcttct tcaccattgt cttgcttct tactgactct gttcatcaa caacaacaac	1380
aacaacctcg gtttcttgta actatttgg ctgagagaga gagctttgcc ttctagtttg	1440
aatttctatt tcttccgctt cttcttctt ttttcttct gttgggttct gcttaggggt	1500
tgtatttcag tttcagggtc tgttcgttgg ttctgaataa tcaatgtctt tgccccctt	1560
ctaattgctcc aagttcagat	1580

<210> 232
 <211> 1584
 <212> DNA
 <213> Arabidopsis thaliana

<400> 232	
cagggtttat ttaacttgcc ctttctcggt tcctcctttt tttcttaaac cactctgctt	60
cctcttcctc tgagaaatca aatcactcac actccaaaaa aaaatctaaa ctttctcaga	120
gtttaatgaa gaagcgctta accacttcca cttgttcttc ttctccatct tcctctgttt	180
cttcttctac tactacttcc tctcctattc agtcggaggc tccaaggcct aaacgagcca	240
aaagggctaa gaaatcttct ctttctggtg ataaatctca taaccgaca agccctgctt	300
ctaccgacg cagctctatc tacagaggag ttagtagaca tagatggact gggagattcg	360
aggctcatct ttgggacaaa agctcttggg attcgattca gaacaagaaa ggcaaacaag	420
tttatctggg agcatatgac agtgaagaag cagcagcaca tacgtacgat ctggctgctc	480
tcaagtactg gggacccgac accatcttga attttccggc agagacgtac acaaaggaat	540
tggaagaaat gcagagagtg acaaaggaag aatatttggc ttctctccgc cgccagagca	600
gtggtttctc cagaggcgct tctaaatatc gcggcgctgc taggcattcac cacaacggaa	660

pctau2012001598-seq1

gatgggaggc	tcggatcgga	agagtgtttg	ggaacaagta	cttgtagctc	ggcacctata	720
atacgagga	ggaagctgct	gcagcatatg	acatggctgc	gattgagtat	cgaggcgcaa	780
acgcggttac	taatttcgac	attagtaatt	acattgaccg	gttaaagaag	aaaggtgttt	840
tcccgttccc	tgtgaaccaa	gctaaccatc	aagagggtat	tcttggtgaa	gccaaacaag	900
aagttgaaac	gagagaagcg	aaggaagagc	ctagagaaga	agtgaacaaa	cagtacgtgg	960
aagaaccacc	gcaagaagaa	gaagagaagg	aagaagagaa	agcagagcaa	caagaagcag	1020
agattgtagg	atattcagaa	gaagcagcag	tggtcaattg	ctgcatagac	tcttcaacca	1080
taatggaaat	ggatcgttgt	ggggacaaca	atgagctggc	ttggaacttc	tgtatgatgg	1140
atacaggggt	ttctccgttt	ttgactgatc	agaatctcgc	gaatgagaat	cccatagagt	1200
atccggagct	attcaatgag	ttagcatttg	aggacaacat	cgacttcatg	ttcgatgatg	1260
ggaagcacga	gtgcttgaac	ttggaaaatc	tggttggttg	cgtgggtggg	agagagagcc	1320
caccctcttc	ttcttcacca	ttgtcttgct	tatctactga	ctctgcttca	tcaacaacaa	1380
caacaacaac	ctcggtttct	tgtaactatt	tggtctgaga	gagagagctt	tgcttcttag	1440
tttgaatttc	tatttcttcc	gcttcttctt	cttttttttc	ttttggtggg	ttctgcttag	1500
ggtttgtatt	tcagtttcag	ggcttggtcg	ttggttctga	ataatcaatg	tctttgcccc	1560
ttttctaattg	ctccaagtgc	agat				1584

<210> 233
 <211> 1470
 <212> DNA
 <213> Arabidopsis thaliana

<400> 233	
cagggtttat	ttacttgcc ctttctcggt tcctcctttt tttcttaaac cactctgctt 60
cctcttcctc	tgagaaatca aatcactcac actccaaaaa aaaatctaaa ctttctcaga 120
gtttaatgaa	gaagcgctta accacttcca cttgttcttc ttctccatct tcctctgttt 180
cttcttctac	tactacttcc tctcctattc agtcggaggc tccaaggcct aaacgagcca 240
aaagggttaa	gaaatcttct ctttctggtg ataaatctca taaccgaca agccctgctt 300
ctaccgagc	cagctctatc tacagaggag ttagtagaca tagatggact gggagattcg 360
aggctcatct	ttgggacaaa agctcttgga attcgattca gaacaagaaa ggcaaacaag 420
tttatctggg	agcatatgac agtgaagaag cagcagcaca tacgtacgat ctggctgctc 480
tcaagtactg	gggacccgac accatcttga attttccggc agagacgtac acaaaggaat 540
tggaagaaat	gcagagagtg acaaaggaag aatatttggc ttctctccgc cgccagagca 600
gtggtttctc	cagaggcgtc tctaaatatc gcggcgtcgc taggcattcac cacaacggaa 660
gatgggaggc	tcggatcgga agagtgtttg ggaacaagta cttgtacctc ggcacctata 720
atacgagga	ggaagctgct gcagcatatg acatggctgc gattgagtat cgaggcgcaa 780
acgcggttac	taatttcgac attagtaatt acattgaccg gttaaagaag aaaggtgttt 840

pctau2012001598-seq1

tcccgttccc	tgtgaaccaa	gctaaccatc	aagagggtat	tcttgttgaa	gccaaacaag	900
aagttgaaac	gagagaagcg	aaggaagagc	ctagagaaga	agtgaacaa	cagtacgtgg	960
aagaaccacc	gcaagaagaa	gaagagaagg	aagaagagaa	agcagagcaa	caagaagcag	1020
agattgtagg	atattcagaa	gaagcagcag	tgggtcaattg	ctgcatagac	tcttcaacca	1080
taatggaaat	ggatcgttgt	ggggacaaca	atgagctggc	ttggaacttc	tgtatgatgg	1140
atacaggggt	ttctccgttt	ttgactgatc	agaatctcgc	gaatgagaat	cccatagagt	1200
atccggagct	attcaatgag	ttagcatttg	aggacaacat	cgacttcatg	ttcgatgatg	1260
ggaagcacga	gtgcttgaac	ttggaaaatc	tggattgttg	cgtgggtggga	agagagagcc	1320
cacctcttc	ttcttcacca	ttgtcttgct	tatctactga	ctctgcttca	tcaacaacaa	1380
caacaacaac	ctcggtttct	tgtactatt	tgtttcaggg	cttgttcggt	ggttctgaat	1440
aatcaatgtc	tttgcccctt	ttctaattgct				1470

<210> 234
 <211> 1397
 <212> DNA
 <213> Arabidopsis lyrata

<400> 234

cagggtttat	ttaacttgcc	ctttctcggt	tcctcctttt	tttcttaaac	cactcttctt	60
tcttctctg	atttcccaga	ttttaatgaa	gaggcgctta	accacttcca	cttcttcttc	120
ttctccatct	tcctctgttt	cttcttctac	tactacttcc	tctcctattc	aatcggaggc	180
tcccaggcct	aaacgagcca	aaagggctaa	gaaatcgtct	ccttctggtg	ataaatcaca	240
taacccgaca	agccctgctt	ctaccagacg	cagctctatc	tacagaggag	tcactagaca	300
tagatggact	gggagattcg	aggctcatct	ttgggacaaa	agctcttgga	attcgattca	360
gaacaagaaa	ggcaaacaag	gagcatatga	cagcgaagaa	gcagcagcac	atacgtacga	420
tctggctgct	ctcaagtact	ggggacccga	caccatcttg	aattttccgg	cagagacgta	480
cacaaaggaa	ttggaagaga	tgcagagagt	gacaaaggaa	gaatatttgg	cttctctccg	540
ccgccagagc	agtggtttct	ccagaggcgt	ctctaaatat	cgcggcgtcg	ctaggcatca	600
ccataacgga	agatgggagg	ctcggattgg	aagagtgttt	gggaacaagt	acttgtaact	660
cggcacctat	aatacgcagg	aggaagctgc	tgcggcgtat	gatatggcgg	cgatagagta	720
tcgaggtgcg	aacgcggtta	ctaattttga	cattagtaat	tacatcgacc	ggttaaagaa	780
gaaaggtgtt	ttcccgttcc	ctgtgaacca	acctaatacat	caagaggcta	ttcttgttga	840
agccaaacaa	gaaatcgaaa	cgagagaagc	taaagaagag	cctagagaag	aagtgaacaa	900
acagtacgtg	gaagaaccac	cacaagaaga	gaaagaagaa	gagaaagcag	agcaacaaga	960
agcagagttt	gtaggatata	aagatgaagg	ggcggtggtc	aattgctgca	tagactcttc	1020
agccataatg	gaaatgaatc	gttggtggga	caacaatgag	ctggcttgga	acttctgtat	1080
gatggattca	gggtttgctc	cgtttttgac	ggatcagaat	ctctcaaatg	agaatcccat	1140
cgagtatccg	gagcttttca	atgagttagc	atttgaggat	aacattgact	tcatgttcga	1200

pctau2012001598-seq1

cgaagccaag aacgactgcc tgagcttggg aaatctggat tgttgcgtgg tgggaagaga	1260
gagcccaact tcttcgtcct caccattgtc ttgcttttct actgactctg cttcatcaac	1320
aacaacaaca acctcgggtt cttgtaacta tttgggcttg ttcgttggtt ctgaataatc	1380
aatgtctttg ccccttt	1397

<210> 235
 <211> 1242
 <212> DNA
 <213> Brassica napus

<400> 235	
atgaagagac ccttaaccac ttctccttct tcctcctctt ctacttcttc ttcggcctgt	60
atacttccga ctcaatcaga gactccaagg cccaaacgag ccaaagggc taagaaatct	120
tctctgcgtt ctgatgttaa accacagaat cccaccagtc ctgcctccac cagacgcagc	180
tctatctaca gaggagtcac tagacataga tggacagga gatacgaagc tcatctatgg	240
gacaaaagct cgtggaattc gattcagaac aagaaaggca aacaagtta tctgggagca	300
tatgacagcg aggaagcagc agcacatacg tacgatctag ctgctctcaa gtactgggggt	360
ccaacacca tcttgaactt tccggttgag acgtacacaa aggagctgga ggagatgcag	420
agatgtacaa aggaagagta tttggcttct ctccgccgcc agagcagtgg tttctctaga	480
ggcgtctcta aatatcgcg cgtcgccagg catcaccata atggaagatg ggaagctcgg	540
attggaaggg tgtttgaaa caagtacttg tacctcggca cctataatac gcaggaggaa	600
gctgcagctg catatgacat ggcggctata gactacagag gtgcaaacgc agtgaccaac	660
ttcgacattg gtaactacat cgaccgggta aagaaaaaag gtgtcttccc gttccccgtg	720
agccaagcta atcatcaaga agctgttctt gctgaaacca aacaagaagt ggaagctaaa	780
gaagagccta cagaagaagt gaagcagtgt gtcgaaaaag aagaagctaa agaagagaag	840
actgagaaaa aacaacaaca agaagtggag gaggcgggtga tcaattgctg cattgattct	900
tcagagagca atgagctggc ttgggacttc tgtatgatgg attcaggggt tgctccgttt	960
ttgactgatt caaatctctc gactgagaat cccattgagt atcctgagct tttcaatgag	1020
atggggtttt aggataacat tgacttcatg ttcgaggaag ggaagcaaga ctgcttgagc	1080
ttggagaatc ttgattgttg cgatggtgtt gttgtggtgg gaagagagag cccaacttca	1140
ttgtcgtctt ctccgttgtc ctgcttgtct actgactctg cttcatcaac aacaacaaca	1200
gcaacaacag taacctctgt ttcttggaaac tattctgtct ga	1242

<210> 236
 <211> 1248
 <212> DNA
 <213> Brassica napus

<400> 236	
atgaagagac ccttaaccac ttctccttct acctcctctt ctacttcttc ttcggcttgt	60
atacttccga ctcaaccaga gactccaagg cccaaacgag ccaaagggc taagaaatct	120

pctau2012001598-seq1

tctattccta	ctgatgttaa	accacagaat	cccaccagtc	ctgcctccac	cagacgcagc	180
tctatctaca	gaggagtcac	tagacataga	tggacagggg	gatacgaggc	tcatctatgg	240
gacaaaagct	cgtggaattc	gattcagaac	aagaaaggca	aacaagttta	tctgggagca	300
tatgacagcg	aggaagcagc	agcgcatacg	tacgatctag	ctgctctcaa	gtactggggg	360
cccgacacca	tcttgaactt	tccggctgag	acgtacacaa	aggagttgga	ggagatgcag	420
agatgtacaa	aggaagagta	tttggcttct	ctccgccgcc	agagcagtg	tttctctaga	480
ggcgtctcta	aatatcgcg	cgtcgccagg	catcaccata	acggaagatg	ggaagctagg	540
attggaaggg	tgtttgga	caagtacttg	tacctcgga	cttataatac	gcaggaggaa	600
gctgcagctg	catatgacat	ggcggctata	gagtacagag	gcgcaaacgc	agtgaccaac	660
ttcgacatta	gtaactacat	cgaccgggta	aagaaaaaag	gtgtcttccc	attccctgtg	720
agccaagcca	atcatcaaga	agctgttctt	gctgaagcca	aacaagaagt	ggaagctaaa	780
gaagagccta	cagaagaagt	gaagcagtg	gtcgaaaaag	aagaaccgca	agaagctaaa	840
gaagagaaga	ctgagaaaaa	acaacaacaa	caagaagtgg	aggaggcggt	ggtcacttgc	900
tgcattgatt	cttcggagag	caatgagctg	gcttgggact	tctgtatgat	ggattcaggg	960
tttgctccgt	ttttgacgga	ttcaaattct	tcgagtgaga	atcccattga	gtatcctgag	1020
cttttcaatg	agatgggggt	tgaggataac	attgacttca	tgttcgagga	agggaagcaa	1080
gactgcttga	gcttgagaaa	tctggattgt	tgcgatgggt	ttgttgtggt	gggaagagag	1140
agcccaactt	cattgtcgtc	ttcaccgttg	tcttgcttgt	ctactgactc	tgcttcatca	1200
acaacaacaa	caacaataac	ctctgtttct	tgtaactatt	ctgtctga		1248

<210> 237
 <211> 868
 <212> DNA
 <213> Glycine max

<400> 237						
gaggctccaa	atgaagaggt	ctccagcatc	ttcttgttca	tcatccactt	cctctgttgg	60
gtttgaagtt	catcatccca	ttgaaaaaag	aaggcctaag	catccaagga	ggaataattt	120
gaagtcacaa	aatgcaagc	agaaccaaac	caccactggg	ggcagaagaa	gctctatcta	180
tagaggagtt	acaaggcata	ggtggacagg	gaggtttgaa	gctcacctat	gggataagag	240
ctcttggaa	aacattcaga	gcaagaagg	taaacaagtt	tatttggggg	catatgatac	300
tgaagaatct	gcagcacgta	cctatgacct	tgcagccctt	aagtactggg	gaaaagatgc	360
caccctgaat	ttcccgatag	aaacttatac	caaggacctc	gaggaaatgg	acaaggtttc	420
aagagaagaa	tatttagctt	ctttgcggcg	caaagcagt	ggcttttcta	gaggcatctc	480
taagtaccgt	ggggttgcta	ggcatcatca	taatggctgc	tgggaagctc	gcattggaag	540
agtatgtgga	aacaagtacc	tctacttggg	aacatataaa	actcaagagg	aggcagcagt	600
ggcatatgac	atggcagcaa	ttgagtaccg	tggagtcaat	gcagtgacca	attttgacat	660

pctau2012001598-seq1

aagcaactac atggacaaaa taaagaagaa aaatgaccaa accctacaac aacaacaaac	720
agaagtacaa acagaaacag ttcctaactc ctctgactct gaagaagcag aagtagaaca	780
acaacacaca acaacaataa ctacaccacc cccatctgaa aatctgcaca tgctaccaca	840
ggaacaccaa gtaggtggat ggggtgtga	868

<210> 238
 <211> 1254
 <212> DNA
 <213> *Jatropha curcas*

<400> 238	
atgaagaggt cttctgcttc atcttgctct tcttcttctt cttcttcttc ttctccatcc	60
tcttcttcgt cttctgcttg ttctgcttcg tcttcttgct tagattcagt atctcctcct	120
aatcaccatc aattacgatc agagaaatca aaatccaaac gcattcgaaa aattcaaacc	180
aagcaagata aatgtcagac tacagctact accaccagtc caagcggcgg cggtaggaga	240
agctccattt acagaggagt caccggcat agatggactg gaaggtttga agctcatctt	300
tgggataaga gttcttgaa taacattcaa aacaagaaag gaaggcaagt ttatttgggg	360
gcttacgaca atgaggaagc agctgctcat acctatgac ttgctgctct caagtactgg	420
ggacaagaca ccactttgaa ttttccgata gaaacatact caaaggagct tgaagagatg	480
caaaagatga gcaaggaaga gtacttagca tctcttcgac ggagaagcag tggattttca	540
agaggagttt ctaagtaccg gggagtagct aggcattcat acaatggccg gtgggaagct	600
cgaattggcc gggtttttgg caataagtat ctctacctcg gaacttaca tacacaagaa	660
gaggcagcag cagcatatga tatggcagca atagagtaca gaggagcaaa tgcagtaacc	720
aattttgatg tcagccatta catagaccgt ttgaagaaga aaggcattcc tttagataaa	780
atcctaccag aaacactttc taaaggctca aaagagtcag aagaaatcga gcgaacctca	840
cccttaccgt tgccatcacc accatcacca tcaataacac cattacacga agaaatagtc	900
tcaccacagc tgcttgaaac tgaatgccca caacatcctc catgtatgga tacttgctact	960
atgatcgta tggaccctat agaagagcac gagcttactt ggagcttctg tctcgattcg	1020
gggttagttc cgctccctgt gcctgacctt ccactagcaa atggctgtga gttaccagac	1080
ttgttgatg acacaggctt tgaagacaat attgacttga tatttgatgc ttgttgcttc	1140
ggaaatgatg ccaaccctgc agatgagaat gggaaagaga ggttgtcttc cgcttcaact	1200
tctccatctt gttccacaac attaaacttct gtttcttgta actattctgt ttga	1254

<210> 239
 <211> 1332
 <212> DNA
 <213> *Ricinus communis*

<400> 239	
atgaagaggt ctctacttc cccgtgctct tcttcttcat catcttcata ttcatcttct	60
tcagcttcat cttcttggtg cggctctgat gatactcctg tcgctcctgg tagtcatcat	120

pctau2012001598-seq1

catcatgatc accaccaatt acgatcgagc aaatcatcca aacgcattag aaaagttaaa	180
aaaaagcagc agaaccataa tattgatcag aataatacta acaccactat tactgctcct	240
acctctgccca gaagaagctc catttacaga ggagtcaccc ggcatagatg gacgggaagg	300
tttgaggcac acctttggga taagagttca tggaataaca ttcaaaacaa gaagggaaga	360
caaggggcct atgataatga agaagcagct gctcatactt atgaccttgc tgccttaag	420
tactggggac cagagacaac cttgaatttc cccatagaaa catatcctaa agagcttgaa	480
gagatgcaaa agatgagcaa ggaagagtac ttggcatctc ttcgacgcca aagcagtggg	540
ttttccagag gagtttctaa gtaccgtggg gttgctaggc atcaccacaa tggtcgatgg	600
gaggcacgaa tcggacgagt ttttggcaac aagtatctct accttggcac ttataacaca	660
caagaggagg cagcagcagc atatgatatg gcagcaattg agtacagagg agcaaagtct	720
gtgaccaatt ttgatatcag caattacata gaccgcttga agaagaaagg catcctttta	780
gaccagatac tccctgatca accacttaga aaatgctcta gcgaatctga agaagcagaa	840
gcagaagcag aagttgaacg actgccgtca ctgccctcat cgatattacc acaagaacag	900
gacacaatct caccacaact tcaatgcaca cagctccttc catccatgga ctcttgcaca	960
atgattaata tggaccctat cgaagataat gagctcacct ggagtttttg cttggattct	1020
gggctgacct ttttctcagt acccgagctc cccctcgaaa atgcttgcca gctaccggac	1080
ttgttcgatg acacaggctt tgaagacaat atagacttga tttttgatgg gtgttgcttt	1140
ggaaatgatg atgacggtgg tgggtggtgcc aaccaccagg aatttatggt ggaaagtaga	1200
ggatgtaggg taggggaggt tgggatttca gggagcatgg aggaggaaaa cgggaaggag	1260
atgtgttgca gttcaagttc gccatcttgt tcaacaacca cctcggtttc ttgttgtaac	1320
tattctgttt ga	1332

<210> 240
 <211> 1209
 <212> DNA
 <213> Populus trichocarpa

<400> 240	
atgaagaggt cttcatcttg ctctcctcc tcctcctct catcttcttc ttgtgttgcc	60
tctgaaagca ttcacaaacc aaaggccaaa cgcattagaa agaatacaaa gagcaatcaa	120
ggcaaatccc agaatgctgc tgctgctgct gctaataaca gtcataactc aggcaaaaga	180
agttccattt atagaggagt caccagacat agatggacag gaaggtttga agctcatctt	240
tgggataaga gttcatggaa cagcattcaa aacaagaagg gaaaacaagg tgcctatgac	300
aatgaggagg cagccgcaca cacctatgat cttgctgccc tgaagtactg ggggtcagag	360
acaaccttga attttccgat agaaacatac actaaagaga tcgaagagat gcaaaagggtg	420
actaaagaag agtacttagc atctcttaga cggcaaagca gtggattctc tagaggagtc	480
tccaaatacc gtgggggtgg taggcatcat cataatggcc gatgggaagc tagaattgga	540
cgagtttatg ggaataagta tctctacctt ggaacttaca atacacaaga agaggcagcg	600

pctau2012001598-seql

gcagcgtatg	atatggcggc	gatacagtat	agaggagcaa	atgcagtgac	caatTTTgat	660
gtagcaatt	acatagaacg	gttagaggaag	aaaggcatcc	ctatagaccg	aatcctccaa	720
gaacaacaac	ttcttaacaa	ctcggtagat	tccagcgtgg	aagtggaagt	tgaacaacca	780
acaccaccac	cacaacaaca	acaagaggaa	caagaacaaa	agatagtctc	gtcgtcttcg	840
caacttcaat	gttcacagct	aaattcaagc	ttggatggca	cgcctcctat	ggtaattatg	900
gacactattg	aagagcacga	gctagcatgg	agctTTTgta	tggattcagg	attgagcctc	960
acaatgcctg	atcttcctct	cgaaaattcc	tgcgagttac	cggacttggt	cgatcataca	1020
gggtttgagg	acaacattga	cttgatattc	gatgcatggt	gctatggaaa	ggaggccaat	1080
ccggctggct	acacattaga	agataatagc	acaggaggag	ttgaagagga	caggttgctg	1140
tctgattctg	tctcaaattc	tccaactagt	tcaacaacca	cctctgtttc	ttgtaactat	1200
tctgtttga						1209

<210> 241

<211> 1546

<212> DNA

<213> *Brachypodium distachyon*

<400> 241

atgaagagat	cccctcctca	gccgtccccg	tctccttcct	cttcgccggc	ctcctcgtcg	60
tcgtccccgt	cctcctccga	ctcgtcttcc	tcgatcgcca	ttccccgcaa	gagagcgagg	120
acggcggcgg	cggcggcggg	cggcggcaag	gccagggcgg	cggccgccaa	gaggcccaag	180
aaagatggca	aggactccgg	ctcctcctcc	aatggcggcg	gcggcggcgg	cgggaagagg	240
agctccatct	acaggggagt	cacgaggcac	aggtggacag	gtagatttga	ggcgcattct	300
tgggacaaga	actgcttcac	ttcgcttcag	aacaagaaga	aaggcaggca	agtctatctg	360
ggggcttatg	atacagagga	ggcagctgct	cgtgcatatg	accttgacgc	tctcaaatac	420
tggggctcctg	aaacaacact	aaatttctca	gcggatgatt	atgggaagga	gaggtcggag	480
atggaggcgg	tgtcgcggga	ggagtacctc	gccgcccttc	gccgccggag	cagcggcttc	540
tcaaggggtg	tctccaagta	ccgcggggtc	gccaggcatc	accataacgg	gcggtgggag	600
gcgcgaatag	ggcgggtgct	ggggaacaag	tacctttacc	tgggcacctt	tgatactcaa	660
gaggaggcag	ccagggccta	tgatcttgca	gccatccaat	accgaggtgc	caatgcagta	720
acgaacttcg	acattagccg	ctacctggac	cagccacagt	tactggaaca	actgcagcag	780
cagcaggggc	cacaggtggt	ggcagcattg	caagaggaag	ctcaacgtga	tcaccaaagt	840
gacaatgcag	tccaagagct	caattcaggg	gaagcgcaga	cacctggtgg	catcgatgag	900
ccgattgcaa	taggcgacag	tactgaagac	atcaacacct	cccttacggt	tgatgacatc	960
attgaagaga	gcctctggag	tccttatgaa	tttgatatca	tggctggagt	gaacgtcagc	1020
aactcgatga	atctgagcga	gttgttcagc	gatgttgctt	ttgagggcaa	cattgggtgc	1080
ctgtttgaag	aatgctccgg	catcgatgac	tgcagcagca	ggcatggtgc	aggtttggca	1140

pctau2012001598-seq1

gcatttgggt	tgtttacgga	aggcgacgat	aagctgaagg	atgtttcaga	gatggagatg	1200
gaagtaaadc	ctcaagcaaa	tgatgtctcc	tgccctccaa	aatgatcac	tgtgtgcaat	1260
tgagcatcct	tgtgcgtgcg	tgagcatctt	tgtgtttctga	tgtccaagga	ggtttttagc	1320
tgtccatggt	gggtatat	cttagaatag	cttctttag	gctgtctgat	cagggtgggag	1380
agattcatgt	gctcctccct	atcctttgtg	aaaaccccg	aggctgctgc	caattgccat	1440
gatatgtcag	tgtttccttt	ctttttcccc	caattagcac	aaggctaata	tactaacctc	1500
aaaagatgtc	ctcacttaac	ctagattcaa	caaattatat	tgtggt		1546

<210> 242
 <211> 1500
 <212> DNA
 <213> Hordeum vulgare

<400> 242	
ggcctcccca	ctcttcctcg gccaccacc attgacacag agatcatgaa gagatccctt 60
cctccccagc	catctccttc ttcttcgccg gcatgctcgc cctcgccgtc ttctccttcc 120
tcgtccgact	cgtcttcgat agccattccc cgcaagcgag cgcgacgca gaaggccggg 180
agcgccaagg	cgaaggccgc gcccaagcgg gccaaagaa actcgggcag gagcaccaag 240
gactctgacg	cctccgccaa tggcgccgcc gcgtctggga agagaagctc catctacagg 300
ggggtcacaa	ggcacagggtg gacaggcaga tttgaggcac acctctggga caagaattgc 360
ttcacttcca	tccagaacaa gaagaaaggg aggcaagtct atctgggggc ttatgacaca 420
gaggaggcag	ctgctcgtgc gtatgacctt gcagctctca aatattgggg gcctgaaacc 480
acactgaatt	tcacggtgga tgagtacgcg aaggagaggt cagagatgga ggcggtgtcg 540
cgggaggagt	acctcgccgc gctccgccgc cggagcagcg gcttctccag gggggtctcc 600
aagtacaggg	gcgtcgccag gcaccaccac aatgggcgat gggaggcgcg aatcggaagg 660
gtgctgggga	acaagtacct gtacctggga acctttgata cccaagagga agcagccagg 720
gcctatgatc	ttgctgcat cgaataccgg ggcgcaaatt ccgtaaccaa ttttgacatt 780
agccgctacc	tggaccagcc gcagttactg gcgcaactgg agcagggggc acagggtggtg 840
ccagcattgc	aagaggagct tcaacatgat caccaaagtg acaatgcagt acaagagctc 900
aactctggtg	aagcgcagaa gccaggtagc gtcagcgagc cgattgcagt ggatgacaca 960
gacaatacag	gagacatcgg tgccccctt gtgtttgaca gcggcgtcga ggagaacctg 1020
tggagccctt	gcatggatta tgatgtggac cccatctttg gaccaacat cagcagctcg 1080
atgaatctga	gcgagtgggt caacgatccc gccttcgaga gcaacatcgg gtacatgttc 1140
gaaggatggt	cggacgtcga tgactgcagc accaggcatg gcgcaggctt gtcagcggtg 1200
ggttttctca	aggaaggcga cgataagctg aaggatgggt cagacatgga ggcagaaata 1260
actcctcagg	caaacgatgt ctctgccct ccaaaaatga tcaccgtgtg taattgagca 1320
tctctgtgct	ctcatgtcca aggaagggtt caggttttca tgttggttat acctcttagc 1380
tagcttcttg	taggctgtct gtctgtctga tcaggtcaga gaaattcatg tgttcattcc 1440

pctau2012001598-seq1

catcctttga gaaaaaaccc aaaggcattc aacttggtat ggcaagaatg ccaatgctcc 1500

<210> 243
 <211> 1669
 <212> DNA
 <213> Sorghum bicolor

<400> 243
 ctcagattca ccgccgcgcg cgtctcctcc ccctgcacca ccaccagtac accgttcatc 60
 gaggagaggg acgggacggg acggagacca ccatggacat ggagagatct cagcagcaga 120
 agtctccgac ggagtctccg ccgccgccgt cgccgtcgtc ctctcgtcc tctgtctccg 180
 cggacaccgt cctgccccct cccgggaaga ggcggagggc ggcaacgacg gccaaaggcca 240
 aggccggcgc caagcctaag agggcccgta aggacgccgc cgccgaggct gatcctcctc 300
 ctctccccgc cgccgccgcc gccgggaaga ggagctccgt ctacagggga gtcaccaggc 360
 acaggtggac ggggaggttc gaggcgcata tctgggacaa gcaactgcctc gccgcgctcc 420
 acaacaagaa gaaaggcagg caagtctatc tgggggcgta tgacagcgag gaggcagctg 480
 ctctgtgccta tgaccttgca gctctcaagt actgggggtcc tgagaccctg ctcaacttcc 540
 ctgtggagga ttactccagc gagatgccgg agatggaggg cgtgtcccgg gaggagtacc 600
 tggcctccct ccgccgcagg agcagcggct tctccagggg cgtctccaag tacagaggcg 660
 tcgccaggca tcaccacaac gggaggtggg aggcacggat tgggcgagtc tttgggaaca 720
 agtacctcta cttgggaaca ttgacactc aagaagaggc agccaaggcc tatgatcttg 780
 cggccattga ataccgaggt gtcaatgctg taaccaactt cgacattagc tgctacctgg 840
 accaccact gttcctggcg cagctccaac aggaaccaca ggtggtgccg gcaactcaacc 900
 aagaagctca acctgatcag agcgaacag aaactatagc ccaagagtct gtgtcaagcg 960
 aagcgaagac accagatgac aatgcagaac ctgatgacaa tgcagagcct gatgacatcg 1020
 cggagcccct catcacagtc gacgacagca tcgaagagag cttatggagc ctttgcattg 1080
 attatgagct agacaccatg tcaaggtcta actttggcag ctcgattaat ctgagcgagt 1140
 ggttcaatga tgcagacttc gacagcaaca ttggatgcct gttcgaatggg tgttctgcgg 1200
 ttgatgaagg cggcaaggat ggtgtaggcc tggcagattt cagtctgttg gaagatttca 1260
 gtctgtttga agcaggtgat ggtcagctga aggatgttct ttcagatatg gaagagggaa 1320
 tccaacctcc aacaatgatc agtgtgtgca attgattctg aaacctatgc gtcgcgtgcg 1380
 tgcctatgtg agatatccaa ggaggttttc acttttttagg tgtcctgtct tgggtatact 1440
 tgatttgtcc aatgccacag agccactgct gcaaaagcag ctgaactttc ttgtagaaag 1500
 cccatggcag cttggcatta gacaagtgtg ttgatgttcc ttaattcttt gaatttgttc 1560
 gaggtgctt gactaacctt aagggtttgt cattgtcgcg cttagcctag attcaactaa 1620
 tcacatcctg aatctgaagt aattgtgcag tatcacaatt ttacatggc 1669

<210> 244

pctau2012001598-seq1

<211> 1576

<212> DNA

<213> Zea mays

<400> 244

```

ctcccccgcc tcgccgccag tcagattcac caccggctcc cctgcacaac cgcgtccgcg      60
ctgcaccacc accgttcacg gaggaggagg ggggacggag accacggaca tggagagatc      120
tcaacggcag tctcctccgc caccgtcgcc gtctctctcc tcgtcctccg tctccgcgga      180
caccgtcctc gtccctcccc gaaagaggcg gagggcggcg acggccaagg ccggcgccga      240
gcctaataag aggatccgca aggaccccg cgcgcgcgcc gcggggaaga ggagctccgt      300
ctacagggga gtcaccaggc acaggtggac gggcaggttc gaggcgcacg tctgggacaa      360
gcactgcctc gccgcgctcc acaacaagaa gaaaggcagg caagtctacc tgggggcgta      420
tgacagcgag gaggcagctg ctctgtccta tgacctcgca gctctcaagt actggggctc      480
tgagactctg ctcaacttcc ctgtggagga ttactccagc gagatgccgg agatggaggc      540
cgtttcccgg gaggagtacc tggcctccct ccgccgcagg agcagcggct tctccagggg      600
cgtctccaag tacagaggcg tcgccaggca tcaccacaac gggaggtggg aggcacggat      660
tgggcgagtc tttgggaaca agtacctcta cttgggaaca tttgacactc aagaagaggc      720
agccaaggcc tatgaccttg cggccattga ataccgtggc gtcaatgctg taaccaactt      780
cgacatcagc tgctacctgg accacccgct gttcctggca cagctccaac aggagccaca      840
ggtggtgccc gcactcaacc aagaacctca acctgatcag agcgaaaccg gaactacaga      900
gcaagagccg gagtcaagcg aagccaagac accggatggc agtgcagaac ccgatgagaa      960
cgcggtgcct gacgacaccg cggagcccct cagcacagtc gacgacagca tcgaagaggg      1020
cttgtggagc ctttgcattg attacgagct agacaccatg tcgagaccaa actttggcag      1080
ctcaatcaat ctgagcgagt ggttcgctga cgcagacttc gactgcaaca tcgggtgcct      1140
gttcgatggg tgttctgcgg ctgacgaagg aagcaaggat ggtgtaggtc tggcagatct      1200
cagtctgttt gaggcaggtg atgtccagct gaaggatgtt ctttcggata tggaagaggg      1260
gatacaacct ccagcgatga tcagtgtgtg caactaatc tgggaaccga ggaggttttc      1320
gctttccagg tgtcctgtct tgggtaatcc ttgatctgtc taatgccaca gtgccactgc      1380
accagagcag ctgagaactt tctttagtaa agcccatggc agtttggcgt tagacaagtg      1440
tgtcgatgtt ctttaattct ttgaatttgc ccctaggctg cttggctaac gttaagggtt      1500
tgtcattgtc tcacttagcc tagattcaac taatcacatc ctgaatctga aaaaaaaaaa      1560
caaaaaaaaa aaaaaa                                1576

```

<210> 245

<211> 1242

<212> DNA

<213> Brachypodium distachyon

<400> 245

```

atggaggcct actgctccac cttggtgaaa gacgagctca tcaatggcgg cggcgggggg      60

```

pctau2012001598-seq1

tcagcaggag	ggatgcggta	ctgcgaggcg	gcgccgcgcg	tctcgccgcc	ggtggcgatc	120
aagagcgtga	agagacggaa	gagggagccg	ccggcggtgt	cggggatgac	gacggtgagc	180
ggcggcggga	aggacggcga	caagtcggcg	ggcaatgcgg	cggccaagag	gagctccagg	240
ttcaggggcg	tcagcaggca	tcgttgact	ggcaggttcg	aggcgcacct	gtgggacaaa	300
ggcacgtgga	acccgacgca	gaagaagaaa	ggaaagcaag	tttatcttg	ggcgtacaac	360
gaggaagaag	cggcggcgag	agcgtacgac	ctggcggcgc	tcaagtactg	gggacctacc	420
acctacacca	acttcccgg	ggtggactac	gagaaagagc	tgaaggatgat	gcagggcggtg	480
tcgaggggaag	aatacctagc	ctccataaga	aggaagagca	acgggttctc	gcgtggcggtg	540
tccaagtatc	gtggcggtgg	gcggcaccac	cacaacgggc	ggtgggaggc	ccggatcggg	600
cgtgtcttcg	ggaacaagta	cctgtacct	ggcacctaca	gcacccagga	ggaggccgcc	660
cgcgcttacg	acatcgccgc	cattgagtac	cggggcatca	acgtgtcac	caacttcgac	720
ctcagctcct	acatccgatg	gctcaagccc	aactccacca	tcaacaccaa	cacccccgcc	780
gccgaattag	ccatcctcgg	cggcggcggc	acgccgcgg	caatgatcac	gccgccgccg	840
acgatgcacg	ttccccgcct	gctaccgccg	ctcgtaaagg	gaaggggtag	ctcgatcgcc	900
gacgatgtgt	ctgccgggtc	ctgcgtgttc	ggcggcccgt	cgccgtcacc	gtcaccaacc	960
acgacggcgc	tcagtttgct	gttgcggtca	tccgtgttcc	aggagttggt	cgcgagcag	1020
cagcctccgt	cgaccgtgga	tgacgatgat	gacatcggtg	gtcacgcggc	ggtgtccgac	1080
gcggcccaa	gagcagcgga	ggagaacgag	gagtcattcg	gcgaggtggt	gtacggtgcc	1140
ggcgaggggg	aggcggcaac	ggcattctcg	tgctccatgt	acgagctggg	cctggacgac	1200
aatttcgcgc	gcatcgagga	gtccctctgg	ggctgcttgt	ga		1242

<210> 246
 <211> 1467
 <212> DNA
 <213> Sorghum bicolor

<400> 246						
atggaacct	acagcctgca	ggtgaaggat	gagctccacg	gcggcggcat	cggcatcgga	60
ggaggagggc	aggggctgta	ctgcggcgct	acgccgaggc	cggcggctcc	tgccggccacc	120
ggcggcggcg	gcggcggcgg	ggacggtgct	gtgaagagta	ataagaggag	taggaagagg	180
gagccgccgc	cgccgccgcc	gtcgtccctg	gtgacgatga	gcaacggcgg	caaggacgag	240
gcggtcgccg	gcagcggcga	caagtcggcg	agtagcaaca	gtaacgcgtc	caagaggagc	300
tcaaggttca	gaggcgtaag	caggcatcgg	tggacgggca	ggttcgaggc	gcacctgtgg	360
gacaaaggga	catggaaccc	aacgcagaag	aagaaaggga	agcaagttta	ccttgagcgc	420
tacaacgagg	aagacgcggc	agcaagagcg	tatgatctgg	cagcgctcaa	gtactgggga	480
cctaccacct	acaccaactt	tccggctcgtg	gactacgaga	gagagctcaa	agtgatgcag	540
aatgtgtcca	aggaagagta	ccttgccctc	atacgaagga	agagcaacgg	cttctccgc	600
ggagtttcca	agtacagggg	cgtcgcacgg	caccaccaca	acggcagatg	ggaggcacgg	660

pctau2012001598-seq1

atcggccggg	tggtcggcaa	caagtacctg	tacctgggca	cgtacagcac	gcaggaggag	720
gcggcgcgcg	cctacgacat	cgccgccatc	gagtaccgcg	gcatcaacgc	cgtcaccaac	780
ttcgacctca	gcacctacat	ccgctggctc	aagcccggcg	gcggcgtcga	ggacagcgcc	840
gccggcaccc	cgacgtcggg	cgtcagggcg	ccgggtatac	cgccggcgtc	gctgtcccta	900
caggctggcg	gcctcctgca	gcacccgcac	ggcgccgcg	ccggcatgct	gcaggctcgac	960
gtcgacgacc	tctaccgcg	gcagctggcc	gcggcgcgctg	gggctgctct	gttctctggc	1020
ggcatcgacg	acgtgggctc	cgtgtacgca	gctggcagcg	cagggcccag	cccgcaggct	1080
ctgtgcgccg	ggaggccgtc	tccatcgccg	tcgccgtcgt	cgtccaccac	ggcgctcagc	1140
ctgctgctgc	gctcgtcagt	gtttcaggag	ctcgtggcgc	ggaacgccgg	cgcgggcgcc	1200
gcgcagcagc	agcagctcgt	ggtcgcggac	gacgacgggg	cagtatcgcc	ggcggacgtt	1260
gttgacgcca	aagttgagca	gccggaggca	gagggcgagc	tgggtcgtca	cggtgaccag	1320
ctctacggcg	cggcgagggc	ggacgaggac	gaggacgcgt	tcgcctgctc	catgtacgag	1380
cttgacgaca	gcttcgcgcg	catggagcag	tcgctctggg	gctgcctacg	atcatcggat	1440
gcacctgata	atatgaacaa	cctctga				1467

<210> 247
 <211> 1332
 <212> DNA
 <213> Sorghum bicolor

<400> 247	
atggagagca	gtggaatgat gatggtgaag agcgagatcg agagctgcgg ctacccgggg 60
ccgtcgtcgt	cgacggcacc tgcggcaggc gtggtcatcg gcggcagcgc gacgacggaa 120
cgcggcgagg	gggggcacca tcatcatcat catcaggtgg tggtagaggag gaggaggagg 180
gagccgcctc	tgctcgcacc aatcgccggc ggcggcattg gcaagccgct gccagcatc 240
accgtcaaga	ggagctccag gttccgtgga gtcagcaggc atcgatggac aggtcgtttt 300
gaggcacatc	tgtgggacaa gaactcgtgg aatccgacgc agcggaagaa gggcaaacag 360
gtctatctgg	gagcttacga tgaggaggag gctgctgcca gacacctacga ccttgctgct 420
ctcaagtact	ggggggccac cacctacacg aacttcccg tcatggacta cgagaaggag 480
ctcaagataa	tggagaacct caciaaggaa gagtaccttg catcactgag gaggaagagc 540
agtggttttt	caagaggagt gtccaagtac agaggcggtg ccagacatca tcagaacggg 600
aggtagggagg	caagaattgg gcgcgttttt gggaacaagt acctttacct aggaacttac 660
agcactcaag	aagaagccgc ccgcgcatac gacatcgcg cgatagagta caaaggcggtg 720
aacgcggtga	ccaacttcga cctgagggtc tacatcacat ggctgaaacc cagcggtgct 780
ccggccgcct	tcaaccggga ggccggcgctc ctgatgcagg cagctccggc cgagcagctc 840
ctccacccgg	cagagacggc ccagatgctc ccacgagtcg gcaaccggtt cctgctggac 900
cacggcgccg	cgccgccggg cagtagcggc ggccggcgggc aggacgcctc catgtcgtcc 960

pctau2012001598-seq1

atggtgtctc ccggtgccgg cggcgccatg aggaggagag gctcctccac cgcgctcagc	1020
ctgctgtctca agtcgtccat gttcaggcag ctggctcgaga agaactccga cgccgaggag	1080
ggcgtcaggg acaggggaaga tgccgccgcc gccgccgctg ctgcgcatcc agctggacca	1140
ggagacgcct atgagtacca caacttcttc cagggagaag cgcctccgga catgtgcat	1200
ctcttctcgt caggcggcgg tggcgaccac gcgcggaatg ctgggttcca tggagagatc	1260
gccgcgtgct acgatgatgg cgaggggttg gatggctgga atgggtttgg caacatgtca	1320
tctctgcaat ag	1332

<210> 248
 <211> 1271
 <212> DNA
 <213> Glycine max

<400> 248	
aggtaaatta cattgttgtc tgctgaaata aaaaacagaa gagtgaaatg gagcttgcac	60
ctgtgaagtc tgaactaagt ccaaggagcc ataggttgct catgatagat ggtagtgagg	120
ttattggcac taagtgtgtc aaaaggcggc gaagagattc atctacggcc gtgttaggtg	180
gcaatggaca acaaggtgaa cagttagaag aacagaagca acttggtggc caatcaactg	240
ccaccactgt gaagagaagc tcaaggttca ggggtgttag cagacacagg tggactggaa	300
ggtttgaagc acatctatgg gataaaggga cctggaatcc cactcagaag aagaaaggaa	360
agcaagtata tttgggagct tataatgatg aagaagctgc agctagagct tatgatttgg	420
ctgcactcaa gtactgggga atatcaactt tcacaaattt tcctgtatct gattatgaga	480
aagaaattga gataatgaaa actgtaacca aagaagaata tcttgcttca ttgagaagga	540
ggagcagtgg tttttccaga ggtgtatcga agtatagagg agtagcaagg catcatcaca	600
atggaagatg ggaagcaagg attgggagag tgttcggtaa caagtacctc taccttggga	660
cttacagtac acaagaagaa gcagctcgcg catatgacat tgcagcaatt gaatacagag	720
gcataaatgc agtgacaaac tttgacttga gcacctacat aagatggcta agaccgggaa	780
cacatcctac tgcttctcat gatcaaaagc ctagcactga tgctcaacct tttgcaacct	840
ctaactccat gcaagcaaga gggaacattg aggtatccaa ctccaacaag aattcattcc	900
cctcaggtaa attggacagt accaagaagc gagacttttc caagtacatg aaccctttga	960
gtccatgcaa caagccatct tccccaacag cattaggact tctcctaaaa tcctcagtgt	1020
ttagagaact gatgcagaga aatctgaact cttctagtga agaagctgaa gaagttgaat	1080
tgaaatatcc acatgagggc aatgatggag ttggagggat ttatgataat gaaaacacca	1140
ataactctta cttttgctct tctaatatca gcagattacc taacttggag tcatcagaag	1200
agagtccatt gcctatgtat catggaactg tgcaatcact atggaatagt gctttcaaca	1260
tgtctaactg a	1271

<210> 249
 <211> 1266

pctau2012001598-seql

<212> DNA
 <213> Glycine max

<400> 249
 caggtaaatt acattgttgt ctgctgagaa aaacagaaga gtgaaatgga gcttgcacct 60
 gtgaagtctg aactaagtcc aaggagccat aggttgggtca ttatagatgg tagtgatggt 120
 attagcacta agtgtgccaa aaggcggcga agagattcgt ctatggccgt gttaggtggc 180
 aatggacaac aaggtgaaca gttagaagaa cagaagcaac ttggtggcca atcaactgcc 240
 accactgtga agagaagctc aaggttcagg ggagttagca gacacaggtg gactggaagg 300
 tttgaagcac atctatggga taaagggacc tggaatccca ctgagaagaa gaaaggaaag 360
 caagtatatt tgggagctta taatgatgaa gaagctgctg ctagagctta tgatttggct 420
 gcactcaagt actggggaac atcaactttc acaaatcttc ctgtatctga ttatgagaaa 480
 gaaattgaga taatgaaaac tgtaaccaa gaagaatacc ttgcttcatt gaggaggagg 540
 agcagcggtt tttctagagg tgtatcaaag tatagaggag tagcaaggca tcatcacaat 600
 ggaagatggg aagcaaggat tgggagagtg ttcggcaaca agtacctcta ccttgggact 660
 tacagtacac aagaagaagc agcccgtgca tatgacattg cagcaattga atacagaggc 720
 ataaatgcag tgacaaattt tgacttgagc acctacatca gatggctaag accaggaaca 780
 catcctactg cttctcatga tcaaaagcct agcactgatg ctcaactttt tgcaacctct 840
 aactccatgc aaacaagagg gaacattgag gtatccaact ccaacatgca ttcattcccc 900
 tcaggtgaat tggacagtac caagaagcga gacttttcca agtacatgaa ccctttgagt 960
 ccttgaaca agccatcttc cccaacagca ttaggacttc tcctaaaatc ctcggtgttt 1020
 agagaactga tgcagagaaa tctgaactct tctagtgaag aagctgatgt tgaattgaaa 1080
 tatccacaag agggcaatga tggggttgga gggatttatg ataatgacaa caccagtaac 1140
 tcttactttt gctcttctaa tatcagcaga ttacctaaact tggagtcatc agaagagtgt 1200
 ccattgccta tgtatcatgg aactatgcaa tcactatgga atagtgtctt caacatgtct 1260
 aactga 1266

<210> 250
 <211> 1257
 <212> DNA
 <213> Populus trichocarpa

<400> 250
 atggagatga caaggaatac tggcgatcag ataagcctag gaaggcgctg cttgtgcatg 60
 atagaagaag aaaggagagc cggggaggct ggtaagtgtt ttaaaagacg acgaagagac 120
 ccttccacat ttgcattaag ctgcaatatt aatgatcaac aaagtgacca gcagcagcag 180
 caacagtcac ttggtgatcg aactgcagct gttgccacta ctgtcaagag aagttcaaga 240
 tttcggggag ttagcaggca ccgttggacc ggacgatttg aagctcattt atgggataaa 300
 ggaacttggg atccaacaca gagaaagaag ggtaaacaag gagcttatga tgaagaagaa 360
 tccgctgcaa gagcatatga tttagctgct ctcaagtact ggggaacgtc aactttcaca 420

pctau2012001598-seq1

aattttccgg	cgtctgacta	cgagaaagag	attgaaataa	tgaaaacggt	aaccaaagaa	480
gaatatctag	cctccttgag	aaggagaagc	agtgggtttt	caagagggtgt	atccaaatac	540
agaggagttg	ctaggcacca	tcacaatggg	agatgggaag	caaggatagg	gagagtgttt	600
ggaaacaaat	atctttatct	tggcacgtac	agcactcaag	aggaagctgc	tcatgcctat	660
gatatagcag	ccattgaata	cagagggatc	aatgctgtaa	ccaatttcga	tttaagcaca	720
tacatcagat	ggctgaagcc	agaagccagt	cttcctgcac	ctcaaactca	agaatcgaag	780
ccagcttcag	accccttgcc	gatggcaacc	ttctccaatc	accttccgag	cgagaaacca	840
accagctat	ccgtttctca	gatggaccct	tctttaatgg	ataattttaa	cacccttaaa	900
aacgaggaca	tttttcacag	gaagacgctt	cccgttagcc	ccctcactag	gtcatcctcc	960
tcgactgccc	ttagcctcct	ctttaaatct	tcaatattca	aagagcttgt	ggagaagaac	1020
ttgaatacca	cctccgaaga	aattgaagaa	aatgattcga	aaaacccgca	taatggcaac	1080
aacaatgctg	gggaggcatt	ctatgatgga	ctcagcccca	ttcctcacac	aggcacttca	1140
accgaggatc	cgttcttggt	ttcagagcaa	ggagagacaa	atacattgcc	accatacagt	1200
ggaatggagc	aatctctttg	gaatggagcc	ttgagcatgc	cttctcgttt	ccattaa	1257

<210> 251
 <211> 1011
 <212> DNA
 <213> Vitis vinifera

<400> 251						
atggagatga	ctacagtga	atctgaatta	ggcctggaaa	ggggccgctt	gtgcactgca	60
gaaactgatg	ctttggaggt	gaccaagtgt	gttaagagaa	ggagaagaga	tccatcagca	120
gttacacccg	gatgcagtaa	acaaggagag	cagcagaagc	aggtgttgct	gcaggctggg	180
caatccatta	ctgccattgc	cactactatg	aagagaagtt	caagattcag	gggtgttagc	240
agacatagat	ggacagggag	gtttgaggct	catttatggg	ataaagggtc	ctggaatgtg	300
acacagagaa	aaaagggaaa	gcaagtctat	cttgagcctt	atgatgaaga	agaatctgct	360
gcaagagcgt	atgacttggc	cgactgaag	tattgggggc	catcaacttt	taccaatttc	420
ccagtatcag	attatgagaa	ggagattgaa	ataatgcagg	gtctaacaaa	ggaggaatac	480
ttggcttctt	taagaaggag	aagcagtggg	ttttcaagag	gagtgtccaa	gtacagagga	540
gttgcaaggc	atcatcacia	cggaagatgg	gaagcaagga	ttgggagggt	gtttggaac	600
aagtatctct	accttgaac	ttacagtacc	caagaagagg	ctgctcatgc	ttatgacatt	660
gcagcaattg	agtacagagg	gatcaatgct	gtaacaaact	tcgagttgag	cacatacgta	720
aggtggctga	ggcccagagc	cactgctctt	accctcagg	aaccaagatc	aaattcaata	780
atgcaagcat	cctctaactg	ccttccaaat	gaggaagtag	agctgtcctt	cctcagcccc	840
aatcctttta	cagtagatga	cttggcaact	cctctaaagc	aagaaaagtt	ccaagagag	900
gtttccatca	gcccttgac	caagtcatct	tcgccaactg	cactaagtct	cctccataga	960

tcttcagtat ttagacaact ggtggagaag aattcgaaca gcatcgaata a 1011

<210> 252
 <211> 1170
 <212> DNA
 <213> Glycine max

<400> 252
 atggcaatga tgaaagagaa tattattgaa gtgagccttg gaaggaggca aatgtctatg 60
 acagaagggtg agttccaagg aacacggagc gtcaagagac gacgaagaga ggttgctgca 120
 gccgcaggca gtgggtgatga taaccaccaa cagcagttac cacaacaaga agttgggtgaa 180
 aatactacag tcaacactac aaagagaagc tcaagatttc gcggtgttag cagacataga 240
 tggacgggtc gatatgaagc tcacttgttg gacaagcttt cttggaacat aacacaaaag 300
 aagaaaggga agcaaggggc ttatgatgaa gaagaatctg ctgccagagc atatgattta 360
 gctgcactta agtattgggg aacttcaact ttcacaaatt ttccaatatc agattatgag 420
 aaagagatac aaataatgca gactatgacc aaagaggagt atttggccac ttttaaggaga 480
 aagagcagtg gcttttcaag aggtgtatca aagtataggg gtgttgcgag gcaccaccac 540
 aatggtagat gggaagcaag aataggaagg gtttttggaa acaaatatct ctaccttgga 600
 acctacagca cccaagagga agctgcgcgt gcttatgaca tagcagccat tgagtacaga 660
 ggaattcatg ctgtaacaaa ctttgatttg agcacttaca taaaatgggt aaagccttca 720
 ggaggaggca ccctagaagc aaatcttgaa tcacatgcag cactagagca tcaaaagggtg 780
 gcatcccat ctaattatgc tctaacagaa gagtctaagt ctttggccct tcacaacagt 840
 tttttcagtc catattctct ggattcccct gtaaagcatg aaagatttgg aaacaaaacc 900
 taccaattct caagcaataa gtcattctct cccactgcac ttggtctcct tctccgctct 960
 tcgttattta gggaattagt tgaaaagaat tcaaatgttt ctggagaaga agatgatggg 1020
 gaagccacaa aggatcaaca gacacaaata gctactgatg atgatctggg tggaatcttc 1080
 ttcgatagct ttagtgatat tccatttgtg tgtgaccca atagatacga cttggaattg 1140
 caagaaagag acctgcactc aatattttga 1170

<210> 253
 <211> 1170
 <212> DNA
 <213> Glycine max

<400> 253
 atggtaatga tgaaagagaa tattattgaa gagaagcttg gaaggagcca aatgtctatg 60
 gtagaagggtg agttccaagg aacatgggggt gtcaagagac gacgaagaga ggttgctgca 120
 gcagcaagca gcggtgatga taaccaccac cagcagttac cacagcaaga agttgggtgaa 180
 aattcttcaa tcagcacaac aaagagaagc tcaagatttc gcggtgttag cagacataga 240
 tggacgggtc ggtatgaagc tcacttgttg gacaaactct cttggaacat aacacagaag 300
 aagaaaggga agcaaggggc ttatgatgaa gaagaatctg ctgccagagc atatgattta 360

pctau2012001598-seq1

gctgcactta agtattgggg aaattcaact ttcaccaatt ttccaatatc agattatgag	420
aaagagatag aaataatgca gactatgact aaagaggagt atttggccac ttttaaggaga	480
aagagcagtg gcttttcaag aggtgtatca aagtatcggg gcgttgcaag gcaccaccac	540
aatggtagat gggaagcaag aataggaagg gtttttggaa acaaatatct ctaccttggc	600
acctacagca cccaagaaga agctgcgcgt gcttatgaca tagcagccat tgagtacaga	660
ggaattcatg ctgtaacaaa ctttgatttg agcacttaca taaaatgggt aaagccttca	720
ggaggaggca cccagaaga aaatcttgaa tcacatgcag tactagagca tcaaaagttg	780
gcatcccat ctaattatgc tctaacagaa gagtctaagt ctttggtcct tccaacagt	840
tttatcagtc cagattctct ggattcacct gtaaagcatg aaagttttgg aaacaaaacc	900
taccaattct caagaaataa gtcattctct cccactgcac ttggtctcct tctgcgctct	960
tcattattta gagaattggt tgaaaagaat tcaaatgttt ctggagaaga agctgatggg	1020
gaagtcacaa aagatcaaca gccacaacta gctagcgatg atgatctgga tggaatcttc	1080
ttcgatagct ttggcgatat tccatttgtg tgtgatccca ctagatacaa cttggaattg	1140
caggaaagag acctgcactc aatattttga	1170

<210> 254
 <211> 1179
 <212> DNA
 <213> *Medicago truncatula*

<400> 254

atggcaatgt tgatagaaaa cgaagtcatg tgtttaggaa agagtcaaag atccatggat	60
ggaaaagagg taaaaggagc acgtcgtgtt aagagacaaa gaagagatgc tattgtacca	120
aaaataggtg atgatgctaa caagatggca cagaaacaag ttggagaaaa ttctacaact	180
aacacttcaa agagaagctc aagatttcgt ggtgttagca ggcatagatg gacgggtcga	240
tttgaagctc acttgtggga caagctttct tggaacacaa cacaaaagaa gaaagggag	300
caaggggctt atgatgaaga agaactctgct gcaagagcat atgatttagc tgcacttaag	360
tattggggaa catcaacttt cactaatttt ccgatatcag attatgataa agagatagag	420
ataatgaata ctatgaccaa agaggagtac ttggccacat taagaagaaa aagcagtggc	480
ttttcaagag gtgtctcaaa gtatagaggt gttgcaaggc accatcataa tggtagatgg	540
gaagcaagga tagggagagt ttttggaac aaatatctct atcttggaa ctatagtaca	600
caagaagaag ctgcacgtgc ttatgatata gctgcaattg agtacagagg aattcatgct	660
gtgaccaatt ttgaattgag tagctacata aaatggttaa agccagaaac tactacagaa	720
gaaaatcatg aatcacaaat attacaaaaa gagtctcgaa cattagcacc ccctaataac	780
tctactctat tgcaagaatc aaagttatta gcccttcaaa agagcttttt catttcaaat	840
gacctgaatt caactgaaaa gcaagaatca tcatattgaaa acaaaaacta tcatattcttg	900
agcaacaagt caacttctcc aaccgcgctt agccttctcc ttcgctcttc gttgtttaga	960
gaattacttg aaaagaattc aatgtatct gaggatgaag ttacaaaaga gcaacaacaa	1020

pctau2012001598-seq1

caacaaataa ctagtgatga tgaacttgga ggaatcttct atgatggtat tgacaatatt	1080
tcatttgact ttgatcccaa cagctgcaac atagaattgc aagaaagaga cttgcattca	1140
atatcttggtt tatatcaata cttaaatttt ggacaataa	1179

<210> 255
 <211> 1161
 <212> DNA
 <213> Populus trichocarpa

<400> 255	
atgatgatga taaagaatga agaaaaccca ggaaggcgcc gcggatgcat tgcagacagt	60
gaagctcagg tagctcggtg tgtcaaaagg cggcgaaggg atcctgctat agttgcactc	120
ggaagtgacg ataatcaaag tcagcagcag atgccccaga acaaaactga tcaaacttct	180
gctgctacca ctgtgaagag aagctcaaga tttcgtggcg tcagcagaca tagatggaca	240
ggtcggtttg aagcccacct gtgggataag ctctcatgga atgtaacaca gaagaagaaa	300
gggaaacaag gagcttatga tgaagaagaa tccgcggcga gagcatatga tttggctgct	360
ttgaagtatt ggggaacatc aacttttacc aattttccga tatcagatta tgagaaagag	420
attgagataa tgcaaactgt aacgaaagag gagtathtag catctctaag aagaaagagc	480
agcggttttt caagaggtgt atccaagtat aggggagttg caaggcacca tcacaatggg	540
agatgggaag cgagaatagg gagagttttc ggaacaaat acctttacct tggcacatac	600
agcacgcaag aggaggctgc acgtgcttat gacattgcag caattgagta cagagggatc	660
aatgcggtga cgaactttga cttgagcact tatatcagat ggataaagcc aggagtagct	720
gctcaagctg ctgcaaatga actacagaca gtcacagatc ctcagacagc agcgacattg	780
actgatacct aactccaag agaggaaacc aaaccctcac tgtttctacc taatcagttc	840
actgcagatt acctgaactc cccaccaaag ctagacgcat ttcagaacaa tatctttgta	900
gattcttcca acaagacttc atcgccact gccctcagcc tcctcctacg gtcttcggta	960
tttagagaat tgggtgagaa gaattcaaat gtttgtgagg aggaaactga tgggaatgag	1020
ataaagaacc aaccaatggc aggtagtgat gatgaatatg gtgggatatt ttatgatgga	1080
attggagata ttccatttgt ttactcttct aacaaataca gcctaggggtt ggaagaaaga	1140
gaactccagt tcgtcttata a	1161

<210> 256
 <211> 1119
 <212> DNA
 <213> Ricinus communis

<400> 256	
atggaaatga tgatggtgaa gaatgaagag atctcaggaa ggcgccgtgc gagtggtaca	60
gaaagtgagg cttacgtagc tcgttgtgtc aaaagacggc gcagagatgc agctgtagtt	120
acagttggtg gtgatgacag tcaaagtcac cagcagcagc agcagcagca gccagagcaa	180
caggctcatc aaatttctgc tgctaccacc gtgaaaagaa gttcaagata tcgtggtgtc	240

pctau2012001598-seq1

agcaggcata	gatggacagg	cagatttgaa	gctcaccttt	gggataagct	ctcgtggaat	300
gtaacacaaa	agaagaaagg	aaaacaagga	gcttatgatg	aagaagaatc	tgcggaagg	360
gcatatgatt	tagctgcatt	gaagtattgg	ggaacatcaa	ctttcaccaa	ttttccgata	420
tcggattatg	agaaagaaat	tgagataatg	caaactgtaa	caaaagagga	atatttggcc	480
tcgttaagaa	gaaagagcag	tggcttctca	aggggggtgt	ccaaatacag	gggagttgca	540
aggcaccacc	ataatgggag	atgggaagca	agaataggga	gagtttttgg	aaacaaatac	600
ctttatcttg	gcacctatag	cacacaagag	gaggccgctc	gtgcatatga	cattgctgca	660
attgaatata	gaggaatcaa	tgcagtgacc	aactttgatt	tgagtactta	catcagatgg	720
ctaaagccag	aagttgctgc	ccaagtcgca	gcaaataaac	cgagactgt	tgcagagtct	780
cggatgctgc	catcgatcaa	taaccgtatt	gcaagagaag	aatctaaacc	ttcattcttc	840
agcgctactc	ctttcagttt	agattgctgg	agttaccctc	gtaaacaaga	agagttccaa	900
aacagaactc	ccattacccc	ttgcagcaag	acgtcatcac	ccactgcact	tagtctcctc	960
cttcggtcct	cgatattcag	ggaattagtt	gagaagaatt	ccaatgtctc	tgaagatgaa	1020
aatgaagggg	aagagacgaa	gaaccaatca	caaataggca	gtgatgatga	gtttggggga	1080
ctgttttacg	aaagaattgg	tgacatccca	tttatttaa			1119

<210> 257

<211> 1137

<212> DNA

<213> Brachypodium distachyon

<400> 257

atggccaagc	aacgcacgga	cagcgccggc	accgatgccg	ccgccgtcca	attgacgaag	60
ccgaagcggg	cgcggaagag	cgtgccgagg	cgaggagacc	cctcgcgccg	cacctcggcg	120
taccgcggcg	tcacgcggca	ccggtggacg	gggcgggttcg	aggcgcacct	gtgggacaag	180
aacacctgga	cccagtcgca	gaggaagaag	aagggcaggc	aagtttatct	cggggcgtac	240
ggcggcgagg	aagcggcggc	gcgcgcctac	gacctggcgg	cgttgaagta	ctggggccgc	300
gacaccgtcc	tcaacttccc	gttgtctaac	tacgacgaag	agtggaaaga	aatggagggg	360
cagtccaggg	agaatacat	cggatccctc	cggaggaaaa	gcactggctt	ctcaagaggg	420
gtgtccaagt	acagaggcgt	cgcaaggcat	caccacaacg	gcaaataggg	ggcgaggatt	480
gggcgtgtct	acggcaacaa	gtatctctac	ctgggcacat	atgggacgca	ggaggaggcg	540
gccatggcgt	acgacatcgc	ggcgatcgag	caccgcgggc	tcaacgccgt	caccaacttc	600
gacgtcagcc	gctacatcga	ctggcaccgc	cgctcttgcc	gcgacctcgg	cgacaacatc	660
atcacacccc	taacgaacct	aaccgtcgac	ctcgaggaag	ccatggcccg	cgacgacgac	720
gacggccaat	ttctgtctcc	gtcccaggcg	acgacgccgc	cgtccacgtc	gtccgcgctc	780
ggcctgctgc	tgctgtcgcc	gaggctcaag	gaggtgatcg	agggatcagg	ggccgcgagc	840
gcgatggccg	cgtcgacttc	cgagagcagc	gccgccgggt	cccccccgcc	gtcttggtct	900

pctau2012001598-seq1

tcgtcttcct gttccccctc cccgccctcg ccctcgact cgccgccgga gacccagcag	960
aagcagcagc agcaggaata cggcgcttcg gctgcggcgg cgcggtgcag cttcccggac	1020
gacgtgcaga cgtacttcgg gtgcgaggac ggctgcgcgg aggtggacac cttcttgttc	1080
ggggacctga gcgcctacgc cgcgcccattg ttccagttcg agctgctgga cgtttga	1137

<210> 258
 <211> 1323
 <212> DNA
 <213> Sorghum bicolor

<400> 258	
atggccaggc cgcgcaagaa cgccggtacc gacgaggata atcctaacgc cgccacgggc	60
gtcagcgtca ccgggaagcc gccgaagctg aagcgggtgc ggaggaaggg cgagccgcgg	120
gagagcagca ccccgctcca gcgcagctcc gcgtaccgcg gcgtcacacg gcaccggtgg	180
acggggcggt tcgaggcgca cctgtgggac aaggacgccc gcaatggctc ccggaacaag	240
aaaggcaagc aaggtgcgta cgacgacgag gaagcggcgg cgcgcgcgca cgacctggcg	300
gcgctcaagt actggggccc cgccaccgtc ctcaacttcc cgctgtgcgg ctacgacgag	360
gagctgaggg agatggaggc gcagcccagg gaggagtaca tcgggtcgct gcgtcggcgg	420
agcagcggct tctcgcgagg ggtgtccaag tacaggggcg ttgccaggca tcatcacaac	480
gggcgatggg aggcgaggat cgggcgcgtg ctgggcaaca agtacctcta cctcggcact	540
ttcgcaacgc aagaggaggc ggcggtggcg tacgacatcg ccgctataga gcaccgtggc	600
ctcaacgcgg tcaccaactt cgacatcagc cactacgtca atcactggca tcgacactgc	660
cacggcccct ccgacgacag cctcggcgtg gtggtggacg acgtcgccgc gttccagcta	720
ccagatgacc tgcccagagt cccggcggcc gccatcggcg tcgaggagac cacgggcggc	780
gacgcggaat tccacaacgg cgaggagggc tacctgcagc accacaccag cggacccttc	840
ggcgcgcagc agcagcttcc tgacgagacg ggagcgctag cggctcatca aatggcgcca	900
aattcgtcgg cgctcgacat ggtgttgagc tctccaagt tcaaggagct gatggagcag	960
gtctccgcgg cggcggcggc ggtggcctcc gagagcagca tcggcggcag catgagcagc	1020
agctcgccgt cgccgtcttt gtccctcgtt tcgccttcgc ctctgcagct gccatcgccg	1080
tcgtctttgt cctcgttctc gccttcgtcg cctctgcagc agccatcgcc gccgctgcag	1140
cagcctgagt tcgtcgaagg agcgcggcg gctcgggtgca gcttcccggg cgacgtgcag	1200
acgttcttcg atttcgagaa cgagagcgac atgagcttca tgtacgcgga ggttgacacg	1260
ttcctcttcg gggacctggg cgcgtacgcg gcgccgatat tccacttcga cctagacgtg	1320
tga	1323

<210> 259
 <211> 1532
 <212> DNA
 <213> Zea mays

<400> 259

pctau2012001598-seq1

ccgctgctca	ccccgtccta	gacgcgaggc	gacgacgcct	ccctatccct	ctcctccgca	60
atggccaggc	cccgaagaa	cggcggtacc	gacgaggatg	atgctaacgc	cgccacggga	120
gccaccggga	agccgaagaa	gctgatgaag	cgggcgcgga	ggaagagcga	gagcccctcc	180
ccgcgcagct	ccgcgtaccg	cggcgtcaca	cggcaccggt	ggacggggcg	gttcgaggcg	240
cacctgtggg	acaaggacgc	ccgcaacggc	tcccggagca	agaaaggcaa	gcaagtttat	300
ctcggcgctg	acgacgacga	ggacgcggcg	gcgcgcgcgc	acgacctggc	ggcgctcaag	360
tactggggcc	ccgccggcac	cgtcctcaac	ttcccgtgta	gctggctacga	cgaggagcgg	420
agggagatgg	aggggcagcc	ccgggaggag	tacgtcgcgt	cgctgcggcg	gcggagcagc	480
ggcttcgcgc	gaggggtgtc	caagtacagg	ggcgttgcca	ggcatcatca	caacggggaga	540
tgggaggcga	ggatcggggc	cgtgttaggc	aacaagtacc	tctatctcgg	cacttacgca	600
acacaagagg	aggcggcagt	ggcgtatgac	atggcggcta	tagaacaccg	cggcttcaac	660
gcggtcacca	acttcgacat	cagccattac	atcaaccact	ggcatcgaca	ttgccatggc	720
ccctgcgacg	gcagcctcgg	cgcgatggac	gtcgcaccga	acgtctcgct	cgaactagac	780
ctactcgagt	gcccggccac	cgtcggcctc	ggcctcgagg	agaccacggg	cgacgatgaa	840
ttccacaacc	gcgaggacta	cctggggcac	cttttcggcg	tgcagcagct	tcccgcagag	900
atggggccgc	cggctcatca	gatggcacca	gcctcgtcgg	cgctcgatct	ggtgttgag	960
tcgcctaggt	tcaaggagct	gatgcaacag	gtttccgcgg	cgggagcctc	cgagaccaac	1020
ggcggcagca	tgcgtagctc	gccatcaacg	tccctgtgtt	cgttctcgcc	ttcgctcttg	1080
gagctgcat	cgccgccgct	acagcaacct	actgaattca	tcgatggcgc	gccgccgcgg	1140
tgcagcttcc	cggatgacgt	gcagtcgttc	ttcgatttca	agaacgataa	cgacatgagc	1200
ttcgtctacg	cggaggttga	cacattcctc	ttcggcgacc	taggcgcata	cgctccgcca	1260
atgttcgact	tcgacctgta	cgagtgatca	ggatccgact	gccagcgttc	caagacgcca	1320
gtgttggtta	aatggcttca	taatttttgt	ttatgatgat	gatgagatga	ctgctatcaa	1380
ggtttacttg	aaagaggtag	gattagttat	ttattctcgc	ccttttttca	aatagctaga	1440
agaggaaaat	atgcttctct	ttttttcttt	tttccctaaa	aggctcaaag	gtgtatacaa	1500
gtcgtcacga	agaactaatt	gacattgatg	ac			1532

<210> 260
 <211> 977
 <212> DNA
 <213> Arabidopsis lyrata

atcctttatt	attctcattc	accggtaatg	gcaaaagtct	ctcggaggag	caagaaaaca	60
atcgttgaag	atgaaatcag	cgataaaaca	gcgtctgcat	ctgaggctgc	gtccattgtc	120
tttaaatacca	aacgcaaacg	taagtcgccg	cctcgaaacg	ctcctcctca	acgcagctcc	180
ccttatagag	gcgtcacaag	gcatagatgg	actgggagat	acgaagcgca	tttgtgggat	240
aagaacagct	ggaacgagac	tcaaaccaaa	aaaggacgac	aagtttatat	aggggcttac	300

pctau2012001598-seql

gacgaagaag aagcagcagc acgtgcctac gacttagcag cattgaagta ctggggacga	360
gacacactct tgaacttccc tttgctgatt tatgacgaag acgtcaaaga aatggaaggc	420
caatccaagg aagagtatat tggatcattg agaagaaaaa gtagtggatt ttctcgaggt	480
gtatcaaaat acagaggcgt tgcaaggcat caccataatg ggagatggga agctagaatt	540
ggaaggggtgt ttggtataaa atatctatat cttggaacat acgccacgca agaagaagca	600
gcaatcgcct acgacatcgc ggcaatagag taccgtggac ttaacgccgt taccaatttc	660
gacgtcagcc gttatctaaa ccctgacgcc gccgattcta agccaattcg aaacgatccc	720
gaatcgtcgg atgataacaa atgtccgaaa tcagaggaaa taatcgaacc atctacatcg	780
cctgaagcta ttacaactcg ccggagcttt cccgacgata tccagacgta ttttgggtgt	840
caagactccg gcaagttagc gacggaggaa gacgttattt tcggtggttt gaattcattc	900
ataaatcctg gtttctataa cgagtttgat tatggacctt gatcgtattt ctttacaagt	960
tttgttttga ttatcta	977

<210> 261
 <211> 1082
 <212> DNA
 <213> Arabidopsis thaliana

<400> 261	
ctctcatitt cataattcca ttttcaggat tgtctctcaa tcttttattc ttctcattca	60
ccggtaatgg caaaagtctc tgggaggagc aagaaaacaa tcgttgacga tgaaatcagc	120
gataaaacag cgtctgcgtc tgagtctgctg tccattgcct taacatccaa acgcaaacgt	180
aagtcgccgc ctcgaaacgc tcctcttcaa cgcagctccc cttacagagg cgtcacaagg	240
catagatgga ctgggagata cgaagcgcct ttgtgggata agaacagctg gaacgataca	300
cagaccaaga aaggacgtca agtttatcta ggggcttacg acgaagaaga agcagcagca	360
cgtgcctacg acttagcagc attgaagtac tggggacgag acacactctt gaacttcctt	420
ttgccgagtt atgacgaaga cgtcaaagaa atggaaggcc aatccaagga agagtatatt	480
ggatcattga gaagaaaaag tagtggattt tctcgcggtg tatcaaaata cagaggcggt	540
gcaaggcatc accataatgg gagatgggaa gctagaattg gaaggggtgt tgccacgcaa	600
gaagaagcag caatcgccta cgacatcgcg gcaatagagt accgtggact taacgccggt	660
accaatttcg acgtcagccg ttatctaaac cctaacgccg ccgcgataa agccgattcc	720
gattctaagc ccattcgaag ccctagtcgc gagcccgaat cgtcggatga taacaaatct	780
ccgaaatcag aggaagtaat cgaaccatct acatcgccgg aagtgattcc aactcgccgg	840
agcttccccg acgatatcca gacgtatttt ggggtgtcaag attccggcaa gttagcgact	900
gaggaagacg taatattcga ttgtttcaat tcttatataa atcctggctt ctataacgag	960
tttgattatg gaccttaatc gtattttcta caagttttgt tttgattatc tacacaatac	1020
atcaaatatt atttcttagt gatgatgata acattctatc ttaaagctct caaaaatttc	1080

tt

1082

<210> 262
 <211> 1196
 <212> DNA
 <213> Arabidopsis lyrata

<400> 262
 attccatcaa tttctcatct ttgttgTTTT cctcttcttt atgtacatcg ccgtcgaagt 60
 ttctccggtg atggaagaga tcacacggaa gagcaagaaa acatccgttg agaatgaaac 120
 cggcgatgat cagtcagcaa catcagtagt cgtaaagct aaacgtaaac gtcgatcgca 180
 accacgagac gtcctccac aacgcagctc cgtccataga ggcgtcacca ggcacgatg 240
 gactggaagg tacgaagcac atttgtggga taagaatagt tggaacgaga ctcagagcaa 300
 gaaaggaaga caaggggcat atgacgagga agatgcagca gcacgtgcct acgacttagc 360
 agcattgaaa tattggggac gagacacat cttgaatttc cctctgtgta attatgaaga 420
 agacatcaag gaaatggaaa gccagtcaaa ggaagagtat attggatctt tgagaaggaa 480
 aagtagtggg ttctcacgtg gtgtatcaaa atacagaggc gttgcaaagc atcaccataa 540
 tgggagatgg gaagctcgaa ttggaagagt gtttggaat aaatatttat accttggaac 600
 ttacgctaca caagaagaag cggctatagc gtatgacatc gcagcaatcg agtaccgtgg 660
 acttaatgcc gttactaact tcgacatcag ccgttatatg aaactcccgg tgccggagaa 720
 ccctatcgac gccgcgaata atctcctgga gagtccgcat tctgattcca gccatttat 780
 aaaccctact cagcagctctg atttatcaca gagtcaatct tcctcagacg ataacgatga 840
 tcggaaaaca aagctcttaa agtcgtcacc tttaaagtca gaggaagtaa tcggaccatc 900
 gacgccaccg gagattgctc cgctcgtcg gagcttcccg gaagatatcc agacgtatct 960
 cgggtgtcaa aactccggca agttaacgac ggaggaagat gacgttatct tcggtgattt 1020
 agattctttc cttacgcctg atttctacag cgagttaaat gattgctaaa gtgttcttct 1080
 gatttgataa gttttgtttt tagttttttg gaatctcggg tgtgaaaatc agcattgaca 1140
 catatcggtt attctttctt gtgacaatct tattataata aagtttgaat cttttt 1196

<210> 263
 <211> 1114
 <212> DNA
 <213> Theellungiella halophila

<400> 263
 gatcgtatat gtcattaaa agtttgaaac cctcttttca ccattgtcgc tcctcattca 60
 ttcttctctc ttaccgggaa tggcaaaagt ctctcagcgg agcaagaaga caatcgtcaa 120
 cgacgaaatc agcgataaaa aagcgggtgc ggttgctct gtgtcttcgt ccgctttcct 180
 taaatcaaaa cgcaaacgga agctgccgcc caaaacgct ccacctcaac gcagctcctc 240
 ctacagaggc gtgacaaggc atagatggac ggggagatac gaagcgcatt tgtgggataa 300
 gaactgctgg aacgagacgc agaccaagaa aggacgtcaa gtttatcttg gggcatatga 360

pctau2012001598-seq1

cgaggaagaa gcagcagcac gtgcctacga cttagcggca ttgaagtact ggggacgaga	420
cacactcttg aacttccctc tacctactta tgaagaagac gtcaaagaaa tggaaggcca	480
ttccagggaa gagtacattg gatctttgag aaggaaaagt agtgggtttt ctcgtggtgt	540
atcaaaatac agaggcgttg caaggcatca ccataatggg agatgggaag ctcgaattgg	600
aaggggtgtt ggtaataaat atctatacct tggaacatac gcaacgcaag aggaagcagc	660
cagagcatat gacatcgcg caatagagta ccgtggactt aacgccgtta ccaacttcga	720
cgtcagccgt tatctaaatc taccggaatc gaaaaaccct agcgccgccc caaatcatct	780
cccggatgaa tccgattatt acgattctat gcccgttcga aacccaaacc acgagccaag	840
atcaccagac ggtcaaaact cgtcagagga taacgactat acgaaaacag aggagacact	900
cgacccggaa gctattccat ctcgccggag ctttcccgac gatatccaga cgtatttttg	960
gtgtcaggat tccggcaagt tagcgacgga ggaagacgtt attttcggtg gttttaattc	1020
tttcataaat cctggtttct ataacgactt tgattatgcg ccttgatagt gttttgttga	1080
aaagtttttt ttgtcccgaa aaaaaaaaaa aaaa	1114

<210> 264
 <211> 1220
 <212> DNA
 <213> Arabidopsis thaliana

<400> 264	
attccatcaa tttctcatcc ttgttgtttt cttcttcttc ttcttctctt tcttttatgt	60
tcatcgccgt cgaagtttct ccggtaatgg aagacatcac acggcagagc aaaaaaactt	120
cggttgagaa tgaaaccggc gatgatcagt cagcaacatc agtagtcctt aaagctaaac	180
gcaaacgccg atcgcaacca cgagacgtc caccccaacg tagctccgtc catagaggcg	240
tcacaaggca tcgatggact ggaaggtacg aagcacattt gtgggataag aatagttgga	300
acgaaactca gaccaagaaa ggaagacaag tatatttagg ggcatatgac gaggaagatg	360
cagcagcacg tgcctacgac ttagcagcat tgaaatatgt gggacgagac accatcttga	420
acttcccttt gtgtaattat gaagaagaca tcaaagaaat ggaaagccag tcaaaggaag	480
agtatatgtg atctttgaga agaaaaagta gtgggttttc acgaggtgta tcaaaataca	540
gaggcgttgc aaagcatcac cacaatggga gatgggaagc tcgaatcgga agagtgtttg	600
gcaataaata tttatacctt ggaacttacg cgacgcaaga agaagcagct atagcgtacg	660
atatcgcagc tatcgagtac cgtggactca acgccgttac taacttcgac atcagccggt	720
atctgaaact cccggtgccg gagaacccta tcgataccgc gaataatctc ctcgagagtc	780
cgcattctga tcttagccca tttataaaac ctaaccacga gtctgactta tcacagagtc	840
aatcttcgtc agaggacaac gatgatcgga aaacaaagct cttgaagtcg tcacctttag	900
tggcagagga ggtaatcgga ccatcgacgc cacctgagat tgctccgcct cgtcggagct	960
tcccggaaga tatccagacg tatttcgggt gtcaaaactc cggcaagtta acggcggagg	1020
aagatgatgt tatcttcggt gatttagatt ctttccttac gcctgatttc tacagcgagt	1080

pctau2012001598-seql

taaatgattg	ctaaagtgtt	gttcttctga	taagttttgt	tttttagttg	ttcagaatct	1140
cggttgtgaa	aatcaacatt	gacacatcga	ttattctttc	ttgtgacaat	cttatataat	1200
aaagtttgaa	tctttttaac					1220

<210>	265
<211>	900
<212>	DNA
<213>	Glycine max

400> 265							
atggcgaaga	aatcacagaa	aagcctcaaa	aacaacaata	acaacaacac	cactaggaag		60
cgcacacgca	aaagcgtccc	aagagactca	cctcctcaac	gaagctcaat	ctataggggt		120
gtcacaaggc	atcgatggac	tggctcgatac	gagggtcatt	tgtgggataa	gaattgctgg		180
aatgaatcac	agagcaagaa	aggaagacaa	gtatatctag	gggcatatga	tgatgaagaa		240
gccgcagcac	gtgcctatga	cttggcagca	ttgaagtact	ggggccaaga	taccattctc		300
aattttccat	tatcaaatta	tgaggagaag	ctcaaggaga	tggaaggcca	gtcaaaagaa		360
gaatatattg	gatccttgcg	aagaaaaagc	agtgggttct	ctcgaggagt	ttcaaaatac		420
agaggcgtag	caagacacca	ccataacgga	agatgggaag	ctcgaattgg	aagggtcttt		480
ggcaacaaat	atctatacct	tggaacttac	gctacacaag	aagaagcagc	cgcagcatat		540
gacatggcag	caatagaata	cagaggactt	aacgccgtta	caaactttga	cctgagccgt		600
tatatcaatt	ggccacgtcc	taaaacagaa	gaaaatcatc	aaaatactcc	tagcaaccaa		660
aatgtaaact	ctaattgctga	actcgagcta	ggttcggcta	gtgacgagat	cacagaagag		720
ggcgtagcac	gatcttcaga	gtcagagtca	aatccttctc	gtcgcacatt	ccctgaagac		780
attcaaacaa	tttttgaaaa	taatcaagat	tcaggcatct	acattgaaaa	tgatgatatc		840
atttttgggtg	acttgggctc	gtttgggtgca	ccaatttttc	atttcgaact	tgatgttttag		900

<210>	266
<211>	1186
<212>	DNA
<213>	Brachypodium distachyon

<400>	266						
cacaatggca	aagccccgca	agaactccgc	cgccgccaac	aacaacaaca	acgacaacag		60
caccaacgct	aacaacgccg	tgcgccaggc	tgccggcggcg	gatgtgcgcg	cgaagccgaa		120
gaagcgcacc	cggaagagcg	tgccgcggga	gtccccctcg	cagcgcagct	ccatctaccg		180
cggcgttaca	cggcaccggt	ggacggggcg	gttcgaggcg	cacctgtggg	acaagaacag		240
ctggaacgag	tcccagaaca	agaagggcaa	gcaagtttac	ctcggagcat	acgacgagga		300
ggaggcagcg	gcccgggCGT	atgacttggc	ggcattgaaa	tactggggcc	ccgacaccat		360
cctcaacttc	ccgctatctg	tatacgatga	tgagttgaaa	gaaatggagg	ggcagtcAAG		420
ggaagagtat	attgggtccc	tgaggaggaa	aagtagcggg	ttctcaagag	gagtctccaa		480
gtaccgcggt	gtcgccaggc	atcaccacaa	cgggagatgg	gaggccccga	tcggccgtgt		540

pctau2012001598-seq1

tttcggcaac	aagtatctct	acctcggcac	ctacgcgacg	caggaggagg	ccgcgatggc	600
gtacgacatg	gcggcgatcg	agtaccgcgg	cctgaacgcc	gtcaccaact	tcgacctcag	660
ccgctacatc	aagtggctcc	gccccggcgg	cggcgtcgac	tcgggtgccg	ccgccgccgc	720
ccgcaacccg	cacccgatgc	tggccggcct	ggcgaccag	gaggagttac	cagcgatcga	780
tcacctctc	gacggcatgg	cgttccagca	gcacggcctc	cattcttctt	cggcggcggc	840
ggcagcagcg	caggagtcc	ccctgccgcc	ggcgtcggc	cacgccccca	cgacgtcggc	900
gctgagcctc	ctgctgcagt	cgcccaagtt	caaggagatg	atcgagcgga	cgccgccgcg	960
ggagacgacg	acgacggcca	cgacgacgtc	gtcgtcttcg	tccccgcgtc	cggcggcgtc	1020
gccgcagtgc	agcttcccg	aggacatcca	gaccttcttc	gggtgcgacg	acggcgtggg	1080
cgtgggagtc	ggtgcagtgg	gatacaccga	cgtggacggc	ctcttcttcg	gggacctctc	1140
cgcgtacgcg	tcgtcgacgg	cgttccattt	cgagctcgac	ttgtga		1186

<210> 267
 <211> 1197
 <212> DNA
 <213> Sorghum bicolor

<400> 267	
atggcggaagc	ctcgcaagaa cagcgccgcc gccacaaca acaacagcag cagcaacggc 60
gccggcgacc	tgacgccgcg cgcaagccg aagcgcacgc ggaagagcgt gccccgggag 120
tccccacgc	agcgcagctc cgtctaccgc ggcgtcacac ggcaccggtg gacggggagg 180
ttcgaggctc	acctgtggga caagaacagc tggaacgagt ccagaacaa gaagggcaag 240
caagtttacc	tcggcgcgta cgacgacgag gaggcagcgg cgcgggcgta tgacttggcg 300
gcattgaagt	actggggccc cgacaccatc ctcaacttcc cggcgtctgc atatgaagga 360
gagatgaaag	gaatggaggg ccagtccagg gaagagtata ttggatcctt gaggaggaaa 420
agcagtgggt	tctccagagg tgtctccaaa taccgaggtg tcgcgagaca tcaccacaac 480
gggagatggg	aggcgaggat cggctcgtgtt ttcggaacaa agtatctcta cctcgggact 540
tacgcgacgc	aggaggaggc ggcgatggcg tacgacatgg cggcgatcga gtaccgtggc 600
ctgaacgccg	tcacaaactt cgacctcagc cgctacatca agtggctccg cccgggcgcc 660
ggcgggatgg	cagccgcggc ggcggcggcg cagaaccgc acccgatgct gggcggcctg 720
gcgcagcagc	tgctgctgcc gccgccggcg gacaccacca ccaccgacgg cgccggcgcc 780
gcggcgttcc	agcacgacca ccacggcgcc gaggcgttcc cgctgccgcc caggacgtcg 840
ctgggccaca	cgcccacgac gtcggcgctc agcctgctgc tgcagtcgcc caagttcaag 900
gagatgatcc	agcggaccga gagcggcacc accaccacta ccaccacgac atcgtcgttg 960
tcctcgtcgc	cgccgcccac gccgtccccg tccccgccgc ggcggtcgcc agcggcgacg 1020
cagccgccgg	tgcaggcggc ggccagggac gcctcgccgc accagcgcgg cttccccgag 1080
gacgtacaga	cattcttcgg ctgcgaggac accgcgggca tcgacgtgga agccctcttc 1140

pctau2012001598-seq1
 tttggcgacc tcgccgcgta cgcgacgccg gcgttccact tcgagatgga cttgtag 1197

<210> 268
 <211> 1736
 <212> DNA
 <213> Zea mays

<400> 268
 gccggccaat gagatggagc agagcaaaga atttcagaga gagagagaga gaggagagag 60
 aaaggccggc ctgtctctcc cgtctctctt ctttcctcct cctttgctca tttggcattt 120
 gccctgctct ccaaacaatca actaaatttg gccgccccat caccaatccc ttcccagtct 180
 ctccctccgc gtcgtcttta aaacggccac ctccaaaggc cccctgtagt tggccaccca 240
 cctctccctc ctcttttctt ctgtcgtctc tctctccgca cgcgcaatgg cgaggcctcg 300
 caagaacagc gccgccgccg ccaacaacaa caacagcaac accaccaacg ccggcaacgc 360
 cgccgtcgac ctcgcgccgc gcgtcaagcc gaagcgcacg cggaagagcg tgccccggga 420
 gtccccctcg cagcgcagct ccgtctaccg cggcgtcaca cggcaccggt ggacggggag 480
 gttcgaggct cacctgtggg acaagaacag ctggaacgag tcccagaaca agaagggcaa 540
 gcaagtttac ctcgccgcgt acgacgacga ggacgcagcg gcgcgcgcgt atgacttggc 600
 ggcatagaag tactggggcc ccgacaccat cctcaacttc cggcgtctg catatgaagc 660
 agagttgaaa gaaatggagg ggcagtccag ggaagagtat attgggtcct taaggaggaa 720
 aagcagtggg ttctccagag gtgtttccaa ataccgaggt gtcgcaaggc atcaccacaa 780
 cgggagatgg gaggcgagga ttggtcgcgt cttcggaac aagtatctct acctcgggac 840
 ttacgggacg caggaggagg cggcgatggc gtacgacatg gcggcgatcg agtaccgggg 900
 cctgaacgcc gtcactaact tcgacctcag ccgctacatc aagtggctcc gcccgggcgc 960
 cggcgccgcg cagaaccgc acccgatgct ggacggcctg gcgcagcagc tgctgctgtc 1020
 gccggagggc accatcgacg gcgccgcgtt ccaccagcag cagcacgacc accgccagca 1080
 gggcgccgcc gagtccccgc tgccgccag ggcgtcgtg ggccacacgc ccaccacgtc 1140
 ggcgtcggc ctgctgctgc agtcgtccaa gttcaaggag atgatccagc gggcgtcggc 1200
 ggccgagagc ggcaccacca ccgtcaccac cacgtcgtcg tcgtcgtcgc agccgccac 1260
 gccgacgcc acgccgtcgc catcgccacc tccgacgcc ccagtgcagc cggccaggga 1320
 cgctcgcca cagtgcagct tccccgagga catacagacc ttcttcggct gcgaggacgt 1380
 cgcgggcgtc ggcgccggcg tcgacgtgga cgccctcttc ttcggcgacc tcgccgcgta 1440
 cgcgtcgccg gcgttccact tcgagatgga cttgtagctc atcacacgcc attactcgtt 1500
 aatcactatc tgggtttttt tttttgtcga taggagaatg gagctctctc gccatttgag 1560
 aaccaagaaa atggtgagac acacaaaaca agagaggcag cagaacagac agacggcgcc 1620
 tagtctgttt tgagagattt atttacagtg gtaactagt acatttactg aacggttaag 1680
 attgtattat tcaaattttt ttcacattgg taacaaggga ttgtttcatt tgtttc 1736

pctau2012001598-seq1

<210> 269
 <211> 1089
 <212> DNA
 <213> Glycine max

<400> 269
 atggccaaaa aatcacagct gcgtacccag aaaaacaatg ctactaatga cgatattaat 60
 cttaacgcaa ccaacactgt aatcaccaag gtgaaacgaa caaggagaag tgtccctaga 120
 gactccccac ctcaacgcag ctcaatatac cgaggagtca ctaggcaccg atggactggc 180
 cgatacgaag ctcatTTgtg ggacaagcat tgctggaatg aatcacagaa taaaaaagga 240
 cgacaaggcg cttatgataa tgaagaggca gcagcacatg cttatgatct agcagcactg 300
 aaatactggg gtcaagatac cattcttaat tttccgttat caaactacct gaatgaactg 360
 aaagaaatgg aggggtcaatc acgggaagag tatatcggat cgctgaggag gaaaagcagt 420
 ggttttttcta gggggatttc taaatacaga ggtgttgcaa ggcatcatca taacggaagg 480
 tgggaggctc ggattggcaa agtttttggc aataaatatc tttatctcgg aacttacgct 540
 acccaagaag aagctgctac tgcctatgac ctggcagcta tagaataccg tggactcaat 600
 gctgtcacca atttcgatct cagccgttac attaagtggc ttaagcccaa caacaccaac 660
 agcaacaatg accagattag tattaatctc actaacataa ataataattg cactaacaac 720
 ttcaccccaa accctgatca agaacaagaa gttagtttct tccacaacca ggattcactc 780
 aataatacta ttgtagagga agccacgttg gtgccacatc agcctcgtcc agcaagtgcc 840
 acgttagcat tggagcttct acttcagtca tccaagttca aggaaatggg ggaaatgaca 900
 tccgtggcca atctttcaac acagatggaa tctgatcagt tgccacagtg cacatttcct 960
 gatcacattc agacatactt tgagtatgaa gattccaata aatatgagga aggggatgat 1020
 ctctgttca agttcagcga gttcagctcc attgtgccgt tttaccattg tgacgagttc 1080
 gagagttga 1089

<210> 270
 <211> 1298
 <212> DNA
 <213> Glycine max

<400> 270
 atggccaaaa aatcacagct gcgtacccag aaaaacaatg ttaccaccaa tgacgataat 60
 aatcttaacg taaccaacac tgtgaccacc aaggtgaaac gaacaaggag aagtgtccct 120
 agagactccc cacctcaacg cagctcaata taccgaggag tcactaggca ccgatggaca 180
 ggccgatacg aagctcattt gtgggacaaa cattgctgga atgaatcaca gaacaaaaaa 240
 gggcgacaag tctaccttgg cgcttatgac aatgaagagg cagcagcaca tgcttatgat 300
 ctagcagcac tgaaatactg ggggtcaagat accattctta attttccgtt atcaaactac 360
 ctgaacgaac tgaaagaaat ggaggggtcaa tcacgggagg agtatatcgg atcgctgagg 420
 aggaaaagca gtggtttttc tcggggaatt tctaaataca gaggtgttgc aaggcatcat 480
 cataacggaa ggtgggaggc tcggattggc aaagtttttg gcaataaata tctttacctc 540

pctau2012001598-seq1

ggaacttatg	ctaccaaga	agaagctgct	actgcctatg	acctggcagc	catagaatac	600
cgtggactca	atgctgtcac	caatttcgat	ctcagccgtt	acattaagtg	gcttaagcct	660
aacaacaaca	ccaacaacgt	tatcgacgac	cagattagta	ttaatctcac	taacataaac	720
aataataata	attgcactaa	cagcttcacc	ccaagtcctg	atcaagaaca	agaagctagc	780
ttcttccaca	acaagattc	actcaataat	actattgtag	aagaagtcac	gttggtgcc	840
catcagcctc	gtccagcgag	tgccacgtca	gcattggagc	ttctacttca	gtcatcaaag	900
ttcaaggaaa	tgatggagat	gacatctgtg	gccaatcttt	catcaacaca	gatggaatct	960
gagttgccac	agtgcacatt	tcctgatcac	attcagacgt	actttgagta	tgaagattcc	1020
aatagatatg	aggaaggaga	tgatctcatg	ttcaagttca	acgagttcag	ctccattgtg	1080
ccgttttacc	aatgtgacga	gttcgagagt	tgaagaagtc	aggtttatat	aatgcatgga	1140
aaaaagaaac	tctgatatgt	ttgtttat	gtttaatttg	ttgattatgt	taaagaccat	1200
attcataaat	ctttagctaa	ttaaggttta	agtttttaga	agagagatca	tgtcattcac	1260
aactattata	ataagtggac	ttgttttcaa	tttgtgaa			1298

<210> 271
 <211> 1160
 <212> DNA
 <213> *Medicago truncatula*

<400> 271						
acctttcaaa	ttaatgcacc	attaatacta	tatataaata	tacatctttg	ctagcttgtt	60
tcttcagtat	acacttcttt	ttttttccaa	tggcgaaaaa	atcgcagaag	caaatagaga	120
aagacgacaa	tgcaagcaat	gataatgata	atctgaatcc	aagtaacact	gtcaccacca	180
aagcaaaacg	gacgaggaaa	agtgtcccta	gaacttctcc	acctcaacgc	agctcaatat	240
acagaggagt	cactaggcac	cgttggacag	gccgatatga	agctcatttg	tgggacaaga	300
attgttggaa	tgagtctcag	aacaagaaag	ggagacaagg	tgcttatgat	aatgaggaga	360
cggcagcaca	tgcttatgat	ctagcagcat	tgaaatattg	gggtcaagat	accattatta	420
attttccgct	atcaaactat	cagaaagaac	taattgaaat	ggagagtcaa	tcaagggagg	480
agtatattgg	atccttaaga	aggaaaagca	gtgggttttc	acggggtgta	tctaaataca	540
gaggtgtagc	aagacatcat	cataacggaa	gggtggaggc	acggattggc	aaagtttttg	600
gcaataagta	tctttatctc	ggaacatatg	ctaccaaga	agaagcagct	acagcttatg	660
acatggcagc	tattgaatac	cgtggactta	atgctgttac	caatttcgac	ctaagccgtt	720
atattaagtg	gcttaagccc	aacaataaca	acaatgacga	caataacaag	tctaataatta	780
atctttgtga	cattaactct	aacagtagtg	ccaacgactc	aaactcta	gaagaattgg	840
aatttagcct	cgtggacaat	gaaatatcac	tcaataattc	tattgacgaa	gccacattgg	900
ttcaacctcg	accaactagt	gccacgtcag	ccctagagct	tttgcttcaa	tcattctaagt	960
ttaaggaaat	ggtggaaatg	gcatcaatga	cttccaatgt	gtcaacgacg	ttggagtctg	1020

pctau2012001598-seq1

atcaattgtc acaatgtgca tttcctgatg acattcaaac gtactttgaa tatgaaaatt	1080
tcaatgatac aatgtttgaa gacctcaact ccattatgcc tacgtttcat tatgactttg	1140
agggtgctga ggttttatag	1160

<210> 272
 <211> 1044
 <212> DNA
 <213> Glycine max

<400> 272	
atggcgaagc agcaaacaca caaaatcaat gcaagcacta acaataacat tagcacaact	60
aacactgtca ccgccaaggt gaaacgaaca aggagaagcg tccctagaga ttccccacct	120
caacgcagct caatatacag aggagtcact aggcaccgat ggactggccg atacgaagct	180
catttgtggg acaagaattg ctggaatgag tcacagaaca agaaaggacg acaaggtgct	240
tatgacgatg aagaggcggc agcacatgct tacgatctag cagcattgaa atactggggt	300
caagatacca ttcttaattt tccgttatcg acctaccaga atgaactgaa agaaatggag	360
ggatcaatcta gggaagagta tattggatca ctaaggagga aaagcagcgg tttttctcgt	420
ggagtttcta aatacagagg tgttgcaaga caccatcata acggaagatg ggaggctcgg	480
attggcaggg ttttcggcaa caaatatctt tacctcggaa cttacgcaac ccaagaagaa	540
gctgccacag catatgacat ggagctatt gaatatcgag gagtcaacgc tgtgaccaat	600
ttcgacctca gccgctatat caagtggctt aagcccaata ataacaacac cacagtcaac	660
tctaattcta ttgactctaa ccctaattgt gagacgaact ttacctcaa ctctaataca	720
caacaaggat tcaatttttt caacagacaa gagtattca ataatagaaga agcggcaatg	780
actcagctc ggccagctgt tgccacgtca gccctaggac ttttgcttca gtcattcaag	840
tttaaggaaa tgatggagat gacttctgct acagatttgt caacaccgcc gtcggagtct	900
gagttgccgt catgcacatt tcccgatgac attcaaacat attttgagtg tgaagattcc	960
catagatatg gagaagggga tgatattatg ttcagcgtgc tcaacggctt tgtgccccct	1020
atgttccatt gtgatgactt ttga	1044

<210> 273
 <211> 1056
 <212> DNA
 <213> Glycine max

<400> 273	
atggcgaagc agcaaacaca cgaaatcaat gcaagcacta acaataacat taacacaact	60
aaaactgtca ccaccaaggt gaaacgaaca aggagaagtg tccccagaaa ttccccacct	120
caacgcagct caatatacag aggagtcact agacaccgat ggacaggctc atacgaagct	180
catttgtggg acaagaattg ctggaatgag tcacagaaca agaaaggacg gcaaggtgct	240
tatgacgatg aagaggccgc agcacatgct tatgatctag cagcattgaa atactggggt	300
caagatacca ttcttaattt tccgttatcg acctaccaga atgaactgaa ggaaatggag	360

pctau2012001598-seq1

ggtcaatcta	gggaagagta	tattggatca	ctaaggagga	aaagcagcgg	tttttctcgt	420
ggagtttcta	aatacagagg	tggtgcaaga	caccatcata	acggaagatg	ggaagctcgg	480
attggcaggg	ttttcggcaa	caaatatctt	tacctcgga	cttatgcaac	ccaagaagaa	540
gctgccacag	catatgacat	ggcagctatt	gaatatcgag	gactcaacgc	tgtgaccaat	600
tttgacctca	gccgctatat	caagtggctt	aagcccaata	acaacaataa	caaagtcaat	660
tctaataatc	ttattgtctc	tatcccta	tgtgcgacca	actttacccc	aaactctaac	720
caacaacaag	gattcaat	tttcaacagt	caagaatcat	tcaataataa	tgaagaagcg	780
gcaatgactc	agcctcggcc	agctgctgcc	acgtcagccc	taggactttt	gcttcaatca	840
tccaagt	tttaaggaa	tgatgact	tctgcatag	atttgtcaac	accgccatca	900
gaatctgagt	tgccgccatg	cacatttccc	gatgacattc	aaacatattt	tgagtgtgaa	960
gattcccata	gatatggaga	aggggatgac	atcatgttta	gcgagctcaa	cggctttgtg	1020
ccccctatgt	tccattgtga	tgactttgag	gcttga			1056

<210> 274
 <211> 1062
 <212> DNA
 <213> Populus trichocarpa

<400> 274

atggcgaaat	tatcacagaa	gaacacaaaa	aacacagcaa	gtaacaataa	taacacgact	60
aatggcgtga	ccaaggtgaa	acgaacaagg	agaagtgttc	caagagactc	tcctcctcaa	120
cgcagctcca	tttatcgtgg	cgttactagg	caccggtgga	ccggccgata	tgaagctcat	180
ttgtgggata	agaattgctg	gaatgaatca	cagaacaaga	aaggacgaca	aggtgcctat	240
gatgatgaag	aagcggcagc	gcatgcttat	gacttagcag	cactgaagta	ctggggctct	300
gagacaattc	ttaattttcc	gttgtcgaca	tacaaaaatg	aactgaaaga	aatggaggggt	360
cagtcgagag	aagagtgc	tgatcattg	agaaggaaaa	gcagtggctt	ttctcgtgga	420
gtttcaaaat	accgaggagt	tgctaggcat	catcacacacg	ggaggtggga	agctcggatt	480
ggtagagtat	tcggcaacaa	atacctatat	ctcgggacat	atgctactca	agaagaagct	540
gccaccgctt	acgatatggc	ggctatagag	taccgtggcc	tgaatgctgt	taccaacttc	600
gatcttagcc	ggtacatcaa	atggcta	cccaatcaaa	ataacactga	taacaataat	660
ggtctggatc	tcctaacc	tatcataggc	actgataatt	caactcatcc	taaccctaac	720
caagagcttg	gaacaacctt	ccttcaaat	aaccagcaaa	cctaccaacc	tagtgaaact	780
acattaactc	aaccacggcc	agcaacgaat	ccttcatctg	ccctagggct	cttacttcaa	840
tcatcgaagt	tcaaggagat	gatggaaatg	acagcagtga	cagattgccc	accaacccca	900
ccgtcgggct	tagacccgac	accatgtagc	ttccttgaag	acgtgcagac	atattttgac	960
tgtctcgatt	caagtaacta	cggggatcaa	ggtgacgaca	tgatttttgg	cgacctcaac	1020
tcgtttgtgc	cacccatgtt	ccagtgcgat	ttcgagacat	aa		1062

pctau2012001598-seql

<210> 275
 <211> 1044
 <212> DNA
 <213> Vitis vinifera

<400> 275
 atggcgaagc tatcacagca aaaccataag aatagtgcaa acagcaatgc tactaatact 60
 actctctctg tgacaaaggt gaagcggact cgcaaaactg tccccagaga ctctcctcct 120
 cagcggagct cgatatatcg gggcgctcac aggcacgcgt ggactggccg atacgaggct 180
 catttggtggg ataagaactg ctggaatgaa tcacagaaca agaaaggaag acaaggtgcc 240
 taccatgatg aagaagccgc agcacatgct tatgacttgg cagctctcaa atactgggggt 300
 ccagaaacca ttctcaattt tccgctatca acatatgaaa aggagttgaa ggaaatggag 360
 ggtctgtcaa gagaagaata cattggatcc ttgagaagaa gaagcagcgg tttttctcga 420
 ggtgtttcaa aatatagagg ggttgcaaga catcatcaca atggcagatg ggaagcgcgg 480
 attggcagag tgttcggcaa caaatacctc tatcttggca catatgctac tcaggaggag 540
 gcagcgactg catatgatat ggctgccata gagtatcgcg gtcttaacgc ggttaccaac 600
 tttgatctca gtcgctacat caaatggtta aagcctaadc aaaataaccc atgcgaacag 660
 cccaataacc ctaaccttga ctctaatttg actccaaacc ctaatcatga ttttggaatc 720
 agcttcctga atcaccaca gacttctggt actgctgcct gcagtgaacc tccattgact 780
 cagaccaggc ctcccattgc ctcatcgccg cttggccttc tacttcagtc ctcaaagtcc 840
 aaagagatga tggagatgac aactgccgcc gaccacctat cgacaccgcc tgagtctgag 900
 ctgccacggt gcagcttccc tgacgacatt cagacgtact tcgagtgccg ggattccggc 960
 agttacgagg aaggagacga tggtatcttc agtgagctca actcgttcat acctcccatg 1020
 ttccaatgtg atttcagtgc ttaa 1044

<210> 276
 <211> 828
 <212> DNA
 <213> Populus trichocarpa

<400> 276
 atgggaaaaa catcaaagca aagcctaaag aactctgcaa acactagcat taatcccgca 60
 accaaggtga agcggactag aaagagtgtt cctcgagact ctctcctca acgcagctcc 120
 atctatagag gcgttacaag gcatcgatgg accggacgtt atgaggctca tttgtgggat 180
 aagaactgct ggaatgagtc acagaacaag aaaggaagac aaggggccta cgatgacgaa 240
 gaagcggctg ggcatgccta tgacttggca gcattaaagt actggggaca ggatactatc 300
 ctcaattttc ctctgtcaac ttacgaagaa gaattcaaag agatggaggg tcattcgaaa 360
 gaagaatata ttggatcact gaggagaaaa agtagtgggt tttcacgcgg tgtgtccaaa 420
 tatagaggtg tagccagaca ccatcataat ggaagatggg aagctcgaat tggcagagtt 480
 ttcggcaaca agtacctcta ccttgaaca tatgccactc aagaagaagc agcaacagca 540
 tatgacatgg cagctataga gtaccgtgga cttaatgctg tgacaaactt tgacctaaagc 600

pctau2012001598-seql

cgctactcct	caaagttcaa	ggaaatgttg	gagagaacat	cagcttctga	ttgccctttg	660
acaccacccg	aatccgaccg	cgatccgcct	cgtaggagct	tcccagatga	catccaaact	720
tacttcgatt	gccaaagactc	cagtagctac	actgatggtg	atgacattat	ttttggagac	780
ctacattcat	ttgcttcacc	gatttttcat	tgcgaattgg	atgggtag		828

<210> 277
 <211> 943
 <212> DNA
 <213> *Lupinus angustifolius*

<400> 277	
tggcaaataa	tgatttttat tttgactgat agtgacctgt tcgttgcaac aaattgatga 60
gcaatgcttt	tttataatgc caactttgta caaaaaagtt ggtgaggcac atctttggga 120
taaaagtaca	tggaaccaa atcagaataa gaaaggaaag caagtttact tgggggcata 180
tgatgatgag	gaagcagcag ctagagcata tgaccttgct gctctgaagt attgggggtcc 240
agggactctc	attaatttcc cagtgactga ttatacaaga gatcttgaag aaatgcagaa 300
tgtttcgaga	gaggaatacc ttgcatcttt acggcggaag agcagtgggt tttcaagagg 360
aatatcaaaa	tatcgtgcac tctccagtcg atgggagcca tcatatagtc gttttgctgg 420
atctgactac	ttcaacagta tgcattatgg cgcaggtgat gattcagccg cagaaaaaga 480
aaagagacat	ggtttcgctg gagatatttg cagcgaactt aaaacgctgg aacagaaagt 540
ccaacctact	gaaccatacc agatgccaga gttaggcagg tcccacaatg agaaaaaaca 600
tagaagttct	gccgtctctg ccttaagtat cttgtcgcaa tctgctgctt ataagagctt 660
gcaagagaaa	gcatcaaaga aacaggaaaa tagcatcgat aatgatgaga atgaaaacaa 720
aaatacagtc	aataagttgg attatggcaa agcagttgag aaatcatcaa tcatgacagt 780
gacagtgatc	gacttgacat tgcaatggga acagttgtgc cattgtgtct caaagaaatg 840
ttttatccat	cgatccgttc ttgtctgcac gcttttgact gcatacgata ctcgttgatc 900
catcatggta	agatcctggt tttgtgagac gtcctcatgg gcc 943

<210> 278
 <211> 1729
 <212> DNA
 <213> *Lupinus angustifolius*

<400> 278	
cgactctata	cgatacaaaa catactctga ttttctctct cttagcggtta gggtttcggt 60
ttcaataacc	atggcttctt cttcttccga tcccggtaag tccgaaattg gtggcggtgc 120
ggcggagact	tcggaagctg cggcggtggc ggtggcggtg acgaacgacc aatcgttatt 180
gtacagggga	ttgaagaaag cgaagaaaga gagaggttgt actgcgaaag aacgcatcag 240
caaaatgcct	ccttgtgccg ctggaaaacg cagctccata taccgtggag ttaccaggca 300
tagatggact	ggtcgctatg aggcacatct tagggataaa agtacatgga accaaaacca 360
gaataagaaa	ggaaagcaag tttacttggg ggcataatgat gatgaggaag cagcagctag 420

pctau2012001598-seql

```

agcatatgac cttgctgctc tgaagtattg ggggtccaggg actctcatta atttcccagt      480
gactgattat acaagagatc ttgaagaaat gcagaatgtt tcgagagagg aataccttgc      540
atctttacgg cggaagagca gtgggtttttc aagaggaata tcaaaatatc gtgcactctc      600
cagtcgatgg gagccatcat atagtcgttt tgctggatct gactacttca acagtatgca      660
ttatggcgca ggtgatgatt cagccgcaga aagtgagtat gcaagtgggt tctgtataga      720
aagaaagatt gatttaactg gtcacattaa atgggtgggga tctaataaaa gtcgtcaacc      780
tgatgctgga acgagattat cagaagaaaa gagacatggg ttcgctggag atatttgcag      840
cgaacctaaa acgctggaac agaaagtcca acctactgaa ccataccaga tgccagagtt      900
aggcaggtcc cacaatgaga aaaaacatag aagttctgcc gtctctgcct taagtatctt      960
gtcgcaatct gctgcttata agagcttgca agagaaagca tcaaagaaac aggaaaatag     1020
caccgataat gatgagaatg aaaacaaaaa cacagtcaat aagttggatc atggcaaagc     1080
agttgagaaa tcatcaaacc acgatgggtg cagtgatcgt gttgacattg aaattggaac     1140
aactggggca ttgtctcttc aaagaaatat ctaccactg acgccattct tgtctgcacc     1200
gcttttgaca gcatacaata ctgttgatcc atcattggta gatcctgttc tatggacatc     1260
tcttgtgcct atgctttctg ctggcctttc ttgtcctact caggttacaa agactgagac     1320
cagttcatct tacaccatct ttcagcctga gggatgatgt ggacattcga ttatctgaca     1380
gtagaaaagg caaacacctg aaagagctga actcctatat caaacagcaa cataagcatc     1440
tgcatctgaa ctctgaagaa ctcaagcagg tatgagagaa tgctcccatg gctttttttt     1500
tgcaagctta catgggtttt agaggccaga gaaaagaaag tagtattact gtgtttttaa     1560
tgataatata tatacaggaa tagcataaat cttttcaagc catacataca tgtgtatcag     1620
ttgatatcac caatagaaaa ttctgtacag agcaagatgt gaaattatgc aagaatatgt     1680
ttattaattt atttgcatag gagtttttga ggtcctaaaa aaaaaaaaaa     1729

```

<210> 279
 <211> 356
 <212> PRT
 <213> Arabidopsis thaliana

<400> 279

Met Asp Trp Glu Ile Arg Gly Ser Ser Leu Gly Gln Lys Leu Leu Glu
 1 5 10 15

Phe Asp Ser Glu Gln Glu Arg Gln Thr Arg Phe Arg Ala Tyr Asp Ser
 20 25 30

Glu Glu Ala Ala Ala His Thr Tyr Asp Leu Ala Ala Leu Lys Tyr Trp
 35 40 45

Gly Pro Asp Thr Ile Leu Asn Phe Pro Ala Glu Thr Tyr Thr Lys Glu
 50 55 60

pctau2012001598-seql

```

Leu Glu Glu Met Gln Arg Val Thr Lys Glu Glu Tyr Leu Ala Ser Leu
65      70      75      80

Arg Arg Gln Ser Ser Gly Phe Ser Arg Gly Val Ser Lys Tyr Arg Gly
85      90      95

Val Ala Arg His His His Asn Gly Arg Trp Glu Ala Arg Ile Gly Arg
100     105     110

Val Phe Gly Asn Lys Tyr Leu Tyr Leu Gly Thr Tyr Asn Thr Gln Glu
115     120     125

Glu Ala Ala Ala Tyr Asp Met Ala Ala Ile Glu Tyr Arg Gly Ala
130     135     140

Asn Ala Val Thr Asn Phe Asp Ile Ser Asn Tyr Ile Asp Arg Leu Lys
145     150     155     160

Lys Lys Gly Val Phe Pro Phe Pro Val Asn Gln Ala Asn His Gln Glu
165     170     175

Gly Ile Leu Val Glu Ala Lys Gln Glu Val Glu Thr Arg Glu Ala Lys
180     185     190

Glu Glu Pro Arg Glu Glu Val Lys Gln Gln Tyr Val Glu Glu Pro Pro
195     200     205

Gln Glu Glu Glu Lys Glu Glu Glu Lys Ala Glu Gln Gln Glu Ala
210     215     220

Glu Ile Val Gly Tyr Ser Glu Glu Ala Ala Val Val Asn Cys Cys Ile
225     230     235     240

Asp Ser Ser Thr Ile Met Glu Met Asp Arg Cys Gly Asp Asn Asn Glu
245     250     255

Leu Ala Trp Asn Phe Cys Met Met Asp Thr Gly Phe Ser Pro Phe Leu
260     265     270

Thr Asp Gln Asn Leu Ala Asn Glu Asn Pro Ile Glu Tyr Pro Glu Leu
275     280     285

Phe Asn Glu Leu Ala Phe Glu Asp Asn Ile Asp Phe Met Phe Asp Asp
290     295     300

Gly Lys His Glu Cys Leu Asn Leu Glu Asn Leu Asp Cys Cys Val Val
305     310     315     320

Gly Arg Glu Ser Pro Pro Ser Ser Ser Ser Pro Leu Ser Cys Leu Ser
325     330     335

```

pctau2012001598-seql

Thr Asp Ser Ala Ser Ser Thr Thr Thr Thr Thr Ser Val Ser Cys
340 345 350

Asn Tyr Leu Val
355

<210> 280
<211> 430
<212> PRT
<213> Arabidopsis thaliana
<400> 280

Met Lys Lys Arg Leu Thr Thr Ser Thr Cys Ser Ser Ser Pro Ser Ser
1 5 10 15

Ser Val Ser Ser Ser Thr Thr Thr Ser Ser Pro Ile Gln Ser Glu Ala
20 25 30

Pro Arg Pro Lys Arg Ala Lys Arg Ala Lys Lys Ser Ser Pro Ser Gly
35 40 45

Asp Lys Ser His Asn Pro Thr Ser Pro Ala Ser Thr Arg Arg Ser Ser
50 55 60

Ile Tyr Arg Gly Val Thr Arg His Arg Trp Thr Gly Arg Phe Glu Ala
65 70 75 80

His Leu Trp Asp Lys Ser Ser Trp Asn Ser Ile Gln Asn Lys Lys Gly
85 90 95

Lys Gln Val Tyr Leu Gly Ala Tyr Asp Ser Glu Glu Ala Ala Ala His
100 105 110

Thr Tyr Asp Leu Ala Ala Leu Lys Tyr Trp Gly Pro Asp Thr Ile Leu
115 120 125

Asn Phe Pro Ala Glu Thr Tyr Thr Lys Glu Leu Glu Glu Met Gln Arg
130 135 140

Val Thr Lys Glu Glu Tyr Leu Ala Ser Leu Arg Arg Gln Ser Ser Gly
145 150 155 160

Phe Ser Arg Gly Val Ser Lys Tyr Arg Gly Val Ala Arg His His His
165 170 175

Asn Gly Arg Trp Glu Ala Arg Ile Gly Arg Val Phe Gly Asn Lys Tyr
180 185 190

Leu Tyr Leu Gly Thr Tyr Asn Thr Gln Glu Glu Ala Ala Ala Ala Tyr
195 200 205

pctau2012001598-seql

Asp Met Ala Ala Ile Glu Tyr Arg Gly Ala Asn Ala Val Thr Asn Phe
210 215 220

Asp Ile Ser Asn Tyr Ile Asp Arg Leu Lys Lys Lys Gly Val Phe Pro
225 230 235 240

Phe Pro Val Asn Gln Ala Asn His Gln Glu Gly Ile Leu Val Glu Ala
245 250 255

Lys Gln Glu Val Glu Thr Arg Glu Ala Lys Glu Glu Pro Arg Glu Glu
260 265 270

Val Lys Gln Gln Tyr Val Glu Glu Pro Pro Gln Glu Glu Glu Lys
275 280 285

Glu Glu Glu Lys Ala Glu Gln Gln Glu Ala Glu Ile Val Gly Tyr Ser
290 295 300

Glu Glu Ala Ala Val Val Asn Cys Cys Ile Asp Ser Ser Thr Ile Met
305 310 315 320

Glu Met Asp Arg Cys Gly Asp Asn Asn Glu Leu Ala Trp Asn Phe Cys
325 330 335

Met Met Asp Thr Gly Phe Ser Pro Phe Leu Thr Asp Gln Asn Leu Ala
340 345 350

Asn Glu Asn Pro Ile Glu Tyr Pro Glu Leu Phe Asn Glu Leu Ala Phe
355 360 365

Glu Asp Asn Ile Asp Phe Met Phe Asp Asp Gly Lys His Glu Cys Leu
370 375 380

Asn Leu Glu Asn Leu Asp Cys Cys Val Val Gly Arg Glu Ser Pro Pro
385 390 395 400

Ser Ser Ser Ser Pro Leu Ser Cys Leu Ser Thr Asp Ser Ala Ser Ser
405 410 415

Thr Thr Thr Thr Thr Thr Ser Val Ser Cys Asn Tyr Leu Val
420 425 430

<210> 281
<211> 430
<212> PRT
<213> Arabidopsis lyrata

<400> 281

Met Lys Arg Arg Leu Thr Thr Ser Thr Ser Ser Ser Ser Pro Ser Ser
1 5 10 15

Ser Val Ser Ser Ser Thr Thr Thr Ser Ser Pro Ile Gln Ser Glu Ala
Page 334

Pro Arg Pro Lys Arg Ala Lys Arg Ala Lys Lys Ser Ser Pro Ser Gly
 35 40 45
 Asp Lys Ser His Asn Pro Thr Ser Pro Ala Ser Thr Arg Arg Ser Ser
 50 55 60
 Ile Tyr Arg Gly Val Thr Arg His Arg Trp Thr Gly Arg Phe Glu Ala
 65 70 75 80
 His Leu Trp Asp Lys Ser Ser Trp Asn Ser Ile Gln Asn Lys Lys Gly
 85 90 95
 Lys Gln Gly Ala Tyr Asp Ser Glu Glu Ala Ala Ala His Thr Tyr Asp
 100 105 110
 Leu Ala Ala Leu Lys Tyr Trp Gly Pro Asp Thr Ile Leu Asn Phe Pro
 115 120 125
 Ala Glu Thr Tyr Thr Lys Glu Leu Glu Glu Met Gln Arg Val Thr Lys
 130 135 140
 Glu Glu Tyr Leu Ala Ser Leu Arg Arg Gln Ser Ser Gly Phe Ser Arg
 145 150 155 160
 Gly Val Ser Lys Tyr Arg Gly Val Ala Arg His His His Asn Gly Arg
 165 170 175
 Trp Glu Ala Arg Ile Gly Arg Val Phe Gly Asn Lys Tyr Leu Tyr Leu
 180 185 190
 Gly Thr Tyr Asn Thr Gln Glu Glu Ala Ala Ala Ala Tyr Asp Met Ala
 195 200 205
 Ala Ile Glu Tyr Arg Gly Ala Asn Ala Val Thr Asn Phe Asp Ile Ser
 210 215 220
 Asn Tyr Ile Asp Arg Leu Lys Lys Lys Gly Val Phe Pro Phe Pro Val
 225 230 235 240
 Asn Gln Pro Asn His Gln Glu Ala Ile Leu Val Glu Ala Lys Gln Glu
 245 250 255
 Ile Glu Thr Arg Glu Ala Lys Glu Glu Pro Arg Glu Glu Val Lys Gln
 260 265 270
 Gln Tyr Val Glu Glu Pro Pro Gln Glu Glu Lys Glu Glu Glu Lys Ala
 275 280 285
 Glu Gln Gln Glu Ala Glu Phe Val Gly Tyr Lys Asp Glu Gly Ala Val
 Page 335

pctau2012001598-seql

290

295

300

Val Asn Cys Cys Ile Asp Ser Ser Ala Ile Met Glu Met Asn Arg Cys
 305 310 315 320

Gly Asp Asn Asn Glu Leu Ala Trp Asn Phe Cys Met Met Asp Ser Gly
 325 330 335

Phe Ala Pro Phe Leu Thr Asp Gln Asn Leu Ser Asn Glu Asn Pro Ile
 340 345 350

Glu Tyr Pro Glu Leu Phe Asn Glu Leu Ala Phe Glu Asp Asn Ile Asp
 355 360 365

Phe Met Phe Asp Glu Ala Lys Asn Asp Cys Leu Ser Leu Glu Asn Leu
 370 375 380

Asp Cys Cys Val Val Gly Arg Glu Ser Pro Thr Ser Ser Ser Ser Pro
 385 390 395 400

Leu Ser Cys Phe Ser Thr Asp Ser Ala Ser Ser Thr Thr Thr Thr Thr
 405 410 415

Ser Val Ser Cys Asn Tyr Leu Gly Leu Phe Val Gly Ser Glu
 420 425 430

<210> 282

<211> 413

<212> PRT

<213> Brassica napus

<400> 282

Met Lys Arg Pro Leu Thr Thr Ser Pro Ser Ser Ser Ser Thr Ser
 1 5 10 15

Ser Ser Ala Cys Ile Leu Pro Thr Gln Ser Glu Thr Pro Arg Pro Lys
 20 25 30

Arg Ala Lys Arg Ala Lys Lys Ser Ser Leu Arg Ser Asp Val Lys Pro
 35 40 45

Gln Asn Pro Thr Ser Pro Ala Ser Thr Arg Arg Ser Ser Ile Tyr Arg
 50 55 60

Gly Val Thr Arg His Arg Trp Thr Gly Arg Tyr Glu Ala His Leu Trp
 65 70 75 80

Asp Lys Ser Ser Trp Asn Ser Ile Gln Asn Lys Lys Gly Lys Gln Val
 85 90 95

Tyr Leu Gly Ala Tyr Asp Ser Glu Glu Ala Ala Ala His Thr Tyr Asp
 100 105 110

pctau2012001598-seql

Leu Ala Ala Leu Lys Tyr Trp Gly Pro Asn Thr Ile Leu Asn Phe Pro
 115 120 125
 Val Glu Thr Tyr Thr Lys Glu Leu Glu Glu Met Gln Arg Cys Thr Lys
 130 135 140
 Glu Glu Tyr Leu Ala Ser Leu Arg Arg Gln Ser Ser Gly Phe Ser Arg
 145 150 155 160
 Gly Val Ser Lys Tyr Arg Gly Val Ala Arg His His His Asn Gly Arg
 165 170 175
 Trp Glu Ala Arg Ile Gly Arg Val Phe Gly Asn Lys Tyr Leu Tyr Leu
 180 185 190
 Gly Thr Tyr Asn Thr Gln Glu Glu Ala Ala Ala Ala Tyr Asp Met Ala
 195 200 205
 Ala Ile Glu Tyr Arg Gly Ala Asn Ala Val Thr Asn Phe Asp Ile Gly
 210 215 220
 Asn Tyr Ile Asp Arg Leu Lys Lys Lys Gly Val Phe Pro Phe Pro Val
 225 230 235 240
 Ser Gln Ala Asn His Gln Glu Ala Val Leu Ala Glu Thr Lys Gln Glu
 245 250 255
 Val Glu Ala Lys Glu Glu Pro Thr Glu Glu Val Lys Gln Cys Val Glu
 260 265 270
 Lys Glu Glu Ala Lys Glu Glu Lys Thr Glu Lys Lys Gln Gln Gln Glu
 275 280 285
 Val Glu Glu Ala Val Ile Thr Cys Cys Ile Asp Ser Ser Glu Ser Asn
 290 295 300
 Glu Leu Ala Trp Asp Phe Cys Met Met Asp Ser Gly Phe Ala Pro Phe
 305 310 315 320
 Leu Thr Asp Ser Asn Leu Ser Ser Glu Asn Pro Ile Glu Tyr Pro Glu
 325 330 335
 Leu Phe Asn Glu Met Gly Phe Glu Asp Asn Ile Asp Phe Met Phe Glu
 340 345 350
 Glu Gly Lys Gln Asp Cys Leu Ser Leu Glu Asn Leu Asp Cys Cys Asp
 355 360 365
 Gly Val Val Val Val Gly Arg Glu Ser Pro Thr Ser Leu Ser Ser Ser
 370 375 380

pctau2012001598-seql

Pro Leu Ser Cys Leu Ser Thr Asp Ser Ala Ser Ser Thr Thr Thr Thr
385 390 395 400

Ala Thr Thr Val Thr Ser Val Ser Trp Asn Tyr Ser Val
405 410

<210> 283
<211> 415
<212> PRT
<213> Brassica napus
<400> 283

Met Lys Arg Pro Leu Thr Thr Ser Pro Ser Thr Ser Ser Ser Thr Ser
1 5 10 15

Ser Ser Ala Cys Ile Leu Pro Thr Gln Pro Glu Thr Pro Arg Pro Lys
20 25 30

Arg Ala Lys Arg Ala Lys Lys Ser Ser Ile Pro Thr Asp Val Lys Pro
35 40 45

Gln Asn Pro Thr Ser Pro Ala Ser Thr Arg Arg Ser Ser Ile Tyr Arg
50 55 60

Gly Val Thr Arg His Arg Trp Thr Gly Arg Tyr Glu Ala His Leu Trp
65 70 75 80

Asp Lys Ser Ser Trp Asn Ser Ile Gln Asn Lys Lys Gly Lys Gln Val
85 90 95

Tyr Leu Gly Ala Tyr Asp Ser Glu Glu Ala Ala Ala His Thr Tyr Asp
100 105 110

Leu Ala Ala Leu Lys Tyr Trp Gly Pro Asp Thr Ile Leu Asn Phe Pro
115 120 125

Ala Glu Thr Tyr Thr Lys Glu Leu Glu Glu Met Gln Arg Cys Thr Lys
130 135 140

Glu Glu Tyr Leu Ala Ser Leu Arg Arg Gln Ser Ser Gly Phe Ser Arg
145 150 155 160

Gly Val Ser Lys Tyr Arg Gly Val Ala Arg His His His Asn Gly Arg
165 170 175

Trp Glu Ala Arg Ile Gly Arg Val Phe Gly Asn Lys Tyr Leu Tyr Leu
180 185 190

Gly Thr Tyr Asn Thr Gln Glu Glu Ala Ala Ala Ala Tyr Asp Met Ala
195 200 205

pctau2012001598-seql

Ala Ile Glu Tyr Arg Gly Ala Asn Ala Val Thr Asn Phe Asp Ile Ser
210 215 220

Asn Tyr Ile Asp Arg Leu Lys Lys Lys Gly Val Phe Pro Phe Pro Val
225 230 235 240

Ser Gln Ala Asn His Gln Glu Ala Val Leu Ala Glu Ala Lys Gln Glu
245 250 255

Val Glu Ala Lys Glu Glu Pro Thr Glu Glu Val Lys Gln Cys Val Glu
260 265 270

Lys Glu Glu Pro Gln Glu Ala Lys Glu Glu Lys Thr Glu Lys Lys Gln
275 280 285

Gln Gln Gln Glu Val Glu Glu Ala Val Val Thr Cys Cys Ile Asp Ser
290 295 300

Ser Glu Ser Asn Glu Leu Ala Trp Asp Phe Cys Met Met Asp Ser Gly
305 310 315 320

Phe Ala Pro Phe Leu Thr Asp Ser Asn Leu Ser Ser Glu Asn Pro Ile
325 330 335

Glu Tyr Pro Glu Leu Phe Asn Glu Met Gly Phe Glu Asp Asn Ile Asp
340 345 350

Phe Met Phe Glu Glu Gly Lys Gln Asp Cys Leu Ser Leu Glu Asn Leu
355 360 365

Asp Cys Cys Asp Gly Val Val Val Val Gly Arg Glu Ser Pro Thr Ser
370 375 380

Leu Ser Ser Ser Pro Leu Ser Cys Leu Ser Thr Asp Ser Ala Ser Ser
385 390 395 400

Thr Thr Thr Thr Thr Ile Thr Ser Val Ser Cys Asn Tyr Ser Val
405 410 415

<210> 284
<211> 285
<212> PRT
<213> Glycine max

<400> 284

Met Lys Arg Ser Pro Ala Ser Ser Cys Ser Ser Ser Thr Ser Ser Val
1 5 10 15

Gly Phe Glu Val His His Pro Ile Glu Lys Arg Arg Pro Lys His Pro
20 25 30

pctau2012001598-seql

Arg Arg Asn₃₅ Asn Leu Lys Ser Gln₄₀ Lys Cys Lys Gln₄₅ Asn Gln Thr Thr

Thr Gly₅₀ Gly Arg Arg Ser₅₅ Ser Ile Tyr Arg Gly Val₆₀ Thr Arg His Arg

Trp₆₅ Thr Gly Arg Phe₇₀ Glu Ala His Leu Trp₇₅ Asp Lys Ser Ser Trp Asn₈₀

Asn Ile Gln Ser₈₅ Lys Lys Gly Lys Gln₉₀ Val Tyr Leu Gly Ala Tyr₉₅ Asp

Thr Glu Glu Ser₁₀₀ Ala Ala Arg Thr Tyr₁₀₅ Asp Leu Ala Ala Leu₁₁₀ Lys Tyr

Trp Gly Lys₁₁₅ Asp Ala Thr Leu Asn₁₂₀ Phe Pro Ile Glu Thr₁₂₅ Tyr Thr Lys

Asp Leu₁₃₀ Glu Glu Met Asp₁₃₅ Lys Val Ser Arg Glu₁₄₀ Glu Tyr Leu Ala Ser

Leu Arg Arg Gln Ser₁₅₀ Ser Gly Phe Ser Arg Gly₁₅₅ Ile Ser Lys Tyr Arg₁₆₀

Gly Val Ala Arg His₁₆₅ His His Asn Gly Arg₁₇₀ Trp Glu Ala Arg Ile Gly₁₇₅

Arg Val Cys Gly₁₈₀ Asn Lys Tyr Leu Tyr₁₈₅ Leu Gly Thr Tyr Lys₁₉₀ Thr Gln

Glu Glu Ala₁₉₅ Ala Val Ala Tyr Asp₂₀₀ Met Ala Ala Ile Glu₂₀₅ Tyr Arg Gly

Val Asn₂₁₀ Ala Val Thr Asn Phe₂₁₅ Asp Ile Ser Asn Tyr₂₂₀ Met Asp Lys Ile

Lys₂₂₅ Lys Lys Asn Asp Gln₂₃₀ Thr Leu Gln Gln₂₃₅ Gln Thr Glu Val Gln₂₄₀

Thr Glu Thr Val₂₄₅ Pro Asn Ser Ser Asp₂₅₀ Ser Glu Glu Ala Glu₂₅₅ Val Glu

Gln Gln His Thr₂₆₀ Thr Thr Ile Thr Thr₂₆₅ Pro Pro Pro Ser Glu₂₇₀ Asn Leu

His Met Leu₂₇₅ Pro Gln Glu His Gln₂₈₀ Val Gly Gly Trp Val₂₈₅

<210> 285
 <211> 417
 <212> PRT
 <213> Jatropha curcas

pctau2012001598-seql

<400> 285

```

Met  Lys  Arg  Ser  Ser  Ala  Ser  Ser  Cys  Ser  Ser  Ser  Ser  Ser  Ser  Ser
1      5      10     15

Ser  Ser  Pro  Ser  Ser  Ser  Ser  Ser  Ser  Ala  Cys  Ser  Ala  Ser  Ser  Ser
20     25     30

Cys  Leu  Asp  Ser  Val  Ser  Pro  Pro  Asn  His  His  Gln  Leu  Arg  Ser  Glu
35     40     45

Lys  Ser  Lys  Ser  Lys  Arg  Ile  Arg  Lys  Ile  Gln  Thr  Lys  Gln  Asp  Lys
50     55     60

Cys  Gln  Thr  Thr  Ala  Thr  Thr  Thr  Ser  Pro  Ser  Gly  Gly  Gly  Arg  Arg
65     70     75     80

Ser  Ser  Ile  Tyr  Arg  Gly  Val  Thr  Arg  His  Arg  Trp  Thr  Gly  Arg  Phe
85     90     95

Glu  Ala  His  Leu  Trp  Asp  Lys  Ser  Ser  Trp  Asn  Asn  Ile  Gln  Asn  Lys
100    105    110

Lys  Gly  Arg  Gln  Val  Tyr  Leu  Gly  Ala  Tyr  Asp  Asn  Glu  Glu  Ala  Ala
115    120    125

Ala  His  Thr  Tyr  Asp  Leu  Ala  Ala  Leu  Lys  Tyr  Trp  Gly  Gln  Asp  Thr
130    135    140

Thr  Leu  Asn  Phe  Pro  Ile  Glu  Thr  Tyr  Ser  Lys  Glu  Leu  Glu  Glu  Met
145    150    155    160

Gln  Lys  Met  Ser  Lys  Glu  Glu  Tyr  Leu  Ala  Ser  Leu  Arg  Arg  Arg  Ser
165    170    175

Ser  Gly  Phe  Ser  Arg  Gly  Val  Ser  Lys  Tyr  Arg  Gly  Val  Ala  Arg  His
180    185    190

His  His  Asn  Gly  Arg  Trp  Glu  Ala  Arg  Ile  Gly  Arg  Val  Phe  Gly  Asn
195    200    205

Lys  Tyr  Leu  Tyr  Leu  Gly  Thr  Tyr  Asn  Thr  Gln  Glu  Glu  Ala  Ala  Ala
210    215    220

Ala  Tyr  Asp  Met  Ala  Ala  Ile  Glu  Tyr  Arg  Gly  Ala  Asn  Ala  Val  Thr
225    230    235    240

Asn  Phe  Asp  Val  Ser  His  Tyr  Ile  Asp  Arg  Leu  Lys  Lys  Lys  Gly  Ile
245    250    255

Pro  Leu  Asp  Lys  Ile  Leu  Pro  Glu  Thr  Leu  Ser  Lys  Gly  Ser  Lys  Glu

```

pctau2012001598-seql

260

265

270

Ser Glu Glu Ile Glu Arg Thr Ser Pro Leu Pro Leu Pro Ser Pro Pro
 275 280 285

Ser Pro Ser Ile Thr Pro Leu His Glu Glu Ile Val Ser Pro Gln Leu
 290 295 300

Leu Glu Thr Glu Cys Pro Gln His Pro Pro Cys Met Asp Thr Cys Thr
 305 310 315 320

Met Ile Val Met Asp Pro Ile Glu Glu His Glu Leu Thr Trp Ser Phe
 325 330 335

Cys Leu Asp Ser Gly Leu Val Pro Leu Pro Val Pro Asp Leu Pro Leu
 340 345 350

Ala Asn Gly Cys Glu Leu Pro Asp Leu Leu Asp Asp Thr Gly Phe Glu
 355 360 365

Asp Asn Ile Asp Leu Ile Phe Asp Ala Cys Cys Phe Gly Asn Asp Ala
 370 375 380

Asn Pro Ala Asp Glu Asn Gly Lys Glu Arg Leu Ser Ser Ala Ser Thr
 385 390 395 400

Ser Pro Ser Cys Ser Thr Thr Leu Thr Ser Val Ser Cys Asn Tyr Ser
 405 410 415

Val

<210> 286
 <211> 443
 <212> PRT
 <213> Ricinus communis

<400> 286

Met Lys Arg Ser Pro Thr Ser Pro Cys Ser Ser Ser Ser Ser Ser Ser
 1 5 10 15

Tyr Ser Ser Ser Ser Ala Ser Ser Ser Cys Val Gly Pro Asp Asp Thr
 20 25 30

Pro Val Ala Pro Gly Ser His His His Asp His His Gln Leu Arg
 35 40 45

Ser Gln Lys Ser Ser Lys Arg Ile Arg Lys Val Lys Lys Lys Gln Gln
 50 55 60

Asn His Asn Ile Asp Gln Asn Asn Thr Asn Thr Thr Ile Thr Ala Pro
 65 70 75 80

pctau2012001598-seql

Thr Ser Ala Arg Arg Ser Ser Ile Tyr Arg Gly Val Thr Arg His Arg
 85 90 95
 Trp Thr Gly Arg Phe Glu Ala His Leu Trp Asp Lys Ser Ser Trp Asn
 100 105 110
 Asn Ile Gln Asn Lys Lys Gly Arg Gln Gly Ala Tyr Asp Asn Glu Glu
 115 120 125
 Ala Ala Ala His Thr Tyr Asp Leu Ala Ala Leu Lys Tyr Trp Gly Pro
 130 135 140
 Glu Thr Thr Leu Asn Phe Pro Ile Glu Thr Tyr Pro Lys Glu Leu Glu
 145 150 155 160
 Glu Met Gln Lys Met Ser Lys Glu Glu Tyr Leu Ala Ser Leu Arg Arg
 165 170 175
 Gln Ser Ser Gly Phe Ser Arg Gly Val Ser Lys Tyr Arg Gly Val Ala
 180 185 190
 Arg His His His Asn Gly Arg Trp Glu Ala Arg Ile Gly Arg Val Phe
 195 200 205
 Gly Asn Lys Tyr Leu Tyr Leu Gly Thr Tyr Asn Thr Gln Glu Glu Ala
 210 215 220
 Ala Ala Ala Tyr Asp Met Ala Ala Ile Glu Tyr Arg Gly Ala Asn Ala
 225 230 235 240
 Val Thr Asn Phe Asp Ile Ser Asn Tyr Ile Asp Arg Leu Lys Lys Lys
 245 250 255
 Gly Ile Leu Leu Asp Gln Ile Leu Pro Asp Gln Pro Leu Arg Lys Cys
 260 265 270
 Ser Ser Glu Ser Glu Glu Ala Glu Ala Glu Ala Glu Val Glu Arg Leu
 275 280 285
 Pro Ser Leu Pro Ser Ser Ile Leu Pro Gln Glu Gln Asp Thr Ile Ser
 290 295 300
 Pro Gln Leu Gln Cys Thr Gln Leu Leu Pro Ser Met Asp Ser Cys Thr
 305 310 315 320
 Met Ile Asn Met Asp Pro Ile Glu Asp Asn Glu Leu Thr Trp Ser Phe
 325 330 335
 Cys Leu Asp Ser Gly Leu Thr Leu Phe Ser Val Pro Glu Leu Pro Leu
 340 345 350

pctau2012001598-seql

Glu Asn Ala Cys Glu Leu Pro Asp Leu Phe Asp Asp Thr Gly Phe Glu
355 360 365

Asp Asn Ile Asp Leu Ile Phe Asp Gly Cys Cys Phe Gly Asn Asp Asp
370 375 380

Asp Gly Gly Gly Gly Ala Asn His Gln Glu Phe Met Val Glu Ser Arg
385 390 395 400

Gly Cys Arg Val Gly Glu Val Gly Ile Ser Gly Ser Met Glu Glu Glu
405 410 415

Asn Gly Lys Glu Met Cys Cys Ser Ser Ser Pro Ser Cys Ser Thr
420 425 430

Thr Thr Ser Val Ser Cys Cys Asn Tyr Ser Val
435 440

<210> 287
<211> 402
<212> PRT
<213> Populus trichocarpa
<400> 287

Met Lys Arg Ser Ser Ser Cys Ser Ser Ser Ser Ser Ser Ser Ser
1 5 10 15

Ser Cys Val Ala Ser Glu Ser Ile His Lys Pro Lys Ala Lys Arg Ile
20 25 30

Arg Lys Asn Gln Lys Ser Asn Gln Gly Lys Ser Gln Asn Ala Ala Ala
35 40 45

Ala Ala Ala Asn Asn Ser His Asn Ser Gly Lys Arg Ser Ser Ile Tyr
50 55 60

Arg Gly Val Thr Arg His Arg Trp Thr Gly Arg Phe Glu Ala His Leu
65 70 75 80

Trp Asp Lys Ser Ser Trp Asn Ser Ile Gln Asn Lys Lys Gly Lys Gln
85 90 95

Gly Ala Tyr Asp Asn Glu Glu Ala Ala Ala His Thr Tyr Asp Leu Ala
100 105 110

Ala Leu Lys Tyr Trp Gly Ser Glu Thr Thr Leu Asn Phe Pro Ile Glu
115 120 125

Thr Tyr Thr Lys Glu Ile Glu Glu Met Gln Lys Val Thr Lys Glu Glu
130 135 140

pctau2012001598-seql

Tyr Leu Ala Ser Leu Arg Arg Gln Ser Ser Gly Phe Ser Arg Gly Val
145 150 155 160
Ser Lys Tyr Arg Gly Val Ala Arg His His His Asn Gly Arg Trp Glu
165 170 175
Ala Arg Ile Gly Arg Val Tyr Gly Asn Lys Tyr Leu Tyr Leu Gly Thr
180 185 190
Tyr Asn Thr Gln Glu Glu Ala Ala Ala Tyr Asp Met Ala Ala Ile
195 200 205
Gln Tyr Arg Gly Ala Asn Ala Val Thr Asn Phe Asp Val Ser Asn Tyr
210 215 220
Ile Glu Arg Leu Arg Lys Lys Gly Ile Pro Ile Asp Arg Ile Leu Gln
225 230 235 240
Glu Gln Gln Leu Leu Asn Asn Ser Val Asp Ser Ser Val Glu Val Glu
245 250 255
Val Glu Gln Pro Thr Pro Pro Pro Gln Gln Gln Gln Glu Glu Gln Glu
260 265 270
Gln Lys Ile Val Ser Ser Ser Ser Gln Leu Gln Cys Ser Gln Leu Asn
275 280 285
Ser Ser Leu Asp Gly Thr Pro Pro Met Val Ile Met Asp Thr Ile Glu
290 295 300
Glu His Glu Leu Ala Trp Ser Phe Cys Met Asp Ser Gly Leu Ser Leu
305 310 315 320
Thr Met Pro Asp Leu Pro Leu Glu Asn Ser Cys Glu Leu Pro Asp Leu
325 330 335
Phe Asp His Thr Gly Phe Glu Asp Asn Ile Asp Leu Ile Phe Asp Ala
340 345 350
Cys Cys Tyr Gly Lys Glu Ala Asn Pro Ala Gly Tyr Thr Leu Glu Asp
355 360 365
Asn Ser Thr Gly Gly Val Glu Glu Asp Arg Leu Ser Ser Asp Ser Val
370 375 380
Ser Asn Ser Pro Thr Ser Ser Thr Thr Thr Ser Val Ser Cys Asn Tyr
385 390 395 400
Ser Val

pctau2012001598-seql

<210> 288
 <211> 409
 <212> PRT
 <213> Vitis vinifera

<400> 288

Met Val Lys Arg Ser Ser Pro Gly Ser Ser Ser Ser Pro Ser Ser Ser
 1 5 10 15

Ser Thr Ser Ser Asp Ala Ala Ser Arg Pro Ala Pro Pro Ser Gly Gly
 20 25 30

Lys Pro Lys Ser Arg Lys Lys Glu Ala Lys Lys Asn Ser Asn Gly Asn
 35 40 45

Gly Ser Asn Ser Lys Asn Lys Arg Thr Ser Ile Tyr Arg Gly Val Thr
 50 55 60

Lys His Arg Trp Thr Gly Arg Phe Glu Ala His Leu Trp Asp Lys Ser
 65 70 75 80

Ser Trp Asn Asp Ile Ser Asn Lys Arg Gly Arg Gln Gly Ala Tyr Tyr
 85 90 95

Asn Glu Glu Ala Ala Ala Arg Thr Tyr Asp Leu Ala Ala Leu Lys Tyr
 100 105 110

Trp Gly Pro Thr Thr Pro Leu Asn Phe Pro Leu Glu Thr Tyr Gln Lys
 115 120 125

Asp Ala Glu Glu Met Glu Lys Met Ser Lys Glu Glu Tyr Leu Ala Leu
 130 135 140

Leu Arg Arg Gln Ser Asn Gly Phe Ser Arg Gly Val Ser Lys His His
 145 150 155 160

His Asn Gly Arg Trp Glu Ala Arg Ile Gly Arg Val Leu Gly Asn Lys
 165 170 175

Tyr Leu Tyr Leu Gly Thr Tyr Ser Thr Gln Glu Glu Ala Ala Ala Ala
 180 185 190

Tyr Asp Met Ala Ala Ile Glu Tyr Arg Gly Leu Asn Ala Val Thr Asn
 195 200 205

Phe Asp Ile Ser Asn Tyr Val Lys Leu Gly Arg Val Glu Ala Gln Val
 210 215 220

Gln Glu Leu Ala Gln Gln Leu Gln Pro Asn Thr Pro Ile Gly Pro Gln
 225 230 235 240

pctau2012001598-seql

Asn Glu Leu Gln Lys Glu Glu Glu Glu Gln Leu Gln Glu Pro Val Leu
245 250 255

Ser Ser Ser Gln His Leu Pro Ser Met Asp Ser Ser Ala Met Glu Ile
260 265 270

Met Asp Pro Ala Asp Asp Pro Asp Leu Pro Trp Asn Phe Cys Ala Tyr
275 280 285

Ser Thr Leu Leu Val Pro Asp Val Pro Leu Gly Lys Gly Gly Glu Leu
290 295 300

Ser Asp Leu Phe Tyr Glu Lys Gly Phe Glu Asp Asn Ile Asp Tyr Met
305 310 315 320

Phe Glu Gly Ala Ala Gly Asn Glu Glu Glu Ser Asn Ser Ala Glu Asn
325 330 335

Gly Val Lys Glu Asn Gly Phe Met His Glu Leu Glu Val Asp Gly Lys
340 345 350

Leu Gln Asn Val Val Gly Phe Phe Phe Leu Ser Phe Phe Phe Leu Pro
355 360 365

Lys Arg Ala Gly Ile Arg Lys Arg Gly Val Asp Ser Cys Met Gln Leu
370 375 380

Phe Leu Tyr Phe Val Phe Leu Phe Tyr Pro Phe Leu Pro Glu Val Ser
385 390 395 400

Lys Phe Leu Phe His Leu Ser Leu Asp
405

<210> 289

<211> 420

<212> PRT

<213> Brachypodium distachyon

<400> 289

Met Lys Arg Ser Pro Pro Gln Pro Ser Pro Ser Pro Ser Ser Ser Pro
1 5 10 15

Ala Ser Ser Ser Ser Ser Pro Ser Ser Asp Ser Ser Ser Ser Ile
20 25 30

Ala Ile Pro Arg Lys Arg Ala Arg Thr Ala Ala Ala Ala Ala Gly Gly
35 40 45

Gly Lys Ala Arg Ala Ala Ala Lys Arg Pro Lys Lys Asp Gly Lys
50 55 60

Asp Ser Gly Ser Ser Ser Asn Gly Gly Gly Gly Gly Gly Gly Lys Arg
Page 347

pctau2012001598-seql

65					70					75					80
Ser	Ser	Ile	Tyr	Arg ₈₅	Gly	Val	Thr	Arg	His ₉₀	Arg	Trp	Thr	Gly	Arg ₉₅	Phe
Glu	Ala	His ₁₀₀	Leu	Trp	Asp	Lys	Asn	Cys ₁₀₅	Phe	Thr	Ser	Leu	Gln ₁₁₀	Asn	Lys
Lys	Lys	Gly ₁₁₅	Arg	Gln	Val	Tyr	Leu ₁₂₀	Gly	Ala	Tyr	Asp	Thr ₁₂₅	Glu	Glu	Ala
Ala	Ala ₁₃₀	Arg	Ala	Tyr	Asp	Leu ₁₃₅	Ala	Ala	Leu	Lys	Tyr ₁₄₀	Trp	Gly	Pro	Glu
Thr ₁₄₅	Thr	Leu	Asn	Phe	Ser ₁₅₀	Ala	Asp	Asp	Tyr	Gly ₁₅₅	Lys	Glu	Arg	Ser	Glu ₁₆₀
Met	Glu	Ala	Val	Ser ₁₆₅	Arg	Glu	Glu	Tyr	Leu ₁₇₀	Ala	Ala	Leu	Arg	Arg ₁₇₅	Arg
Ser	Ser	Gly	Phe ₁₈₀	Ser	Arg	Gly	Val	Ser ₁₈₅	Lys	Tyr	Arg	Gly	Val ₁₉₀	Ala	Arg
His	His	His ₁₉₅	Asn	Gly	Arg	Trp	Glu ₂₀₀	Ala	Arg	Ile	Gly	Arg ₂₀₅	Val	Leu	Gly
Asn	Lys ₂₁₀	Tyr	Leu	Tyr	Leu	Gly ₂₁₅	Thr	Phe	Asp	Thr	Gln ₂₂₀	Glu	Glu	Ala	Ala
Arg ₂₂₅	Ala	Tyr	Asp	Leu	Ala ₂₃₀	Ala	Ile	Gln	Tyr	Arg ₂₃₅	Gly	Ala	Asn	Ala	Val ₂₄₀
Thr	Asn	Phe	Asp	Ile ₂₄₅	Ser	Arg	Tyr	Leu	Asp ₂₅₀	Gln	Pro	Gln	Leu	Leu ₂₅₅	Glu
Gln	Leu	Gln	Gln ₂₆₀	Gln	Gln	Gly	Pro	Gln ₂₆₅	Val	Val	Ala	Ala	Leu ₂₇₀	Gln	Glu
Glu	Ala	Gln ₂₇₅	Arg	Asp	His	Gln	Ser ₂₈₀	Asp	Asn	Ala	Val	Gln ₂₈₅	Glu	Leu	Asn
Ser	Gly ₂₉₀	Glu	Ala	Gln	Thr	Pro ₂₉₅	Gly	Gly	Ile	Asp	Glu ₃₀₀	Pro	Ile	Ala	Ile
Gly ₃₀₅	Asp	Ser	Thr	Glu	Asp ₃₁₀	Ile	Asn	Thr	Ser	Leu ₃₁₅	Thr	Val	Asp	Asp	Ile ₃₂₀
Ile	Glu	Glu	Ser	Leu ₃₂₅	Trp	Ser	Pro	Tyr	Glu ₃₃₀	Phe	Asp	Ile	Met	Ala ₃₃₅	Gly
Val	Asn	Val	Ser	Asn	Ser	Met	Asn	Leu	Ser	Glu	Leu	Phe	Ser	Asp	Val

pctau2012001598-seql

340

345

350

Ala Phe Glu Gly Asn Ile Gly Cys Leu Phe Glu Glu Cys Ser Gly Ile
 355 360 365

Asp Asp Cys Ser Ser Arg His Gly Ala Gly Leu Ala Ala Phe Gly Leu
 370 375 380

Phe Thr Glu Gly Asp Asp Lys Leu Lys Asp Val Ser Glu Met Glu Met
 385 390 395 400

Glu Val Asn Pro Gln Ala Asn Asp Val Ser Cys Pro Pro Lys Met Ile
 405 410 415

Thr Val Cys Asn
 420

<210> 290
 <211> 423
 <212> PRT
 <213> Hordeum vulgare

<400> 290

Met Lys Arg Ser Pro Pro Pro Gln Pro Ser Pro Ser Ser Ser Pro Ala
 1 5 10 15

Cys Ser Pro Ser Pro Ser Ser Pro Ser Ser Asp Ser Ser Ser Ile
 20 25 30

Ala Ile Pro Arg Lys Arg Ala Arg Thr Gln Lys Ala Gly Ser Ala Lys
 35 40 45

Ala Lys Ala Ala Pro Lys Arg Ala Lys Lys Asp Ser Gly Arg Ser Thr
 50 55 60

Lys Asp Ser Asp Ala Ser Ala Asn Gly Ala Ala Ala Ser Gly Lys Arg
 65 70 75 80

Ser Ser Ile Tyr Arg Gly Val Thr Arg His Arg Trp Thr Gly Arg Phe
 85 90 95

Glu Ala His Leu Trp Asp Lys Asn Cys Phe Thr Ser Ile Gln Asn Lys
 100 105 110

Lys Lys Gly Arg Gln Val Tyr Leu Gly Ala Tyr Asp Thr Glu Glu Ala
 115 120 125

Ala Ala Arg Ala Tyr Asp Leu Ala Ala Leu Lys Tyr Trp Gly Pro Glu
 130 135 140

Thr Thr Leu Asn Phe Thr Val Asp Glu Tyr Ala Lys Glu Arg Ser Glu
 145 150 155 160

pctau2012001598-seql

Met Glu Ala Val Ser Arg Glu Glu Tyr Leu Ala Ala Leu Arg Arg Arg
165 170 175
Ser Ser Gly Phe Ser Arg Gly Val Ser Lys Tyr Arg Gly Val Ala Arg
180 185 190
His His His Asn Gly Arg Trp Glu Ala Arg Ile Gly Arg Val Leu Gly
195 200 205
Asn Lys Tyr Leu Tyr Leu Gly Thr Phe Asp Thr Gln Glu Glu Ala Ala
210 215 220
Arg Ala Tyr Asp Leu Ala Ala Ile Glu Tyr Arg Gly Ala Asn Ala Val
225 230 235 240
Thr Asn Phe Asp Ile Ser Arg Tyr Leu Asp Gln Pro Gln Leu Leu Ala
245 250 255
Gln Leu Glu Gln Gly Pro Gln Val Val Pro Ala Leu Gln Glu Glu Leu
260 265 270
Gln His Asp His Gln Ser Asp Asn Ala Val Gln Glu Leu Asn Ser Gly
275 280 285
Glu Ala Gln Lys Pro Gly Ser Val Ser Glu Pro Ile Ala Val Asp Asp
290 295 300
Thr Asp Asn Thr Gly Asp Ile Gly Ala Pro Leu Val Phe Asp Ser Gly
305 310 315 320
Val Glu Glu Asn Leu Trp Ser Pro Cys Met Asp Tyr Asp Val Asp Pro
325 330 335
Ile Phe Gly Pro Asn Ile Ser Ser Ser Met Asn Leu Ser Glu Trp Phe
340 345 350
Asn Asp Pro Ala Phe Glu Ser Asn Ile Gly Tyr Met Phe Glu Gly Cys
355 360 365
Ser Asp Val Asp Asp Cys Ser Thr Arg His Gly Ala Gly Leu Ser Ala
370 375 380
Leu Gly Phe Leu Lys Glu Gly Asp Asp Lys Leu Lys Asp Gly Ser Asp
385 390 395 400
Met Glu Ala Glu Ile Thr Pro Gln Ala Asn Asp Val Ser Cys Pro Pro
405 410 415
Lys Met Ile Thr Val Cys Asn
420

pctau2012001598-seql

<210> 291
 <211> 443
 <212> PRT
 <213> Oryza sativa

<400> 291

```

Met Ala Lys Arg Ser Ser Pro Asp Pro Ala Ser Ser Ser Pro Ser Ala
 1          5          10          15

Ser Ser Ser Pro Ser Ser Pro Ser Ser Ser Ser Glu Asp Ser Ser
 20          25          30

Ser Pro Met Ser Met Pro Cys Lys Arg Arg Ala Arg Pro Arg Thr Asp
 35          40          45

Lys Ser Thr Gly Lys Ala Lys Arg Pro Lys Lys Glu Ser Lys Glu Val
 50          55          60

Val Asp Pro Ser Ser Asn Gly Gly Gly Gly Gly Lys Arg Ser Ser Ile
 65          70          75

Tyr Arg Gly Val Thr Arg His Arg Trp Thr Gly Arg Phe Glu Ala His
 85          90          95

Leu Trp Asp Lys Asn Cys Ser Thr Ser Leu Gln Asn Lys Lys Lys Gly
100          105          110

Arg Gln Gly Ala Tyr Asp Ser Glu Glu Ala Ala Ala Arg Ala Tyr Asp
115          120          125

Leu Ala Ala Leu Lys Tyr Trp Gly Pro Glu Thr Val Leu Asn Phe Pro
130          135          140

Leu Glu Glu Tyr Glu Lys Glu Arg Ser Glu Met Glu Gly Val Ser Arg
145          150          155

Glu Glu Tyr Leu Ala Ser Leu Arg Arg Arg Ser Ser Gly Phe Ser Arg
165          170          175

Gly Val Ser Lys Tyr Arg Gly Val Ala Arg His His His Asn Gly Arg
180          185          190

Trp Glu Ala Arg Ile Gly Arg Val Leu Gly Asn Lys Tyr Leu Tyr Leu
195          200          205

Gly Thr Phe Asp Thr Gln Glu Glu Ala Ala Lys Ala Tyr Asp Leu Ala
210          215          220

Ala Ile Glu Tyr Arg Gly Ala Asn Ala Val Thr Asn Phe Asp Ile Ser
225          230          235          240
    
```


pctau2012001598-seql

Cys Tyr Leu Asp Gln Pro Gln Leu Leu Ala Gln Leu Gln Gln Glu Pro
245 250 255

Gln Leu Leu Ala Gln Leu Gln Gln Glu Pro Gln Val Val Pro Ala Leu
260 265 270

His Glu Glu Pro Gln Asp Asp Asp Arg Ser Glu Asn Ala Val Gln Glu
275 280 285

Leu Ser Ser Ser Glu Ala Asn Thr Ser Ser Asp Asn Asn Glu Pro Leu
290 295 300

Ala Ala Asp Asp Ser Ala Glu Cys Met Asn Glu Pro Leu Pro Ile Val
305 310 315 320

Asp Gly Ile Glu Glu Ser Leu Trp Ser Pro Cys Leu Asp Tyr Glu Leu
325 330 335

Asp Thr Met Pro Gly Ala Tyr Phe Ser Asn Ser Met Asn Phe Ser Glu
340 345 350

Trp Phe Asn Asp Glu Ala Phe Glu Gly Gly Met Glu Tyr Leu Phe Glu
355 360 365

Gly Cys Ser Ser Ile Thr Glu Gly Gly Asn Ser Met Asp Asn Ser Gly
370 375 380

Val Thr Glu Tyr Asn Leu Phe Glu Glu Cys Asn Met Leu Glu Lys Asp
385 390 395 400

Ile Ser Asp Phe Leu Asp Lys Asp Ile Ser Asp Phe Leu Asp Lys Asp
405 410 415

Ile Ser Ile Ser Asp Gly Glu Arg Ile Ser Pro Gln Ala Asn Asn Ile
420 425 430

Ser Cys Pro Gln Lys Met Ile Ser Val Cys Asn
435 440

<210> 292
<211> 420
<212> PRT
<213> Sorghum bicolor

<400> 292

Met Asp Met Glu Arg Ser Gln Gln Gln Lys Ser Pro Thr Glu Ser Pro
1 5 10 15

Pro Pro Pro Ser Pro Ser Ser Ser Ser Ser Val Ser Ala Asp Thr
20 25 30

pctau2012001598-seql

Val	Leu	Pro	Pro	Pro	Gly	Lys	Arg	Arg	Arg	Ala	Ala	Thr	Thr	Ala	Lys
		35					40					45			
Ala	Lys	Ala	Gly	Ala	Lys	Pro	Lys	Arg	Ala	Arg	Lys	Asp	Ala	Ala	Ala
	50					55					60				
Ala	Ala	Asp	Pro	Pro	Pro	Pro	Pro	Ala	Ala	Ala	Ala	Ala	Gly	Lys	Arg
65					70					75					80
Ser	Ser	Val	Tyr	Arg	Gly	Val	Thr	Arg	His	Arg	Trp	Thr	Gly	Arg	Phe
				85					90					95	
Glu	Ala	His	Leu	Trp	Asp	Lys	His	Cys	Leu	Ala	Ala	Leu	His	Asn	Lys
			100					105					110		
Lys	Lys	Gly	Arg	Gln	Val	Tyr	Leu	Gly	Ala	Tyr	Asp	Ser	Glu	Glu	Ala
		115					120					125			
Ala	Ala	Arg	Ala	Tyr	Asp	Leu	Ala	Ala	Leu	Lys	Tyr	Trp	Gly	Pro	Glu
	130					135					140				
Thr	Leu	Leu	Asn	Phe	Pro	Val	Glu	Asp	Tyr	Ser	Ser	Glu	Met	Pro	Glu
145					150					155					160
Met	Glu	Gly	Val	Ser	Arg	Glu	Glu	Tyr	Leu	Ala	Ser	Leu	Arg	Arg	Arg
				165					170					175	
Ser	Ser	Gly	Phe	Ser	Arg	Gly	Val	Ser	Lys	Tyr	Arg	Gly	Val	Ala	Arg
			180					185					190		
His	His	His	Asn	Gly	Arg	Trp	Glu	Ala	Arg	Ile	Gly	Arg	Val	Phe	Gly
		195					200					205			
Asn	Lys	Tyr	Leu	Tyr	Leu	Gly	Thr	Phe	Asp	Thr	Gln	Glu	Glu	Ala	Ala
	210					215					220				
Lys	Ala	Tyr	Asp	Leu	Ala	Ala	Ile	Glu	Tyr	Arg	Gly	Val	Asn	Ala	Val
225					230					235					240
Thr	Asn	Phe	Asp	Ile	Ser	Cys	Tyr	Leu	Asp	His	Pro	Leu	Phe	Leu	Ala
				245					250					255	
Gln	Leu	Gln	Gln	Glu	Pro	Gln	Val	Val	Pro	Ala	Leu	Asn	Gln	Glu	Ala
			260					265					270		
Gln	Pro	Asp	Gln	Ser	Glu	Thr	Glu	Thr	Ile	Ala	Gln	Glu	Ser	Val	Ser
		275					280					285			
Ser	Glu	Ala	Lys	Thr	Pro	Asp	Asp	Asn	Ala	Glu	Pro	Asp	Asp	Asn	Ala
	290					295					300				

pctau2012001598-seql

Glu Pro Asp Asp Ile Ala Glu Pro Leu Ile Thr Val Asp Asp Ser Ile
305 310 315 320

Glu Glu Ser Leu Trp Ser Pro Cys Met Asp Tyr Glu Leu Asp Thr Met
325 330 335

Ser Arg Ser Asn Phe Gly Ser Ser Ile Asn Leu Ser Glu Trp Phe Asn
340 345 350

Asp Ala Asp Phe Asp Ser Asn Ile Gly Cys Leu Phe Asp Gly Cys Ser
355 360 365

Ala Val Asp Glu Gly Gly Lys Asp Gly Val Gly Leu Ala Asp Phe Ser
370 375 380

Leu Leu Glu Asp Phe Ser Leu Phe Glu Ala Gly Asp Gly Gln Leu Lys
385 390 395 400

Asp Val Leu Ser Asp Met Glu Glu Gly Ile Gln Pro Pro Thr Met Ile
405 410 415

Ser Val Cys Asn
420

<210> 293
<211> 395
<212> PRT
<213> Zea mays

<400> 293

Met Glu Arg Ser Gln Arg Gln Ser Pro Pro Pro Pro Ser Pro Ser Ser
1 5 10 15

Ser Ser Ser Ser Val Ser Ala Asp Thr Val Leu Val Pro Pro Gly Lys
20 25 30

Arg Arg Arg Ala Ala Thr Ala Lys Ala Gly Ala Glu Pro Asn Lys Arg
35 40 45

Ile Arg Lys Asp Pro Ala Ala Ala Ala Gly Lys Arg Ser Ser Val
50 55 60

Tyr Arg Gly Val Thr Arg His Arg Trp Thr Gly Arg Phe Glu Ala His
65 70 75 80

Leu Trp Asp Lys His Cys Leu Ala Ala Leu His Asn Lys Lys Lys Gly
85 90 95

Arg Gln Val Tyr Leu Gly Ala Tyr Asp Ser Glu Glu Ala Ala Ala Arg
100 105 110

Ala Tyr Asp Leu Ala Ala Leu Lys Tyr Trp Gly Pro Glu Thr Leu Leu
Page 354

115

120

125

Asn Phe Pro Val Glu Asp Tyr Ser Ser Glu Met Pro Glu Met Glu Ala
 130 135 140

Val Ser Arg Glu Glu Tyr Leu Ala Ser Leu Arg Arg Arg Ser Ser Gly
 145 150 155 160

Phe Ser Arg Gly Val Ser Lys Tyr Arg Gly Val Ala Arg His His His
 165 170 175

Asn Gly Arg Trp Glu Ala Arg Ile Gly Arg Val Phe Gly Asn Lys Tyr
 180 185 190

Leu Tyr Leu Gly Thr Phe Asp Thr Gln Glu Glu Ala Ala Lys Ala Tyr
 195 200 205

Asp Leu Ala Ala Ile Glu Tyr Arg Gly Val Asn Ala Val Thr Asn Phe
 210 215 220

Asp Ile Ser Cys Tyr Leu Asp His Pro Leu Phe Leu Ala Gln Leu Gln
 225 230 235 240

Gln Glu Pro Gln Val Val Pro Ala Leu Asn Gln Glu Pro Gln Pro Asp
 245 250 255

Gln Ser Glu Thr Gly Thr Thr Glu Gln Glu Pro Glu Ser Ser Glu Ala
 260 265 270

Lys Thr Pro Asp Gly Ser Ala Glu Pro Asp Glu Asn Ala Val Pro Asp
 275 280 285

Asp Thr Ala Glu Pro Leu Ser Thr Val Asp Asp Ser Ile Glu Glu Gly
 290 295 300

Leu Trp Ser Pro Cys Met Asp Tyr Glu Leu Asp Thr Met Ser Arg Pro
 305 310 315 320

Asn Phe Gly Ser Ser Ile Asn Leu Ser Glu Trp Phe Ala Asp Ala Asp
 325 330 335

Phe Asp Cys Asn Ile Gly Cys Leu Phe Asp Gly Cys Ser Ala Ala Asp
 340 345 350

Glu Gly Ser Lys Asp Gly Val Gly Leu Ala Asp Phe Ser Leu Phe Glu
 355 360 365

Ala Gly Asp Val Gln Leu Lys Asp Val Leu Ser Asp Met Glu Glu Gly
 370 375 380

Ile Gln Pro Pro Ala Met Ile Ser Val Cys Asn
 Page 355

pctau2012001598-seql

385

390

395

<210> 294

<211> 413

<212> PRT

<213> Brachypodium distachyon

<400> 294

Met Glu Ala Tyr Cys Ser Thr Leu Val Lys Asp Glu Leu Ile Asn Gly
1 5 10 15

Gly Gly Gly Gly Ser Ala Gly Gly Met Arg Tyr Cys Glu Ala Ala Pro
20 25 30

Arg Val Ser Pro Pro Val Ala Ile Lys Ser Val Lys Arg Arg Lys Arg
35 40 45

Glu Pro Pro Ala Val Ser Gly Met Thr Thr Val Ser Gly Gly Gly Lys
50 55 60

Asp Gly Asp Lys Ser Ala Gly Asn Ala Ala Ala Lys Arg Ser Ser Arg
65 70 75 80

Phe Arg Gly Val Ser Arg His Arg Trp Thr Gly Arg Phe Glu Ala His
85 90 95

Leu Trp Asp Lys Gly Thr Trp Asn Pro Thr Gln Lys Lys Lys Gly Lys
100 105 110

Gln Val Tyr Leu Gly Ala Tyr Asn Glu Glu Glu Ala Ala Ala Arg Ala
115 120 125

Tyr Asp Leu Ala Ala Leu Lys Tyr Trp Gly Pro Thr Thr Tyr Thr Asn
130 135 140

Phe Pro Val Val Asp Tyr Glu Lys Glu Leu Lys Val Met Gln Gly Val
145 150 155 160

Ser Arg Glu Glu Tyr Leu Ala Ser Ile Arg Arg Lys Ser Asn Gly Phe
165 170 175

Ser Arg Gly Val Ser Lys Tyr Arg Gly Val Ala Arg His His His Asn
180 185 190

Gly Arg Trp Glu Ala Arg Ile Gly Arg Val Phe Gly Asn Lys Tyr Leu
195 200 205

Tyr Leu Gly Thr Tyr Ser Thr Gln Glu Glu Ala Ala Arg Ala Tyr Asp
210 215 220

Ile Ala Ala Ile Glu Tyr Arg Gly Ile Asn Ala Val Thr Asn Phe Asp
225 230 235 240

pctau2012001598-seql

Leu Ser Ser Tyr Ile Arg Trp Leu Lys Pro Asn Ser Thr Ile Asn Thr
245 250 255

Asn Thr Pro Ala Ala Glu Leu Ala Ile Leu Gly Gly Gly Gly Thr Pro
260 265 270

Ala Ala Leu Ile Thr Pro Pro Pro Thr Met His Val Pro Arg Leu Leu
275 280 285

Pro Pro Leu Val Lys Gly Arg Gly Ser Ser Ile Ala Asp Asp Val Ser
290 295 300

Ala Gly Ser Cys Val Phe Gly Gly Pro Ser Pro Ser Pro Ser Pro Thr
305 310 315 320

Thr Thr Ala Leu Ser Leu Leu Leu Arg Ser Ser Val Phe Gln Glu Leu
325 330 335

Val Ala Gln Gln Gln Pro Pro Ser Thr Val Asp Asp Asp Asp Asp Ile
340 345 350

Gly Gly His Ala Ala Val Ser Asp Ala Ala Gln Arg Ala Ala Glu Glu
355 360 365

Asn Glu Glu Ser Phe Gly Glu Val Leu Tyr Gly Ala Gly Glu Gly Glu
370 375 380

Ala Ala Thr Ala Phe Ser Cys Ser Met Tyr Glu Leu Gly Leu Asp Asp
385 390 395 400

Asn Phe Ala Arg Ile Glu Glu Ser Leu Trp Gly Cys Leu
405 410

<210> 295
<211> 423
<212> PRT
<213> Brachypodium sylvaticum

<400> 295

Met Glu Ala Tyr Cys Ser Ser Leu Val Lys Asp Glu Leu Ile Asn Gly
1 5 10 15

Gly Gly Gly Gly Ala Gly Gly Met Arg Tyr Cys Glu Ala Ala Pro Arg
20 25 30

Val Ser Pro Pro Val Ala Ile Lys Ser Val Lys Arg Arg Lys Arg Glu
35 40 45

Pro Pro Ala Val Ser Gly Met Thr Thr Val Ser Gly Gly Gly Gly Gly
50 55 60

pctau2012001598-seql

Asn Gly Lys Asp Gly Asp Lys Ser Ala Gly Asn Ala Ala Ala Ala Lys
 65 70 75 80
 Arg Ser Ser Arg Phe Arg Gly Val Ser Arg His Arg Trp Thr Gly Arg
 85 90 95
 Phe Glu Ala His Leu Trp Asp Lys Gly Thr Trp Asn Pro Thr Gln Lys
 100 105 110
 Lys Lys Gly Lys Gln Val Tyr Leu Gly Ala Tyr Asn Glu Glu Glu Ala
 115 120 125
 Ala Ala Arg Ala Tyr Asp Leu Ala Ala Leu Lys Tyr Trp Gly Pro Thr
 130 135 140
 Thr Tyr Thr Asn Phe Pro Val Val Asp Tyr Glu Lys Glu Leu Lys Val
 145 150 155 160
 Met Gln Gly Val Ser Arg Glu Glu Tyr Leu Ala Ser Ile Arg Arg Lys
 165 170 175
 Ser Asn Gly Phe Ser Arg Gly Val Ser Lys Tyr Arg Gly Val Ala Arg
 180 185 190
 His His His Asn Gly Arg Trp Glu Ala Arg Ile Gly Arg Val Phe Gly
 195 200 205
 Asn Lys Tyr Leu Tyr Leu Gly Thr Tyr Ser Thr Gln Glu Glu Ala Ala
 210 215 220
 Arg Ala Tyr Asp Ile Ala Ala Ile Glu Tyr Arg Gly Ile Asn Ala Val
 225 230 235 240
 Thr Asn Phe Asp Leu Ser Ser Tyr Ile Arg Trp Leu Lys Pro Asn Ser
 245 250 255
 Ala Ala Asn Thr Asn Thr Pro Pro Ala Ala Ala Glu Leu Ala Ile
 260 265 270
 Leu Gly Gly Ala Pro Ala Ala Leu Ile Ser Pro Ala Pro Ala Pro Thr
 275 280 285
 Thr Met Arg Val Pro Arg Leu Leu Pro Pro Leu Val Arg Gly Arg Gly
 290 295 300
 Gly Ser Ile Pro Asp Asp Val Ser Ala Gly Gly Ser Cys Val Phe Gly
 305 310 315 320
 Ser Pro Ser Pro Ser Pro Ser Pro Thr Thr Thr Ser Ala Leu Ser Leu
 325 330 335

pctau2012001598-seql

Leu Leu Arg Ser Ser Val Phe Gln Glu Leu Val Ala Gln Gln Gln Pro
340 345 350

Pro Ser Ile Val Asp Asp Asp Asp Gly Val Gly Gly Gln Glu Ala Val
355 360 365

Ser Asp Ala Ala Glu Arg Ala Ala Glu Glu Asn Glu Glu Ser Phe Gly
370 375 380

Glu Val Leu Tyr Gly Ala Gly Glu Gly Glu Ala Ala Ala Ala Phe Ser
385 390 395 400

Cys Ser Met Tyr Glu Leu Gly Leu Asp Asp Ser Phe Ala Arg Ile Glu
405 410 415

Glu Ser Leu Trp Gly Cys Leu
420

<210> 296
<211> 399
<212> PRT
<213> Oryza sativa

<400> 296

Met Glu Thr Tyr Gly Leu Val Lys Asp Glu Leu Leu His Gly Ile Gly
1 5 10 15

Gly Gly Gln Gly Arg Leu Tyr Cys Glu Val Lys Pro Thr Ala Ala Pro
20 25 30

Ala Val Ile Thr Ala Ala Gly Gly Gly Ala Lys Ser Val Lys Arg Arg
35 40 45

Lys Arg Glu Pro Ser Ala Ala Ala Met Ser Ala Val Thr Val Ala Gly
50 55 60

Asn Gly Lys Glu Ala Gly Gly Ser Asn Ala Ala Asn Lys Arg Ser Ser
65 70 75 80

Arg Phe Arg Gly Val Ser Arg His Arg Trp Thr Gly Arg Phe Glu Ala
85 90 95

His Leu Trp Asp Lys Gly Thr Trp Asn Pro Thr Gln Lys Lys Lys Gly
100 105 110

Lys Gln Val Tyr Leu Gly Ala Tyr Asn Glu Glu Asp Ala Ala Ala Arg
115 120 125

Ala Tyr Asp Leu Ala Ala Leu Lys Tyr Trp Gly Pro Thr Thr Tyr Thr
130 135 140

pctau2012001598-seql

Asn Phe Pro Val Ala Asp Tyr Glu Lys Glu Leu Lys Leu Met Gln Gly
145 150 155 160

Val Ser Lys Glu Glu Tyr Leu Ala Ser Ile Arg Arg Lys Ser Asn Gly
165 170 175

Phe Ser Arg Gly Val Ser Lys Tyr Arg Gly Val Ala Arg His His His
180 185 190

Asn Gly Arg Trp Glu Ala Arg Ile Gly Arg Val Phe Gly Asn Lys Tyr
195 200 205

Leu Tyr Leu Gly Thr Tyr Ser Thr Gln Glu Glu Ala Ala Arg Ala Tyr
210 215 220

Asp Ile Ala Ala Ile Glu Tyr Arg Gly Ile Asn Ala Val Thr Asn Phe
225 230 235 240

Asp Leu Ser Thr Tyr Ile Arg Trp Leu Lys Pro Pro Ser Ser Ser Ser
245 250 255

Ala Ala Gly Thr Pro His His His Gly Gly Gly Met Val Val Gly Ala
260 265 270

Asp Arg Val Leu Ala Pro Ala Gln Ser Tyr Pro Ile Ser Ala Ala Ala
275 280 285

Asp Asp Asp Val Ala Gly Cys Trp Arg Pro Leu Pro Ser Pro Ser Ser
290 295 300

Ser Thr Thr Thr Ala Leu Ser Leu Leu Leu Arg Ser Ser Met Phe Gln
305 310 315 320

Glu Leu Val Ala Arg Gln Pro Val Val Glu Gly Asp Asp Gly Gln Leu
325 330 335

Ala Val Val Ser Gly Asp Asp Ala Asp Ala Asp Ala Asp Ser Asp Val
340 345 350

Lys Glu Pro Pro Pro Glu Ser Glu Tyr Gly Glu Val Phe Ala Ser Asp
355 360 365

Glu Ala Ala Ala Ala Ala Ala Tyr Gly Cys Ser Met Tyr Glu Leu Asp
370 375 380

Asp Ser Phe Ala Leu Ile Asp Asp Ser Val Trp Asn Cys Leu Ile
385 390 395

<210> 297
<211> 488
<212> PRT
<213> Sorghum bicolor

pctau2012001598-seql

<400> 297

```

Met Glu Thr Tyr Ser Leu Gln Val Lys Asp Glu Leu His Gly Gly Gly
 1      5      10      15

Ile Gly Ile Gly Gly Gly Gln Gly Leu Tyr Cys Gly Ala Thr Pro
 20      25      30

Arg Pro Ala Ala Pro Ala Ala Thr Gly Gly Gly Gly Gly Gly Asp
 35      40      45

Gly Ala Val Lys Ser Asn Lys Arg Ser Arg Lys Arg Glu Pro Pro Pro
 50      55      60

Pro Pro Pro Ser Ser Leu Val Thr Met Ser Asn Gly Gly Lys Asp Glu
 65      70      75

Ala Val Ala Gly Ser Gly Asp Lys Ser Ala Ser Ser Asn Ser Asn Ala
 85      90      95

Ser Lys Arg Ser Ser Arg Phe Arg Gly Val Ser Arg His Arg Trp Thr
100      105      110

Gly Arg Phe Glu Ala His Leu Trp Asp Lys Gly Thr Trp Asn Pro Thr
115      120      125

Gln Lys Lys Lys Gly Lys Gln Val Tyr Leu Gly Ala Tyr Asn Glu Glu
130      135      140

Asp Ala Ala Ala Arg Ala Tyr Asp Leu Ala Ala Leu Lys Tyr Trp Gly
145      150      155      160

Pro Thr Thr Tyr Thr Asn Phe Pro Val Val Asp Tyr Glu Arg Glu Leu
165      170      175

Lys Val Met Gln Asn Val Ser Lys Glu Glu Tyr Leu Ala Ser Ile Arg
180      185      190

Arg Lys Ser Asn Gly Phe Ser Arg Gly Val Ser Lys Tyr Arg Gly Val
195      200      205

Ala Arg His His His Asn Gly Arg Trp Glu Ala Arg Ile Gly Arg Val
210      215      220

Phe Gly Asn Lys Tyr Leu Tyr Leu Gly Thr Tyr Ser Thr Gln Glu Glu
225      230      235      240

Ala Ala Arg Ala Tyr Asp Ile Ala Ala Ile Glu Tyr Arg Gly Ile Asn
245      250      255

Ala Val Thr Asn Phe Asp Leu Ser Thr Tyr Ile Arg Trp Leu Lys Pro

```

pctau2012001598-seql

260

265

270

Gly Gly Gly Val Glu Asp Ser Ala Ala Gly Thr Pro Thr Ser Gly Val
 275 280 285

Arg Ala Pro Gly Ile Pro Pro Ala Ser Leu Ser Leu Gln Ala Gly Gly
 290 295 300

Leu Leu Gln His Pro His Gly Ala Ala Ala Gly Met Leu Gln Val Asp
 305 310 315 320

Val Asp Asp Leu Tyr Arg Gly Gln Leu Ala Ala Ala Arg Gly Ala Ala
 325 330 335

Leu Phe Ser Gly Gly Ile Asp Asp Val Gly Ser Val Tyr Ala Ala Gly
 340 345 350

Ser Ala Gly Pro Ser Pro Thr Ala Leu Cys Ala Gly Arg Pro Ser Pro
 355 360 365

Ser Pro Ser Pro Ser Ser Ser Thr Thr Ala Leu Ser Leu Leu Leu Arg
 370 375 380

Ser Ser Val Phe Gln Glu Leu Val Ala Arg Asn Ala Gly Gly Gly Ala
 385 390 395 400

Ala Gln Gln Gln Gln Leu Val Val Ala Asp Asp Asp Gly Ala Val Ser
 405 410 415

Pro Ala Asp Val Val Asp Ala Lys Val Glu Gln Pro Glu Ala Glu Gly
 420 425 430

Glu Leu Gly Arg His Gly Asp Gln Leu Tyr Gly Ala Ala Arg Ala Asp
 435 440 445

Glu Asp Glu Asp Ala Phe Ala Cys Ser Met Tyr Glu Leu Asp Asp Ser
 450 455 460

Phe Ala Arg Met Glu Gln Ser Leu Trp Gly Cys Leu Arg Ser Ser Asp
 465 470 475 480

Ala Pro Asp Asn Met Asn Asn Leu
 485

<210> 298

<211> 443

<212> PRT

<213> Sorghum bicolor

<400> 298

Met Glu Ser Ser Gly Met Met Met Val Lys Ser Glu Ile Glu Ser Cys
 1 5 10 15

pctau2012001598-seql

Gly Tyr Pro Gly₂₀ Pro Ser Ser Ser Thr₂₅ Ala Pro Ala Ala Gly₃₀ Val Val
 Ile Gly Gly₃₅ Ser Ala Thr Thr Glu₄₀ Arg Gly Glu Gly₄₅ His His His
 His His₅₀ His Gln Val Val Val₅₅ Arg Arg Arg Arg Arg₆₀ Glu Pro Pro Leu
 Leu₆₅ Ala Pro Ile Ala Gly₇₀ Gly Gly Ile Gly₇₅ Lys Pro Leu Pro Ser Ile₈₀
 Thr Val Lys Arg₈₅ Ser Ser Arg Phe Arg Gly₉₀ Val Ser Arg His Arg₉₅ Trp
 Thr Gly Arg Phe₁₀₀ Glu Ala His Leu Trp₁₀₅ Asp Lys Asn Ser Trp₁₁₀ Asn Pro
 Thr Gln Arg₁₁₅ Lys Lys Gly Lys Gln₁₂₀ Val Tyr Leu Gly Ala₁₂₅ Tyr Asp Glu
 Glu Glu₁₃₀ Ala Ala Ala Arg Ala₁₃₅ Tyr Asp Leu Ala Ala₁₄₀ Leu Lys Tyr Trp
 Gly Pro Thr Thr Tyr Thr₁₅₀ Asn Phe Pro Val Met₁₅₅ Asp Tyr Glu Lys Glu₁₆₀
 Leu Lys Ile Met Glu₁₆₅ Asn Leu Thr Lys Glu₁₇₀ Glu Tyr Leu Ala Ser₁₇₅ Leu
 Arg Arg Lys Ser₁₈₀ Ser Gly Phe Ser Arg₁₈₅ Gly Val Ser Lys Tyr₁₉₀ Arg Gly
 Val Ala Arg₁₉₅ His His Gln Asn Gly₂₀₀ Arg Trp Glu Ala Arg₂₀₅ Ile Gly Arg
 Val Phe Gly Asn Lys Tyr Leu₂₁₅ Tyr Leu Gly Thr Tyr₂₂₀ Ser Thr Gln Glu
 Glu Ala Ala Arg Ala Tyr₂₃₀ Asp Ile Ala Ala Ile₂₃₅ Glu Tyr Lys Gly Val₂₄₀
 Asn Ala Val Thr Asn₂₄₅ Phe Asp Leu Arg Ser₂₅₀ Tyr Ile Thr Trp Leu₂₅₅ Lys
 Pro Ser Gly Ala₂₆₀ Pro Ala Ala Phe Asn₂₆₅ Pro Glu Ala Ala Leu₂₇₀ Leu Met
 Gln Ala Ala₂₇₅ Pro Ala Glu Gln Leu₂₈₀ Leu His Pro Ala Glu₂₈₅ Thr Ala Gln

pctau2012001598-seql

Met Leu Pro Arg Val Gly Asn Pro Phe Leu Leu Asp His Gly Ala Ala
290 295 300

Pro Pro Gly Ser Ser Gly Gly Gly Gly Gln Asp Ala Ser Met Ser Ser
305 310 315 320

Met Val Ser Pro Gly Ala Gly Gly Gly Met Arg Arg Arg Gly Ser Ser
325 330 335

Thr Ala Leu Ser Leu Leu Leu Lys Ser Ser Met Phe Arg Gln Leu Val
340 345 350

Glu Lys Asn Ser Asp Ala Glu Glu Gly Val Arg Asp Arg Glu Asp Ala
355 360 365

Ala Ala Ala Ala Ala Ala Ala His Pro Ala Gly Pro Gly Asp Ala Tyr
370 375 380

Glu Tyr His Asn Phe Phe Gln Gly Glu Ala Pro Pro Asp Met Cys Asp
385 390 395 400

Leu Phe Ser Ser Gly Gly Gly Gly Asp His Ala Arg Asn Ala Gly Phe
405 410 415

His Gly Glu Ile Ala Ala Cys Tyr Asp Asp Gly Glu Gly Leu Asp Gly
420 425 430

Trp Asn Gly Phe Gly Asn Met Ser Ser Leu Gln
435 440

<210> 299
<211> 407
<212> PRT
<213> Glycine max

<400> 299

Met Glu Leu Ala Pro Val Lys Ser Glu Leu Ser Pro Arg Ser His Arg
1 5 10 15

Leu Leu Met Ile Asp Gly Ser Glu Val Ile Gly Thr Lys Cys Val Lys
20 25 30

Arg Arg Arg Arg Asp Ser Ser Thr Ala Val Leu Gly Gly Asn Gly Gln
35 40 45

Gln Gly Glu Gln Leu Glu Glu Gln Lys Gln Leu Gly Gly Gln Ser Thr
50 55 60

Ala Thr Thr Val Lys Arg Ser Ser Arg Phe Arg Gly Val Ser Arg His
65 70 75 80

pctau2012001598-seql

Arg Trp Thr Gly Arg Phe Glu Ala His Leu Trp Asp Lys Gly Thr Trp
85 90 95
Asn Pro Thr Gln Lys Lys Lys Gly Lys Gln Val Tyr Leu Gly Ala Tyr
100 105 110
Asn Asp Glu Glu Ala Ala Ala Arg Ala Tyr Asp Leu Ala Ala Leu Lys
115 120 125
Tyr Trp Gly Ile Ser Thr Phe Thr Asn Phe Pro Val Ser Asp Tyr Glu
130 135 140
Lys Glu Ile Glu Ile Met Lys Thr Val Thr Lys Glu Glu Tyr Leu Ala
145 150 155 160
Ser Leu Arg Arg Arg Ser Ser Gly Phe Ser Arg Gly Val Ser Lys Tyr
165 170 175
Arg Gly Val Ala Arg His His His Asn Gly Arg Trp Glu Ala Arg Ile
180 185 190
Gly Arg Val Phe Gly Asn Lys Tyr Leu Tyr Leu Gly Thr Tyr Ser Thr
195 200 205
Gln Glu Glu Ala Ala Arg Ala Tyr Asp Ile Ala Ala Ile Glu Tyr Arg
210 215 220
Gly Ile Asn Ala Val Thr Asn Phe Asp Leu Ser Thr Tyr Ile Arg Trp
225 230 235 240
Leu Arg Pro Gly Thr His Pro Thr Ala Ser His Asp Gln Lys Pro Ser
245 250 255
Thr Asp Ala Gln Pro Phe Ala Thr Ser Asn Ser Met Gln Ala Arg Gly
260 265 270
Asn Ile Glu Val Ser Asn Ser Asn Lys Asn Ser Phe Pro Ser Gly Lys
275 280 285
Leu Asp Ser Thr Lys Lys Arg Asp Phe Ser Lys Tyr Met Asn Pro Leu
290 295 300
Ser Pro Cys Asn Lys Pro Ser Ser Pro Thr Ala Leu Gly Leu Leu Leu
305 310 315 320
Lys Ser Ser Val Phe Arg Glu Leu Met Gln Arg Asn Leu Asn Ser Ser
325 330 335
Ser Glu Glu Ala Glu Glu Val Glu Leu Lys Tyr Pro His Glu Gly Asn
340 345 350

pctau2012001598-seql

Asp Gly Val Gly Gly Ile Tyr Asp Asn Glu Asn Thr Asn Asn Ser Tyr
355 360 365

Phe Cys Ser Ser Asn Ile Ser Arg Leu Pro Asn Leu Glu Ser Ser Glu
370 375 380

Glu Ser Pro Leu Pro Met Tyr His Gly Thr Val Gln Ser Leu Trp Asn
385 390 395 400

Ser Ala Phe Asn Met Ser Asn
405

<210> 300
<211> 406
<212> PRT
<213> Glycine max

<400> 300

Met Glu Leu Ala Pro Val Lys Ser Glu Leu Ser Pro Arg Ser His Arg
1 5 10 15

Leu Val Ile Ile Asp Gly Ser Asp Val Ile Ser Thr Lys Cys Ala Lys
20 25 30

Arg Arg Arg Arg Asp Ser Ser Met Ala Val Leu Gly Gly Asn Gly Gln
35 40 45

Gln Gly Glu Gln Leu Glu Glu Gln Lys Gln Leu Gly Gly Gln Ser Thr
50 55 60

Ala Thr Thr Val Lys Arg Ser Ser Arg Phe Arg Gly Val Ser Arg His
65 70 75 80

Arg Trp Thr Gly Arg Phe Glu Ala His Leu Trp Asp Lys Gly Thr Trp
85 90 95

Asn Pro Thr Gln Lys Lys Lys Gly Lys Gln Val Tyr Leu Gly Ala Tyr
100 105 110

Asn Asp Glu Glu Ala Ala Ala Arg Ala Tyr Asp Leu Ala Ala Leu Lys
115 120 125

Tyr Trp Gly Thr Ser Thr Phe Thr Asn Phe Pro Val Ser Asp Tyr Glu
130 135 140

Lys Glu Ile Glu Ile Met Lys Thr Val Thr Lys Glu Glu Tyr Leu Ala
145 150 155 160

Ser Leu Arg Arg Arg Ser Ser Gly Phe Ser Arg Gly Val Ser Lys Tyr
165 170 175

pctau2012001598-seql

Arg Gly Val Ala Arg His His His Asn Gly Arg Trp Glu Ala Arg Ile
180 185 190

Gly Arg Val Phe Gly Asn Lys Tyr Leu Tyr Leu Gly Thr Tyr Ser Thr
195 200 205

Gln Glu Glu Ala Ala Arg Ala Tyr Asp Ile Ala Ala Ile Glu Tyr Arg
210 215 220

Gly Ile Asn Ala Val Thr Asn Phe Asp Leu Ser Thr Tyr Ile Arg Trp
225 230 235 240

Leu Arg Pro Gly Thr His Pro Thr Ala Ser His Asp Gln Lys Pro Ser
245 250 255

Thr Asp Ala Gln Leu Phe Ala Thr Ser Asn Ser Met Gln Thr Arg Gly
260 265 270

Asn Ile Glu Val Ser Asn Ser Asn Met His Ser Phe Pro Ser Gly Glu
275 280 285

Leu Asp Ser Thr Lys Lys Arg Asp Phe Ser Lys Tyr Met Asn Pro Leu
290 295 300

Ser Pro Cys Asn Lys Pro Ser Ser Pro Thr Ala Leu Gly Leu Leu Leu
305 310 315 320

Lys Ser Ser Val Phe Arg Glu Leu Met Gln Arg Asn Leu Asn Ser Ser
325 330 335

Ser Glu Glu Ala Asp Val Glu Leu Lys Tyr Pro Gln Glu Gly Asn Asp
340 345 350

Gly Val Gly Gly Ile Tyr Asp Asn Asp Asn Thr Ser Asn Ser Tyr Phe
355 360 365

Cys Ser Ser Asn Ile Ser Arg Leu Pro Asn Leu Glu Ser Ser Glu Glu
370 375 380

Cys Pro Leu Pro Met Tyr His Gly Thr Met Gln Ser Leu Trp Asn Ser
385 390 395 400

Ala Phe Asn Met Ser Asn
405

<210> 301
<211> 418
<212> PRT
<213> Populus trichocarpa

<400> 301

Met Glu Met Thr Arg Asn Thr Gly Asp Gln Ile Ser Leu Gly Arg Arg
Page 367

pctau2012001598-seql

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

275

280

285

Asp Pro Ser Leu Met Asp Asn Leu Asn Thr Pro Lys Asn Glu Asp Ile
 290 295 300

Phe His Arg Lys Thr Leu Pro Val Ser Pro Leu Thr Arg Ser Ser Ser
 305 310 315 320

Ser Thr Ala Leu Ser Leu Leu Phe Lys Ser Ser Ile Phe Lys Glu Leu
 325 330 335

Val Glu Lys Asn Leu Asn Thr Thr Ser Glu Glu Ile Glu Glu Asn Asp
 340 345 350

Ser Lys Asn Pro His Asn Gly Asn Asn Asn Ala Gly Glu Ala Phe Tyr
 355 360 365

Asp Gly Leu Ser Pro Ile Pro His Thr Gly Thr Ser Thr Glu Asp Pro
 370 375 380

Phe Leu Cys Ser Glu Gln Gly Glu Thr Asn Thr Leu Pro Pro Tyr Ser
 385 390 395 400

Gly Met Glu Gln Ser Leu Trp Asn Gly Ala Leu Ser Met Pro Ser Arg
 405 410 415

Phe His

<210> 302
 <211> 336
 <212> PRT
 <213> Vitis vinifera

<400> 302

Met Glu Met Thr Thr Val Lys Ser Glu Leu Gly Leu Glu Arg Gly Arg
 1 5 10 15

Leu Cys Thr Ala Glu Thr Asp Ala Leu Glu Val Thr Lys Cys Val Lys
 20 25 30

Arg Arg Arg Arg Asp Pro Ser Ala Val Thr Pro Gly Cys Ser Lys Gln
 35 40 45

Gly Glu Gln Gln Lys Gln Val Leu Leu Gln Ala Gly Gln Ser Ile Thr
 50 55 60

Ala Ile Ala Thr Thr Met Lys Arg Ser Ser Arg Phe Arg Gly Val Ser
 65 70 75 80

Arg His Arg Trp Thr Gly Arg Phe Glu Ala His Leu Trp Asp Lys Gly
 85 90 95

pctau2012001598-seql

Ser Trp Asn Val Thr Gln Arg Lys Lys Gly Lys Gln Val Tyr Leu Gly
100 105 110

Ala Tyr Asp Glu Glu Glu Ser Ala Ala Arg Ala Tyr Asp Leu Ala Ala
115 120 125

Leu Lys Tyr Trp Gly Pro Ser Thr Phe Thr Asn Phe Pro Val Ser Asp
130 135 140

Tyr Glu Lys Glu Ile Glu Ile Met Gln Gly Leu Thr Lys Glu Glu Tyr
145 150 155 160

Leu Ala Ser Leu Arg Arg Arg Ser Ser Gly Phe Ser Arg Gly Val Ser
165 170 175

Lys Tyr Arg Gly Val Ala Arg His His His Asn Gly Arg Trp Glu Ala
180 185 190

Arg Ile Gly Arg Val Phe Gly Asn Lys Tyr Leu Tyr Leu Gly Thr Tyr
195 200 205

Ser Thr Gln Glu Glu Ala Ala His Ala Tyr Asp Ile Ala Ala Ile Glu
210 215 220

Tyr Arg Gly Ile Asn Ala Val Thr Asn Phe Glu Leu Ser Thr Tyr Val
225 230 235 240

Arg Trp Leu Arg Pro Arg Ala Thr Ala Leu Thr Pro Gln Glu Pro Arg
245 250 255

Ser Asn Ser Ile Met Gln Ala Ser Ser Asn Cys Leu Pro Asn Glu Glu
260 265 270

Val Glu Leu Ser Phe Leu Ser Pro Asn Pro Phe Thr Val Asp Asp Leu
275 280 285

Ala Thr Pro Leu Lys Gln Glu Lys Phe Gln Arg Glu Val Ser Ile Ser
290 295 300

Pro Cys Thr Lys Ser Ser Ser Pro Thr Ala Leu Ser Leu Leu His Arg
305 310 315 320

Ser Ser Val Phe Arg Gln Leu Val Glu Lys Asn Ser Asn Ser Ile Glu
325 330 335

<210> 303
<211> 389
<212> PRT
<213> Glycine max

<400> 303

pctau2012001598-seql

```

Met Ala Met Met Lys Glu Asn Ile Ile Glu Val Ser Leu Gly Arg Arg
1      5      10      15

Gln Met Ser Met Thr Glu Gly Glu Phe Gln Gly Thr Arg Ser Val Lys
20      25      30

Arg Arg Arg Arg Glu Val Ala Ala Ala Ala Gly Ser Gly Asp Asp Asn
35      40      45

His Gln Gln Gln Leu Pro Gln Gln Glu Val Gly Glu Asn Thr Thr Val
50      55      60

Asn Thr Thr Lys Arg Ser Ser Arg Phe Arg Gly Val Ser Arg His Arg
65      70      75      80

Trp Thr Gly Arg Tyr Glu Ala His Leu Trp Asp Lys Leu Ser Trp Asn
85      90      95

Ile Thr Gln Lys Lys Lys Gly Lys Gln Gly Ala Tyr Asp Glu Glu Glu
100     105     110

Ser Ala Ala Arg Ala Tyr Asp Leu Ala Ala Leu Lys Tyr Trp Gly Thr
115     120     125

Ser Thr Phe Thr Asn Phe Pro Ile Ser Asp Tyr Glu Lys Glu Ile Gln
130     135     140

Ile Met Gln Thr Met Thr Lys Glu Glu Tyr Leu Ala Thr Leu Arg Arg
145     150     155     160

Lys Ser Ser Gly Phe Ser Arg Gly Val Ser Lys Tyr Arg Gly Val Ala
165     170     175

Arg His His His Asn Gly Arg Trp Glu Ala Arg Ile Gly Arg Val Phe
180     185     190

Gly Asn Lys Tyr Leu Tyr Leu Gly Thr Tyr Ser Thr Gln Glu Glu Ala
195     200     205

Ala Arg Ala Tyr Asp Ile Ala Ala Ile Glu Tyr Arg Gly Ile His Ala
210     215     220

Val Thr Asn Phe Asp Leu Ser Thr Tyr Ile Lys Trp Leu Lys Pro Ser
225     230     235     240

Gly Gly Gly Thr Leu Glu Ala Asn Leu Glu Ser His Ala Ala Leu Glu
245     250     255

His Gln Lys Val Ala Ser Pro Ser Asn Tyr Ala Leu Thr Glu Glu Ser
260     265     270

```

pctau2012001598-seql

Lys Ser Leu Ala Leu His Asn Ser Phe Phe Ser Pro Tyr Ser Leu Asp
275 280 285

Ser Pro Val Lys His Glu Arg Phe Gly Asn Lys Thr Tyr Gln Phe Ser
290 295 300

Ser Asn Lys Ser Ser Ser Pro Thr Ala Leu Gly Leu Leu Leu Arg Ser
305 310 315 320

Ser Leu Phe Arg Glu Leu Val Glu Lys Asn Ser Asn Val Ser Gly Glu
325 330 335

Glu Asp Asp Gly Glu Ala Thr Lys Asp Gln Gln Thr Gln Ile Ala Thr
340 345 350

Asp Asp Asp Leu Gly Gly Ile Phe Phe Asp Ser Phe Ser Asp Ile Pro
355 360 365

Phe Val Cys Asp Pro Asn Arg Tyr Asp Leu Glu Leu Gln Glu Arg Asp
370 375 380

Leu His Ser Ile Phe
385

<210> 304
<211> 389
<212> PRT
<213> Glycine max

<400> 304

Met Val Met Met Lys Glu Asn Ile Ile Glu Glu Lys Leu Gly Arg Ser
1 5 10 15

Gln Met Ser Met Val Glu Gly Glu Phe Gln Gly Thr Trp Gly Val Lys
20 25 30

Arg Arg Arg Arg Glu Val Ala Ala Ala Ser Ser Gly Asp Asp Asn
35 40 45

His His Gln Gln Leu Pro Gln Gln Glu Val Gly Glu Asn Ser Ser Ile
50 55 60

Ser Thr Thr Lys Arg Ser Ser Arg Phe Arg Gly Val Ser Arg His Arg
65 70 75 80

Trp Thr Gly Arg Tyr Glu Ala His Leu Trp Asp Lys Leu Ser Trp Asn
85 90 95

Ile Thr Gln Lys Lys Lys Gly Lys Gln Gly Ala Tyr Asp Glu Glu Glu
100 105 110

pctau2012001598-seql

Ser Ala Ala Arg Ala Tyr Asp Leu Ala Ala Leu Lys Tyr Trp Gly Asn
115 120 125
Ser Thr Phe Thr Asn Phe Pro Ile Ser Asp Tyr Glu Lys Glu Ile Glu
130 135 140
Ile Met Gln Thr Met Thr Lys Glu Glu Tyr Leu Ala Thr Leu Arg Arg
145 150 155 160
Lys Ser Ser Gly Phe Ser Arg Gly Val Ser Lys Tyr Arg Gly Val Ala
165 170 175
Arg His His His Asn Gly Arg Trp Glu Ala Arg Ile Gly Arg Val Phe
180 185 190
Gly Asn Lys Tyr Leu Tyr Leu Gly Thr Tyr Ser Thr Gln Glu Glu Ala
195 200 205
Ala Arg Ala Tyr Asp Ile Ala Ala Ile Glu Tyr Arg Gly Ile His Ala
210 215 220
Val Thr Asn Phe Asp Leu Ser Thr Tyr Ile Lys Trp Leu Lys Pro Ser
225 230 235 240
Gly Gly Gly Thr Pro Glu Glu Asn Leu Glu Ser His Ala Val Leu Glu
245 250 255
His Gln Lys Leu Ala Ser Pro Ser Asn Tyr Ala Leu Thr Glu Glu Ser
260 265 270
Lys Ser Leu Val Leu Pro Asn Ser Phe Ile Ser Pro Asp Ser Leu Asp
275 280 285
Ser Pro Val Lys His Glu Ser Phe Gly Asn Lys Thr Tyr Gln Phe Ser
290 295 300
Arg Asn Lys Ser Ser Ser Pro Thr Ala Leu Gly Leu Leu Leu Arg Ser
305 310 315 320
Ser Leu Phe Arg Glu Leu Val Glu Lys Asn Ser Asn Val Ser Gly Glu
325 330 335
Glu Ala Asp Gly Glu Val Thr Lys Asp Gln Gln Pro Gln Leu Ala Ser
340 345 350
Asp Asp Asp Leu Asp Gly Ile Phe Phe Asp Ser Phe Gly Asp Ile Pro
355 360 365
Phe Val Cys Asp Pro Thr Arg Tyr Asn Leu Glu Leu Gln Glu Arg Asp
370 375 380

pctau2012001598-seql

Leu His Ser Ile Phe
385

<210> 305
<211> 392
<212> PRT
<213> Medicago truncatula

<400> 305

Met Ala Met Leu Ile Glu Asn Glu Val Met Cys Leu Gly Lys Ser Gln
1 5 10 15

Arg Ser Met Asp Gly Lys Glu Val Lys Gly Ala Arg Arg Val Lys Arg
20 25 30

Gln Arg Arg Asp Ala Ile Val Pro Lys Ile Gly Asp Asp Ala Asn Lys
35 40 45

Met Ala Gln Lys Gln Val Gly Glu Asn Ser Thr Thr Asn Thr Ser Lys
50 55 60

Arg Ser Ser Arg Phe Arg Gly Val Ser Arg His Arg Trp Thr Gly Arg
65 70 75 80

Phe Glu Ala His Leu Trp Asp Lys Leu Ser Trp Asn Thr Thr Gln Lys
85 90 95

Lys Lys Gly Lys Gln Gly Ala Tyr Asp Glu Glu Glu Ser Ala Ala Arg
100 105 110

Ala Tyr Asp Leu Ala Ala Leu Lys Tyr Trp Gly Thr Ser Thr Phe Thr
115 120 125

Asn Phe Pro Ile Ser Asp Tyr Asp Lys Glu Ile Glu Ile Met Asn Thr
130 135 140

Met Thr Lys Glu Glu Tyr Leu Ala Thr Leu Arg Arg Lys Ser Ser Gly
145 150 155 160

Phe Ser Arg Gly Val Ser Lys Tyr Arg Gly Val Ala Arg His His His
165 170 175

Asn Gly Arg Trp Glu Ala Arg Ile Gly Arg Val Phe Gly Asn Lys Tyr
180 185 190

Leu Tyr Leu Gly Thr Tyr Ser Thr Gln Glu Glu Ala Ala Arg Ala Tyr
195 200 205

Asp Ile Ala Ala Ile Glu Tyr Arg Gly Ile His Ala Val Thr Asn Phe
210 215 220

Glu Leu Ser Ser Tyr Ile Lys Trp Leu Lys Pro Glu Thr Thr Thr Glu
Page 374

pctau2012001598-seq1

225 230 235 240

Glu Asn His Glu Ser Gln Ile Leu Gln Lys Glu Ser Arg Thr Leu Ala
245 250 255

Pro Pro Asn Asn Ser Thr Leu Leu Gln Glu Ser Lys Leu Leu Ala Leu
260 265 270

Gln Lys Ser Phe Phe Ile Pro Asn Asp Leu Asn Ser Thr Glu Lys Gln
275 280 285

Glu Ser Ser Phe Glu Asn Lys Asn Tyr His Phe Leu Ser Asn Lys Ser
290 295 300

Thr Ser Pro Thr Ala Leu Ser Leu Leu Leu Arg Ser Ser Leu Phe Arg
305 310 315 320

Glu Leu Leu Glu Lys Asn Ser Asn Val Ser Glu Asp Glu Val Thr Lys
325 330 335

Glu Gln Gln Gln Gln Gln Ile Thr Ser Asp Asp Glu Leu Gly Gly Ile
340 345 350

Phe Tyr Asp Gly Ile Asp Asn Ile Ser Phe Asp Phe Asp Pro Asn Ser
355 360 365

Cys Asn Ile Glu Leu Gln Glu Arg Asp Leu His Ser Ile Ser Cys Leu
370 375 380

Tyr Gln Tyr Leu Asn Phe Gly Gln
385 390

<210> 306
<211> 386
<212> PRT
<213> Populus trichocarpa

<400> 306

Met Met Met Ile Lys Asn Glu Glu Asn Pro Gly Arg Arg Arg Gly Cys
1 5 10 15

Ile Ala Asp Ser Glu Ala Gln Val Ala Arg Cys Val Lys Arg Arg Arg
20 25 30

Arg Asp Pro Ala Ile Val Ala Leu Gly Ser Asp Asp Asn Gln Ser Gln
35 40 45

Gln Gln Met Pro Gln Lys Gln Thr Asp Gln Thr Ser Ala Ala Thr Thr
50 55 60

Val Lys Arg Ser Ser Arg Phe Arg Gly Val Ser Arg His Arg Trp Thr
65 70 75 80

pctau2012001598-seql

Gly Arg Phe Glu Ala His Leu Trp Asp Lys Leu Ser Trp Asn Val Thr
 85 90 95
 Gln Lys Lys Lys Gly Lys Gln Gly Ala Tyr Asp Glu Glu Glu Ser Ala
 100 105 110
 Ala Arg Ala Tyr Asp Leu Ala Ala Leu Lys Tyr Trp Gly Thr Ser Thr
 115 120 125
 Phe Thr Asn Phe Pro Ile Ser Asp Tyr Glu Lys Glu Ile Glu Ile Met
 130 135 140
 Gln Thr Val Thr Lys Glu Glu Tyr Leu Ala Ser Leu Arg Arg Lys Ser
 145 150 155 160
 Ser Gly Phe Ser Arg Gly Val Ser Lys Tyr Arg Gly Val Ala Arg His
 165 170 175
 His His Asn Gly Arg Trp Glu Ala Arg Ile Gly Arg Val Phe Gly Asn
 180 185 190
 Lys Tyr Leu Tyr Leu Gly Thr Tyr Ser Thr Gln Glu Glu Ala Ala Arg
 195 200 205
 Ala Tyr Asp Ile Ala Ala Ile Glu Tyr Arg Gly Ile Asn Ala Val Thr
 210 215 220
 Asn Phe Asp Leu Ser Thr Tyr Ile Arg Trp Ile Lys Pro Gly Val Ala
 225 230 235 240
 Ala Gln Ala Ala Ala Asn Glu Leu Gln Thr Val Thr Asp Pro Gln Thr
 245 250 255
 Ala Ala Thr Leu Thr Asp Thr Tyr Thr Pro Arg Glu Glu Thr Lys Pro
 260 265 270
 Ser Leu Phe Leu Pro Asn Gln Phe Thr Ala Asp Tyr Leu Asn Ser Pro
 275 280 285
 Pro Lys Leu Asp Ala Phe Gln Asn Asn Ile Phe Val Asp Ser Ser Asn
 290 295 300
 Lys Thr Ser Ser Pro Thr Ala Leu Ser Leu Leu Leu Arg Ser Ser Val
 305 310 315 320
 Phe Arg Glu Leu Val Glu Lys Asn Ser Asn Val Cys Glu Glu Glu Thr
 325 330 335
 Asp Gly Asn Glu Ile Lys Asn Gln Pro Met Ala Gly Ser Asp Asp Glu
 340 345 350

pctau2012001598-seql

Tyr Gly Gly Ile Phe Tyr Asp Gly Ile Gly Asp Ile Pro Phe Val Tyr
355 360 365

Ser Ser Asn Lys Tyr Ser Leu Gly Leu Glu Glu Arg Glu Leu Gln Phe
370 375 380

Val Leu
385

<210> 307
<211> 372
<212> PRT
<213> Ricinus communis

<400> 307

Met Glu Met Met Met Val Lys Asn Glu Glu Ile Ser Gly Arg Arg Arg
1 5 10 15

Ala Ser Val Thr Glu Ser Glu Ala Tyr Val Ala Arg Cys Val Lys Arg
20 25 30

Arg Arg Arg Asp Ala Ala Val Val Thr Val Gly Gly Asp Asp Ser Gln
35 40 45

Ser His Gln Gln Gln Gln Gln Gln Gln Pro Glu Gln Gln Ala His Gln
50 55 60

Ile Ser Ala Ala Thr Thr Val Lys Arg Ser Ser Arg Tyr Arg Gly Val
65 70 75 80

Ser Arg His Arg Trp Thr Gly Arg Phe Glu Ala His Leu Trp Asp Lys
85 90 95

Leu Ser Trp Asn Val Thr Gln Lys Lys Lys Gly Lys Gln Gly Ala Tyr
100 105 110

Asp Glu Glu Glu Ser Ala Ala Arg Ala Tyr Asp Leu Ala Ala Leu Lys
115 120 125

Tyr Trp Gly Thr Ser Thr Phe Thr Asn Phe Pro Ile Ser Asp Tyr Glu
130 135 140

Lys Glu Ile Glu Ile Met Gln Thr Val Thr Lys Glu Glu Tyr Leu Ala
145 150 155 160

Ser Leu Arg Arg Lys Ser Ser Gly Phe Ser Arg Gly Val Ser Lys Tyr
165 170 175

Arg Gly Val Ala Arg His His His Asn Gly Arg Trp Glu Ala Arg Ile
180 185 190

pctau2012001598-seql

Gly Arg Val Phe Gly Asn Lys Tyr Leu Tyr Leu Gly Thr Tyr Ser Thr
195 200 205

Gln Glu Glu Ala Ala Arg Ala Tyr Asp Ile Ala Ala Ile Glu Tyr Arg
210 215 220

Gly Ile Asn Ala Val Thr Asn Phe Asp Leu Ser Thr Tyr Ile Arg Trp
225 230 235 240

Leu Lys Pro Glu Val Ala Ala Gln Val Ala Ala Asn Glu Pro Gln Thr
245 250 255

Val Ala Glu Ser Arg Met Leu Pro Ser Ile Asn Asn Arg Ile Ala Arg
260 265 270

Glu Glu Ser Lys Pro Ser Phe Phe Ser Ala Thr Pro Phe Ser Leu Asp
275 280 285

Cys Trp Ser Tyr Pro Arg Lys Gln Glu Glu Phe Gln Asn Arg Thr Pro
290 295 300

Ile Thr Pro Cys Ser Lys Thr Ser Ser Pro Thr Ala Leu Ser Leu Leu
305 310 315 320

Leu Arg Ser Ser Ile Phe Arg Glu Leu Val Glu Lys Asn Ser Asn Val
325 330 335

Ser Glu Asp Glu Asn Glu Gly Glu Glu Thr Lys Asn Gln Ser Gln Ile
340 345 350

Gly Ser Asp Asp Glu Phe Gly Gly Leu Phe Tyr Glu Arg Ile Gly Asp
355 360 365

Ile Pro Phe Ile
370

<210> 308
<211> 404
<212> PRT
<213> Vitis vinifera

<400> 308

Met Glu Met Met Arg Val Lys Ser Glu Glu Asn Leu Gly Arg Arg Arg
1 5 10 15

Met Cys Val Ala Asp Ala Glu Ala Gln Gly Thr Arg Cys Val Lys Arg
20 25 30

Arg Arg Arg Asp Pro Ala Ile Val Thr Leu Gly Cys Asp Asp Gln Ser
35 40 45

pctau2012001598-seq1

pctau2012001598-seql

Phe Arg Glu Leu Val Glu Lys Asn Ser Asn Ala Pro Glu Asp Glu Thr
325 330 335

Asp Ala Glu Asp Thr Lys Asn Gln Gln Gln Val Gly Ser Asp Asp Glu
340 345 350

Tyr Gly Ile Phe Tyr Asp Gly Ile Gly Asp Ile Pro Phe Val Cys Pro
355 360 365

Ser Asn Gly Asp Arg Asn Glu Leu Gln Glu Arg Leu Pro Leu Pro Phe
370 375 380

Thr Ile Ser Gln Gly Asn Pro Tyr Gly Thr Ala Val Leu Thr Ser Met
385 390 395 400

Gln Ser Ile Asn

<210> 309

<211> 378

<212> PRT

<213> Brachypodium distachyon

<400> 309

Met Ala Lys Gln Arg Thr Asp Ser Ala Gly Thr Asp Ala Ala Ala Val
1 5 10 15

Gln Leu Thr Lys Pro Lys Arg Thr Arg Lys Ser Val Pro Arg Arg Glu
20 25 30

Ser Pro Ser Arg Arg Thr Ser Ala Tyr Arg Gly Val Thr Arg His Arg
35 40 45

Trp Thr Gly Arg Phe Glu Ala His Leu Trp Asp Lys Asn Thr Trp Thr
50 55 60

Gln Ser Gln Arg Lys Lys Lys Gly Arg Gln Val Tyr Leu Gly Ala Tyr
65 70 75 80

Gly Gly Glu Glu Ala Ala Ala Arg Ala Tyr Asp Leu Ala Ala Leu Lys
85 90 95

Tyr Trp Gly Arg Asp Thr Val Leu Asn Phe Pro Leu Ser Asn Tyr Asp
100 105 110

Glu Glu Trp Lys Glu Met Glu Gly Gln Ser Arg Glu Glu Tyr Ile Gly
115 120 125

Ser Leu Arg Arg Lys Ser Thr Gly Phe Ser Arg Gly Val Ser Lys Tyr
130 135 140

Arg Gly Val Ala Arg His His His Asn Gly Lys Trp Glu Ala Arg Ile
Page 380

pctau2012001598-seql

145 150 155 160

Gly Arg Val Tyr Gly Asn Lys Tyr Leu Tyr Leu Gly Thr Tyr Gly Thr
165 170 175

Gln Glu Glu Ala Ala Met Ala Tyr Asp Ile Ala Ala Ile Glu His Arg
180 185 190

Gly Leu Asn Ala Val Thr Asn Phe Asp Val Ser Arg Tyr Ile Asp Trp
195 200 205

His Arg Arg Leu Cys Arg Asp Leu Gly Asp Asn Ile Ile Thr Pro Leu
210 215 220

Thr Asn Pro Thr Val Asp Leu Glu Glu Ala Met Ala Gly Asp Asp Asp
225 230 235 240

Asp Gly Gln Phe Leu Leu Pro Ser Gln Ala Thr Thr Pro Pro Ser Thr
245 250 255

Ser Ser Ala Leu Gly Leu Leu Leu Leu Ser Pro Arg Leu Lys Glu Val
260 265 270

Ile Glu Gly Ser Gly Ala Ala Ser Ala Met Ala Ala Ser Thr Ser Glu
275 280 285

Ser Ser Ala Ala Gly Ser Pro Pro Pro Ser Trp Ser Ser Ser Ser Cys
290 295 300

Ser Pro Ser Pro Pro Ser Pro Ser His Ser Pro Pro Glu Thr Gln Gln
305 310 315 320

Lys Gln Gln Gln Gln Glu Tyr Gly Ala Ser Ala Ala Ala Ala Arg Cys
325 330 335

Ser Phe Pro Asp Asp Val Gln Thr Tyr Phe Gly Cys Glu Asp Gly Cys
340 345 350

Ala Glu Val Asp Thr Phe Leu Phe Gly Asp Leu Ser Ala Tyr Ala Ala
355 360 365

Pro Met Phe Gln Phe Glu Leu Leu Asp Val
370 375

<210> 310
<211> 416
<212> PRT
<213> Oryza sativa

<400> 310

Met Ala Lys Arg Arg Ser Asn Gly Glu Thr Ala Ala Ala Ser Ser Asp
1 5 10 15

pctau2012001598-seql

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

pctau2012001598-seql

Ala Ala Ser Thr Ser Pro Ile Gly Leu Leu Leu Arg Ser Pro Lys Phe
290 295 300

Lys Glu Ile Ile Glu Gln Ala Ala Ala Ala Ala Ala Ser Ser Ser Gly
305 310 315 320

Ser Ser Ser Ser Ser Ser Thr Asp Ser Pro Ser Ser Ser Ser Ser Ser
325 330 335

Ser Leu Ser Pro Ser Pro Leu Pro Ser Pro Pro Pro Gln Gln Gln Pro
340 345 350

Thr Val Pro Lys Asp Asp Gln Tyr Asn Val Asp Met Ser Ser Val Ala
355 360 365

Ala Ala Arg Cys Ser Phe Pro Asp Asp Val Gln Thr Tyr Phe Gly Leu
370 375 380

Asp Asp Asp Gly Phe Gly Tyr Pro Glu Val Asp Thr Phe Leu Phe Gly
385 390 395 400

Asp Leu Gly Ala Tyr Ala Ala Pro Met Phe Gln Phe Glu Leu Asp Val
405 410 415

<210> 311
<211> 440
<212> PRT
<213> Sorghum bicolor

<400> 311

Met Ala Arg Pro Arg Lys Asn Ala Gly Thr Asp Glu Asp Asn Pro Asn
1 5 10 15

Ala Ala Thr Gly Val Ser Val Thr Gly Lys Pro Pro Lys Leu Lys Arg
20 25 30

Val Arg Arg Lys Gly Glu Pro Arg Glu Ser Ser Thr Pro Ser Gln Arg
35 40 45

Ser Ser Ala Tyr Arg Gly Val Thr Arg His Arg Trp Thr Gly Arg Phe
50 55 60

Glu Ala His Leu Trp Asp Lys Asp Ala Arg Asn Gly Ser Arg Asn Lys
65 70 75 80

Lys Gly Lys Gln Gly Ala Tyr Asp Asp Glu Glu Ala Ala Ala Arg Ala
85 90 95

His Asp Leu Ala Ala Leu Lys Tyr Trp Gly Pro Ala Thr Val Leu Asn
100 105 110

pctau2012001598-seql

Phe Pro Leu Cys Gly Tyr Asp Glu Glu Leu Arg Glu Met Glu Ala Gln
115 120 125
Pro Arg Glu Glu Tyr Ile Gly Ser Leu Arg Arg Arg Ser Ser Gly Phe
130 135 140
Ser Arg Gly Val Ser Lys Tyr Arg Gly Val Ala Arg His His His Asn
145 150 155 160
Gly Arg Trp Glu Ala Arg Ile Gly Arg Val Leu Gly Asn Lys Tyr Leu
165 170 175
Tyr Leu Gly Thr Phe Ala Thr Gln Glu Glu Ala Ala Val Ala Tyr Asp
180 185 190
Ile Ala Ala Ile Glu His Arg Gly Leu Asn Ala Val Thr Asn Phe Asp
195 200 205
Ile Ser His Tyr Val Asn His Trp His Arg His Cys His Gly Pro Ser
210 215 220
Asp Asp Ser Leu Gly Val Val Val Asp Asp Val Ala Ala Phe Gln Leu
225 230 235 240
Pro Asp Asp Leu Pro Glu Cys Pro Ala Ala Ala Ile Gly Val Glu Glu
245 250 255
Thr Thr Gly Gly Asp Ala Glu Phe His Asn Gly Glu Glu Gly Tyr Leu
260 265 270
Gln His His Thr Ser Gly Pro Phe Gly Ala Gln Gln Gln Leu Pro Asp
275 280 285
Glu Thr Gly Ala Leu Ala Ala His Gln Met Ala Pro Asn Ser Ser Ala
290 295 300
Leu Asp Met Val Leu Gln Ser Pro Lys Phe Lys Glu Leu Met Glu Gln
305 310 315 320
Val Ser Ala Ala Ala Ala Ala Val Ala Ser Glu Ser Ser Ile Gly Gly
325 330 335
Ser Met Ser Ser Ser Ser Pro Ser Pro Ser Leu Ser Ser Phe Ser Pro
340 345 350
Ser Pro Leu Gln Leu Pro Ser Pro Ser Ser Leu Ser Ser Phe Ser Pro
355 360 365
Ser Ser Pro Leu Gln Gln Pro Ser Pro Pro Leu Gln Gln Pro Glu Phe
370 375 380

pctau2012001598-seql

Val Glu Gly Ala Pro Ala Ala Arg Cys Ser Phe Pro Asp Asp Val Gln
385 390 395 400

Thr Phe Phe Asp Phe Glu Asn Glu Ser Asp Met Ser Phe Met Tyr Ala
405 410 415

Glu Val Asp Thr Phe Leu Phe Gly Asp Leu Gly Ala Tyr Ala Ala Pro
420 425 430

Ile Phe His Phe Asp Leu Asp Val
435 440

<210> 312
<211> 408
<212> PRT
<213> Zea mays

<400> 312

Met Ala Arg Pro Arg Lys Asn Gly Gly Thr Asp Glu Asp Asp Ala Asn
1 5 10 15

Ala Ala Thr Gly Ala Thr Gly Lys Pro Lys Lys Leu Met Lys Arg Ala
20 25 30

Arg Arg Lys Ser Glu Ser Pro Ser Pro Arg Ser Ser Ala Tyr Arg Gly
35 40 45

Val Thr Arg His Arg Trp Thr Gly Arg Phe Glu Ala His Leu Trp Asp
50 55 60

Lys Asp Ala Arg Asn Gly Ser Arg Ser Lys Lys Gly Lys Gln Val Tyr
65 70 75 80

Leu Gly Ala Tyr Asp Asp Glu Asp Ala Ala Arg Ala His Asp Leu
85 90 95

Ala Ala Leu Lys Tyr Trp Gly Pro Ala Gly Thr Val Leu Asn Phe Pro
100 105 110

Leu Ser Gly Tyr Asp Glu Glu Arg Arg Glu Met Glu Gly Gln Pro Arg
115 120 125

Glu Glu Tyr Val Ala Ser Leu Arg Arg Arg Ser Ser Gly Phe Ala Arg
130 135 140

Gly Val Ser Lys Tyr Arg Gly Val Ala Arg His His His Asn Gly Arg
145 150 155 160

Trp Glu Ala Arg Ile Gly Arg Val Leu Gly Asn Lys Tyr Leu Tyr Leu
165 170 175

pctau2012001598-seql

Gly Thr Tyr Ala Thr Gln Glu Glu Ala Ala Val Ala Tyr Asp Met Ala
180 185 190

Ala Ile Glu His Arg Gly Phe Asn Ala Val Thr Asn Phe Asp Ile Ser
195 200 205

His Tyr Ile Asn His Trp His Arg His Cys His Gly Pro Cys Asp Gly
210 215 220

Ser Leu Gly Ala Met Asp Val Ala Pro Asn Val Ser Leu Glu Leu Asp
225 230 235 240

Leu Leu Glu Cys Pro Ala Thr Val Gly Leu Gly Leu Glu Glu Thr Thr
245 250 255

Gly Asp Asp Glu Phe His Asn Arg Glu Asp Tyr Leu Gly His Leu Phe
260 265 270

Gly Val Gln Gln Leu Pro Asp Glu Met Gly Pro Pro Ala His Gln Met
275 280 285

Ala Pro Ala Ser Ser Ala Leu Asp Leu Val Leu Gln Ser Pro Arg Phe
290 295 300

Lys Glu Leu Met Gln Gln Val Ser Ala Ala Gly Ala Ser Glu Thr Asn
305 310 315 320

Gly Gly Ser Met Arg Ser Ser Pro Ser Thr Ser Leu Cys Ser Phe Ser
325 330 335

Pro Ser Pro Leu Glu Leu Pro Ser Pro Pro Leu Gln Gln Pro Thr Glu
340 345 350

Phe Ile Asp Gly Ala Pro Pro Arg Cys Ser Phe Pro Asp Asp Val Gln
355 360 365

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

Glu Val Asp Thr Phe Leu Phe Gly Asp Leu Gly Ala Tyr Ala Pro Pro
385 390 395 400

Met Phe Asp Phe Asp Leu Tyr Glu
405

<210> 313

<211> 304

<212> PRT

<213> Arabidopsis lyrata

<400> 313

Met Ala Lys Val Ser Arg Arg Ser Lys Lys Thr Ile Val Glu Asp Glu
Page 386

pctau2012001598-seql

```

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

```

275

Asn Ser Phe Ile Asn Pro Gly Phe Tyr Asn Glu Phe Asp Tyr Gly Pro
290 295 300

<210> 314
<211> 308
<212> PRT
<213> Arabidopsis thaliana

<400> 314

Met Ala Lys Val Ser Gly Arg Ser Lys Lys Thr Ile Val Asp Asp Glu
1 5 10 15

Ile Ser Asp Lys Thr Ala Ser Ala Ser Glu Ser Ala Ser Ile Ala Leu
20 25 30

Thr Ser Lys Arg Lys Arg Lys Ser Pro Pro Arg Asn Ala Pro Leu Gln
35 40 45

Arg Ser Ser Pro Tyr Arg Gly Val Thr Arg Trp Thr Gly Arg Tyr Glu
50 55 60

Ala His Leu Trp Asp Lys Asn Ser Trp Asn Asp Thr Gln Thr Lys Lys
65 70 75 80

Gly Arg Gln Gly Ala Tyr Asp Glu Glu Glu Ala Ala Ala Arg Ala Tyr
85 90 95

Asp Leu Ala Ala Leu Lys Tyr Trp Gly Arg Asp Thr Leu Leu Asn Phe
100 105 110

Pro Leu Pro Ser Tyr Asp Glu Asp Val Lys Glu Met Glu Gly Gln Ser
115 120 125

Lys Glu Glu Tyr Ile Gly Ser Leu Arg Arg Lys Ser Ser Gly Phe Ser
130 135 140

Arg Gly Val Ser Lys Tyr Arg Gly Val Ala Arg His His His Asn Gly
145 150 155 160

Arg Trp Glu Ala Arg Ile Gly Arg Val Phe Gly Asn Lys Tyr Leu Tyr
165 170 175

Leu Gly Thr Tyr Ala Thr Gln Glu Glu Ala Ala Ile Ala Tyr Asp Ile
180 185 190

Ala Ala Ile Glu Tyr Arg Gly Leu Asn Ala Val Thr Asn Phe Asp Val
195 200 205

Ser Arg Tyr Leu Asn Pro Asn Ala Ala Ala Asp Lys Ala Asp Ser Asp
210 215 220

pctau2012001598-seql

Ser Lys Pro Ile Arg Ser Pro Ser Arg Glu Pro Glu Ser Ser Asp Asp
225 230 235 240

Asn Lys Ser Pro Lys Ser Glu Glu Val Ile Glu Pro Ser Thr Ser Pro
245 250 255

Glu Val Ile Pro Thr Arg Arg Ser Phe Pro Asp Asp Ile Gln Thr Tyr
260 265 270

Phe Gly Cys Gln Asp Ser Gly Lys Leu Ala Thr Glu Glu Asp Val Ile
275 280 285

Phe Asp Cys Phe Asn Ser Tyr Ile Asn Pro Gly Phe Tyr Asn Glu Phe
290 295 300

Asp Tyr Gly Pro
305

<210> 315
<211> 303
<212> PRT
<213> Arabidopsis thaliana

<400> 315

Met Ala Lys Val Ser Gly Arg Ser Lys Lys Thr Ile Val Asp Asp Glu
1 5 10 15

Ile Ser Asp Lys Thr Ala Ser Ala Ser Glu Ser Ala Ser Ile Ala Leu
20 25 30

Thr Ser Lys Arg Lys Arg Lys Ser Pro Pro Arg Asn Ala Pro Leu Gln
35 40 45

Arg Ser Ser Pro Tyr Arg Gly Val Thr Arg His Arg Trp Thr Gly Arg
50 55 60

Tyr Glu Ala His Leu Trp Asp Lys Asn Ser Trp Asn Asp Thr Gln Thr
65 70 75 80

Lys Lys Gly Arg Gln Val Tyr Leu Gly Ala Tyr Asp Glu Glu Glu Ala
85 90 95

Ala Ala Arg Ala Tyr Asp Leu Ala Ala Leu Lys Tyr Trp Gly Arg Asp
100 105 110

Thr Leu Leu Asn Phe Pro Leu Pro Ser Tyr Asp Glu Asp Val Lys Glu
115 120 125

Met Glu Gly Gln Ser Lys Glu Glu Tyr Ile Gly Ser Leu Arg Arg Lys
130 135 140

pctau2012001598-seql

Ser Ser Gly Phe Ser Arg Gly Val Ser Lys Tyr Arg Gly Val Ala Arg
145 150 155 160

His His His Asn Gly Arg Trp Glu Ala Arg Ile Gly Arg Val Phe Ala
165 170 175

Thr Gln Glu Glu Ala Ala Ile Ala Tyr Asp Ile Ala Ala Ile Glu Tyr
180 185 190

Arg Gly Leu Asn Ala Val Thr Asn Phe Asp Val Ser Arg Tyr Leu Asn
195 200 205

Pro Asn Ala Ala Ala Asp Lys Ala Asp Ser Asp Ser Lys Pro Ile Arg
210 215 220

Ser Pro Ser Arg Glu Pro Glu Ser Ser Asp Asp Asn Lys Ser Pro Lys
225 230 235 240

Ser Glu Glu Val Ile Glu Pro Ser Thr Ser Pro Glu Val Ile Pro Thr
245 250 255

Arg Arg Ser Phe Pro Asp Asp Ile Gln Thr Tyr Phe Gly Cys Gln Asp
260 265 270

Ser Gly Lys Leu Ala Thr Glu Glu Asp Val Ile Phe Asp Cys Phe Asn
275 280 285

Ser Tyr Ile Asn Pro Gly Phe Tyr Asn Glu Phe Asp Tyr Gly Pro
290 295 300

<210> 316
<211> 332
<212> PRT
<213> Arabidopsis lyrata

<400> 316

Met Glu Glu Ile Thr Arg Lys Ser Lys Lys Thr Ser Val Glu Asn Glu
1 5 10 15

Thr Gly Asp Asp Gln Ser Ala Thr Ser Val Val Val Lys Ala Lys Arg
20 25 30

Lys Arg Arg Ser Gln Pro Arg Asp Ala Pro Pro Gln Arg Ser Ser Val
35 40 45

His Arg Gly Val Thr Arg His Arg Trp Thr Gly Arg Tyr Glu Ala His
50 55 60

Leu Trp Asp Lys Asn Ser Trp Asn Glu Thr Gln Ser Lys Lys Gly Arg
65 70 75 80

pctau2012001598-seql

Gln Gly Ala Tyr Asp₈₅ Glu Glu Asp Ala₉₀ Ala Ala Arg Ala Tyr Asp₉₅ Leu

Ala Ala Leu Lys₁₀₀ Tyr Trp Gly Arg Asp₁₀₅ Thr Ile Leu Asn Phe₁₁₀ Pro Leu

Cys Asn Tyr₁₁₅ Glu Glu Asp Ile Lys₁₂₀ Glu Met Glu Ser Gln₁₂₅ Ser Lys Glu

Glu Tyr₁₃₀ Ile Gly Ser Leu Arg₁₃₅ Arg Lys Ser Ser Gly₁₄₀ Phe Ser Arg Gly

Val Ser Lys Tyr Arg Gly₁₅₀ Val Ala Lys His₁₅₅ His Asn Gly Arg Trp₁₆₀

Glu Ala Arg Ile Gly₁₆₅ Arg Val Phe Gly Asn₁₇₀ Lys Tyr Leu Tyr Leu₁₇₅ Gly

Thr Tyr Ala Thr₁₈₀ Gln Glu Glu Ala₁₈₅ Ile Ala Tyr Asp Ile₁₉₀ Ala Ala

Ile Glu Tyr₁₉₅ Arg Gly Leu Asn Ala₂₀₀ Val Thr Asn Phe Asp₂₀₅ Ile Ser Arg

Tyr Met₂₁₀ Lys Leu Pro Val Pro₂₁₅ Glu Asn Pro Ile Asp₂₂₀ Ala Ala Asn Asn

Leu Leu Glu Ser Pro His₂₃₀ Ser Asp Ser Ser Pro₂₃₅ Phe Ile Asn Pro Thr₂₄₀

His Glu Ser Asp Leu₂₄₅ Ser Gln Ser Gln Ser₂₅₀ Ser Ser Asp Asp Asn₂₅₅ Asp

Asp Arg Lys Thr₂₆₀ Lys Leu Leu Lys Ser₂₆₅ Ser Pro Leu Asn Ala₂₇₀ Glu Glu

Val Ile Gly₂₇₅ Pro Ser Thr Pro₂₈₀ Glu Ile Ala Pro Pro₂₈₅ Arg Arg Ser

Phe Pro₂₉₀ Glu Asp Ile Gln Thr₂₉₅ Tyr Phe Gly Cys Gln₃₀₀ Asn Ser Gly Lys

Leu Thr Thr Glu Glu Asp₃₁₀ Asp Val Ile Phe Gly₃₁₅ Asp Leu Asp Ser Phe₃₂₀

Leu Thr Pro Asp Phe₃₂₅ Tyr Ser Glu Leu Asn Asp Cys₃₃₀

<210> 317

<211> 328

<212> PRT

<213> Thellungiella halophila

pctau2012001598-seql

<400> 317

Met Ala Lys Val Ser Gln Arg Ser Lys Lys Thr Ile Val Asn Asp Glu
1 5 10 15

Ile Ser Asp Lys Lys Ala Val Ala Val Ala Ser Val Ser Ser Ser Ala
20 25 30

Phe Leu Lys Ser Lys Arg Lys Arg Lys Leu Pro Pro Gln Asn Ala Pro
35 40 45

Pro Gln Arg Ser Ser Ser Tyr Arg Gly Val Thr Arg His Arg Trp Thr
50 55 60

Gly Arg Tyr Glu Ala His Leu Trp Asp Lys Asn Cys Trp Asn Glu Thr
65 70 75 80

Gln Thr Lys Lys Gly Arg Gln Val Tyr Leu Gly Ala Tyr Asp Glu Glu
85 90 95

Glu Ala Ala Ala Arg Ala Tyr Asp Leu Ala Ala Leu Lys Tyr Trp Gly
100 105 110

Arg Asp Thr Leu Leu Asn Phe Pro Leu Pro Thr Tyr Glu Glu Asp Val
115 120 125

Lys Glu Met Glu Gly His Ser Arg Glu Glu Tyr Ile Gly Ser Leu Arg
130 135 140

Arg Lys Ser Ser Gly Phe Ser Arg Gly Val Ser Lys Tyr Arg Gly Val
145 150 155 160

Ala Arg His His His Asn Gly Arg Trp Glu Ala Arg Ile Gly Arg Val
165 170 175

Phe Gly Asn Lys Tyr Leu Tyr Leu Gly Thr Tyr Ala Thr Gln Glu Glu
180 185 190

Ala Ala Arg Ala Tyr Asp Ile Ala Ala Ile Glu Tyr Arg Gly Leu Asn
195 200 205

Ala Val Thr Asn Phe Asp Val Ser Arg Tyr Leu Asn Leu Pro Glu Ser
210 215 220

Lys Asn Pro Ser Ala Ala Ala Asn His Leu Pro Asp Glu Ser Asp Tyr
225 230 235 240

Tyr Asp Ser Met Pro Val Arg Asn Pro Asn His Glu Pro Arg Ser Pro
245 250 255

Asp Gly Gln Thr Ser Ser Glu Asp Asn Asp Tyr Thr Lys Thr Glu Glu
Page 392

260

265

270

Thr Leu Asp Pro Glu Ala Ile Pro Ser Arg Arg Ser Phe Pro Asp Asp
 275 280 285

Ile Gln Thr Tyr Phe Gly Cys Gln Asp Ser Gly Lys Leu Ala Thr Glu
 290 295 300

Glu Asp Val Ile Phe Gly Gly Phe Asn Ser Phe Ile Asn Pro Gly Phe
 305 310 315 320

Tyr Asn Asp Phe Asp Tyr Ala Pro
 325

<210> 318

<211> 345

<212> PRT

<213> Arabidopsis thaliana

<400> 318

Met Phe Ile Ala Val Glu Val Ser Pro Val Met Glu Asp Ile Thr Arg
 1 5 10 15

Gln Ser Lys Lys Thr Ser Val Glu Asn Glu Thr Gly Asp Asp Gln Ser
 20 25 30

Ala Thr Ser Val Val Leu Lys Ala Lys Arg Lys Arg Arg Ser Gln Pro
 35 40 45

Arg Asp Ala Pro Pro Gln Arg Ser Ser Val His Arg Gly Val Thr Arg
 50 55 60

His Arg Trp Thr Gly Arg Tyr Glu Ala His Leu Trp Asp Lys Asn Ser
 65 70 75 80

Trp Asn Glu Thr Gln Thr Lys Lys Gly Arg Gln Val Tyr Leu Gly Ala
 85 90 95

Tyr Asp Glu Glu Asp Ala Ala Ala Arg Ala Tyr Asp Leu Ala Ala Leu
 100 105 110

Lys Tyr Trp Gly Arg Asp Thr Ile Leu Asn Phe Pro Leu Cys Asn Tyr
 115 120 125

Glu Glu Asp Ile Lys Glu Met Glu Ser Gln Ser Lys Glu Glu Tyr Ile
 130 135 140

Gly Ser Leu Arg Arg Lys Ser Ser Gly Phe Ser Arg Gly Val Ser Lys
 145 150 155 160

Tyr Arg Gly Val Ala Lys His His His Asn Gly Arg Trp Glu Ala Arg
 165 170 175

pctau2012001598-seql

Ile Gly Arg Val Phe Gly Asn Lys Tyr Leu Tyr Leu Gly Thr Tyr Ala
180 185 190

Thr Gln Glu Glu Ala Ala Ile Ala Tyr Asp Ile Ala Ala Ile Glu Tyr
195 200 205

Arg Gly Leu Asn Ala Val Thr Asn Phe Asp Ile Ser Arg Tyr Leu Lys
210 215 220

Leu Pro Val Pro Glu Asn Pro Ile Asp Thr Ala Asn Asn Leu Leu Glu
225 230 235 240

Ser Pro His Ser Asp Leu Ser Pro Phe Ile Lys Pro Asn His Glu Ser
245 250 255

Asp Leu Ser Gln Ser Gln Ser Ser Ser Glu Asp Asn Asp Asp Arg Lys
260 265 270

Thr Lys Leu Leu Lys Ser Ser Pro Leu Val Ala Glu Glu Val Ile Gly
275 280 285

Pro Ser Thr Pro Pro Glu Ile Ala Pro Pro Arg Arg Ser Phe Pro Glu
290 295 300

Asp Ile Gln Thr Tyr Phe Gly Cys Gln Asn Ser Gly Lys Leu Thr Ala
305 310 315 320

Glu Glu Asp Asp Val Ile Phe Gly Asp Leu Asp Ser Phe Leu Thr Pro
325 330 335

Asp Phe Tyr Ser Glu Leu Asn Asp Cys
340 345

<210> 319
<211> 299
<212> PRT
<213> Glycine max

<400> 319

Met Ala Lys Lys Ser Gln Lys Ser Leu Lys Asn Asn Asn Asn Asn Asn
1 5 10 15

Thr Thr Arg Lys Arg Thr Arg Lys Ser Val Pro Arg Asp Ser Pro Pro
20 25 30

Gln Arg Ser Ser Ile Tyr Arg Gly Val Thr Arg His Arg Trp Thr Gly
35 40 45

Arg Tyr Glu Ala His Leu Trp Asp Lys Asn Cys Trp Asn Glu Ser Gln
50 55 60

pctau2012001598-seql

Ser Lys Lys Gly Arg Gln Val Tyr Leu Gly Ala Tyr Asp Asp Glu Glu
65 70 75 80

Ala Ala Ala Arg Ala Tyr Asp Leu Ala Ala Leu Lys Tyr Trp Gly Gln
85 90 95

Asp Thr Ile Leu Asn Phe Pro Leu Ser Asn Tyr Glu Glu Lys Leu Lys
100 105 110

Glu Met Glu Gly Gln Ser Lys Glu Glu Tyr Ile Gly Ser Leu Arg Arg
115 120 125

Lys Ser Ser Gly Phe Ser Arg Gly Val Ser Lys Tyr Arg Gly Val Ala
130 135 140

Arg His His His Asn Gly Arg Trp Glu Ala Arg Ile Gly Arg Val Phe
145 150 155 160

Gly Asn Lys Tyr Leu Tyr Leu Gly Thr Tyr Ala Thr Gln Glu Glu Ala
165 170 175

Ala Ala Ala Tyr Asp Met Ala Ala Ile Glu Tyr Arg Gly Leu Asn Ala
180 185 190

Val Thr Asn Phe Asp Leu Ser Arg Tyr Ile Asn Trp Pro Arg Pro Lys
195 200 205

Thr Glu Glu Asn His Gln Asn Thr Pro Ser Asn Gln Asn Val Asn Ser
210 215 220

Asn Ala Glu Leu Glu Leu Gly Ser Ala Ser Asp Glu Ile Thr Glu Glu
225 230 235 240

Gly Val Ala Arg Ser Ser Glu Ser Glu Ser Asn Pro Ser Arg Arg Thr
245 250 255

Phe Pro Glu Asp Ile Gln Thr Ile Phe Glu Asn Asn Gln Asp Ser Gly
260 265 270

Ile Tyr Ile Glu Asn Asp Asp Ile Ile Phe Gly Asp Leu Gly Ser Phe
275 280 285

Gly Ala Pro Ile Phe His Phe Glu Leu Asp Val
290 295

<210> 320

<211> 393

<212> PRT

<213> Brachypodium distachyon

<400> 320

pctau2012001598-seql

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

pctau2012001598-seql

His Ser Ser Ser Ala Ala Ala Ala Ala Ala Gln Glu Phe Pro Leu Pro
275 280 285

Pro Ala Leu Gly His Ala Pro Thr Thr Ser Ala Leu Ser Leu Leu Leu
290 295 300

Gln Ser Pro Lys Phe Lys Glu Met Ile Glu Arg Thr Ser Ala Ala Glu
305 310 315 320

Thr Thr Thr Thr Ala Thr Thr Thr Ser Ser Ser Ser Pro Arg Pro
325 330 335

Ala Ala Ser Pro Gln Cys Ser Phe Pro Glu Asp Ile Gln Thr Phe Phe
340 345 350

Gly Cys Asp Asp Gly Val Gly Val Gly Val Gly Ala Val Gly Tyr Thr
355 360 365

Asp Val Asp Gly Leu Phe Phe Gly Asp Leu Ser Ala Tyr Ala Ser Ser
370 375 380

Thr Ala Phe His Phe Glu Leu Asp Leu
385 390

<210> 321
<211> 379
<212> PRT
<213> Oryza sativa

<400> 321

Met Ala Lys Pro Arg Lys Asn Ser Thr Thr Thr Asn Thr Ser Ser Ser
1 5 10 15

Gly Val Ala Ala Ala Ala Ala Ala Ala Ala Val Lys Pro Lys Arg Thr
20 25 30

Arg Lys Ser Val Pro Arg Glu Ser Pro Ser Gln Arg Ser Ser Val Tyr
35 40 45

Arg Gly Val Thr Arg His Arg Trp Thr Gly Arg Phe Glu Ala His Leu
50 55 60

Trp Asp Lys Asn Ser Trp Asn Glu Ser Gln Asn Lys Lys Gly Lys Gln
65 70 75 80

Val Tyr Leu Gly Ala Tyr Asp Asp Glu Glu Ala Ala Ala Arg Ala Tyr
85 90 95

Asp Leu Ala Ala Leu Lys Tyr Trp Gly Pro Asp Thr Ile Leu Asn Phe
100 105 110

Pro Leu Ser Ala Tyr Glu Gly Glu Leu Lys Glu Met Glu Gly Gln Ser

115

120

125

Arg Glu Glu Tyr Ile Gly Ser Leu Arg Arg Lys Ser Ser Gly Phe Ser
 130 135 140

Arg Gly Val Ser Lys Tyr Arg Gly Val Ala Arg His His His Asn Gly
 145 150 155 160

Arg Trp Glu Ala Arg Ile Gly Arg Val Phe Gly Asn Lys Tyr Leu Tyr
 165 170 175

Leu Gly Thr Tyr Ala Thr Gln Glu Glu Ala Ala Met Ala Tyr Asp Met
 180 185 190

Ala Ala Ile Glu Tyr Arg Gly Leu Asn Ala Val Thr Asn Phe Asp Leu
 195 200 205

Ser Arg Tyr Ile Lys Trp Leu Arg Pro Gly Ala Asp Gly Ala Gly Ala
 210 215 220

Ala Gln Asn Pro His Pro Met Leu Gly Ala Leu Ser Ala Gln Asp Leu
 225 230 235 240

Pro Ala Ile Asp Leu Asp Ala Met Ala Ser Ser Phe Gln His Asp Gly
 245 250 255

His Gly Ala Ala Ala Ala Ala Ala Gln Leu Ile Pro Ala Arg His Ser
 260 265 270

Leu Gly His Thr Pro Thr Thr Ser Ala Leu Ser Leu Leu Leu Gln Ser
 275 280 285

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

Thr Thr Ser Ser Thr Thr Thr Ser Ser Ser Ser Pro Ser Pro Pro Gln
 305 310 315 320

Ala Thr Lys Asp Asp Gly Ala Ser Pro Gln Cys Ser Phe Pro Lys Asp
 325 330 335

Ile Gln Thr Tyr Phe Gly Cys Ala Ala Glu Asp Gly Ala Ala Gly Ala
 340 345 350

Gly Tyr Ala Asp Val Asp Gly Leu Phe Phe Gly Asp Leu Thr Ala Tyr
 355 360 365

Ala Ser Pro Ala Phe His Phe Glu Leu Asp Leu
 370 375

<210> 322

pctau2012001598-seql

<211> 398

<212> PRT

<213> Sorghum bicolor

<400> 322

Met Ala Lys Pro Arg Lys Asn Ser Ala Ala Ala Asn Asn Asn Asn Ser
 1 5 10 15

Ser Ser Asn Gly Ala Gly Asp Leu Thr Pro Arg Ala Lys Pro Lys Arg
 20 25 30

Thr Arg Lys Ser Val Pro Arg Glu Ser Pro Thr Gln Arg Ser Ser Val
 35 40 45

Tyr Arg Gly Val Thr Arg His Arg Trp Thr Gly Arg Phe Glu Ala His
 50 55 60

Leu Trp Asp Lys Asn Ser Trp Asn Glu Ser Gln Asn Lys Lys Gly Lys
 65 70 75 80

Gln Val Tyr Leu Gly Ala Tyr Asp Asp Glu Glu Ala Ala Ala Arg Ala
 85 90 95

Tyr Asp Leu Ala Ala Leu Lys Tyr Trp Gly Pro Asp Thr Ile Leu Asn
 100 105 110

Phe Pro Ala Ser Ala Tyr Glu Gly Glu Met Lys Gly Met Glu Gly Gln
 115 120 125

Ser Arg Glu Glu Tyr Ile Gly Ser Leu Arg Arg Lys Ser Ser Gly Phe
 130 135 140

Ser Arg Gly Val Ser Lys Tyr Arg Gly Val Ala Arg His His His Asn
 145 150 155 160

Gly Arg Trp Glu Ala Arg Ile Gly Arg Val Phe Gly Asn Lys Tyr Leu
 165 170 175

Tyr Leu Gly Thr Tyr Ala Thr Gln Glu Glu Ala Ala Met Ala Tyr Asp
 180 185 190

Met Ala Ala Ile Glu Tyr Arg Gly Leu Asn Ala Val Thr Asn Phe Asp
 195 200 205

Leu Ser Arg Tyr Ile Lys Trp Leu Arg Pro Gly Ala Gly Gly Met Ala
 210 215 220

Ala Ala Ala Ala Ala Ala Gln Asn Pro His Pro Met Leu Gly Gly Leu
 225 230 235 240

Ala Gln Gln Leu Leu Leu Pro Pro Pro Ala Asp Thr Thr Thr Thr Asp
 245 250 255

pctau2012001598-seql

Gly Ala Gly Ala Ala Ala Phe Gln His Asp His His Gly Ala Glu Ala
260 265 270

Phe Pro Leu Pro Pro Arg Thr Ser Leu Gly His Thr Pro Thr Thr Ser
275 280 285

Ala Leu Ser Leu Leu Leu Gln Ser Pro Lys Phe Lys Glu Met Ile Gln
290 295 300

Arg Thr Glu Ser Gly Thr Thr Thr Thr Thr Thr Thr Ser Ser Leu
305 310 315 320

Ser Ser Ser Pro Pro Pro Thr Pro Ser Pro Ser Pro Pro Arg Arg Ser
325 330 335

Pro Ala Pro Thr Gln Pro Pro Val Gln Ala Ala Ala Arg Asp Ala Ser
340 345 350

Pro His Gln Arg Gly Phe Pro Glu Asp Val Gln Thr Phe Phe Gly Cys
355 360 365

Glu Asp Thr Ala Gly Ile Asp Val Glu Ala Leu Phe Phe Gly Asp Leu
370 375 380

Ala Ala Tyr Ala Thr Pro Ala Phe His Phe Glu Met Asp Leu
385 390 395

<210> 323
<211> 396
<212> PRT
<213> Zea mays

<400> 323

Met Ala Arg Pro Arg Lys Asn Ser Ala Ala Ala Ala Asn Asn Asn Asn
1 5 10 15

Ser Asn Thr Thr Asn Ala Gly Asn Ala Ala Val Asp Leu Ala Ala Arg
20 25 30

Val Lys Pro Lys Arg Thr Arg Lys Ser Val Pro Arg Glu Ser Pro Ser
35 40 45

Gln Arg Ser Ser Val Tyr Arg Gly Val Thr Arg His Arg Trp Thr Gly
50 55 60

Arg Phe Glu Ala His Leu Trp Asp Lys Asn Ser Trp Asn Glu Ser Gln
65 70 75 80

Asn Lys Lys Gly Lys Gln Val Tyr Leu Gly Ala Tyr Asp Asp Glu Asp
85 90 95

pctau2012001598-seql

Ala Ala Ala Arg Ala Tyr Asp Leu Ala Ala Leu Lys Tyr Trp Gly Pro
100 105 110
Asp Thr Ile Leu Asn Phe Pro Ala Ser Ala Tyr Glu Ala Glu Leu Lys
115 120 125
Glu Met Glu Gly Gln Ser Arg Glu Glu Tyr Ile Gly Ser Leu Arg Arg
130 135 140
Lys Ser Ser Gly Phe Ser Arg Gly Val Ser Lys Tyr Arg Gly Val Ala
145 150 155 160
Arg His His His Asn Gly Arg Trp Glu Ala Arg Ile Gly Arg Val Phe
165 170 175
Gly Asn Lys Tyr Leu Tyr Leu Gly Thr Tyr Gly Thr Gln Glu Glu Ala
180 185 190
Ala Met Ala Tyr Asp Met Ala Ala Ile Glu Tyr Arg Gly Leu Asn Ala
195 200 205
Val Thr Asn Phe Asp Leu Ser Arg Tyr Ile Lys Trp Leu Arg Pro Gly
210 215 220
Ala Gly Ala Ala Gln Asn Pro His Pro Met Leu Asp Gly Leu Ala Gln
225 230 235 240
Gln Leu Leu Leu Ser Pro Glu Gly Thr Ile Asp Gly Ala Ala Phe His
245 250 255
Gln Gln Gln His Asp His Arg Gln Gln Gly Ala Ala Glu Leu Pro Leu
260 265 270
Pro Pro Arg Ala Ser Leu Gly His Thr Pro Thr Thr Ser Ala Leu Gly
275 280 285
Leu Leu Leu Gln Ser Ser Lys Phe Lys Glu Met Ile Gln Arg Ala Ser
290 295 300
Ala Ala Glu Ser Gly Thr Thr Thr Val Thr Thr Thr Ser Ser Ser Ser
305 310 315 320
Ser Gln Pro Pro Thr Pro Thr Pro Thr Pro Ser Pro Ser Pro Pro Pro
325 330 335
Thr Pro Pro Val Gln Pro Ala Arg Asp Ala Ser Pro Gln Cys Ser Phe
340 345 350
Pro Glu Asp Ile Gln Thr Phe Phe Gly Cys Glu Asp Val Ala Gly Val
355 360 365

pctau2012001598-seql

Gly Ala Gly Val Asp Val Asp Ala Leu Phe Phe Gly Asp Leu Ala Ala
370 375 380

Tyr Ala Ser Pro Ala Phe His Phe Glu Met Asp Leu
385 390 395

<210> 324
<211> 362
<212> PRT
<213> Glycine max
<400> 324

Met Ala Lys Lys Ser Gln Leu Arg Thr Gln Lys Asn Asn Ala Thr Asn
1 5 10 15

Asp Asp Ile Asn Leu Asn Ala Thr Asn Thr Val Ile Thr Lys Val Lys
20 25 30

Arg Thr Arg Arg Ser Val Pro Arg Asp Ser Pro Pro Gln Arg Ser Ser
35 40 45

Ile Tyr Arg Gly Val Thr Arg His Arg Trp Thr Gly Arg Tyr Glu Ala
50 55 60

His Leu Trp Asp Lys His Cys Trp Asn Glu Ser Gln Asn Lys Lys Gly
65 70 75 80

Arg Gln Gly Ala Tyr Asp Asn Glu Glu Ala Ala Ala His Ala Tyr Asp
85 90 95

Leu Ala Ala Leu Lys Tyr Trp Gly Gln Asp Thr Ile Leu Asn Phe Pro
100 105 110

Leu Ser Asn Tyr Leu Asn Glu Leu Lys Glu Met Glu Gly Gln Ser Arg
115 120 125

Glu Glu Tyr Ile Gly Ser Leu Arg Arg Lys Ser Ser Gly Phe Ser Arg
130 135 140

Gly Ile Ser Lys Tyr Arg Gly Val Ala Arg His His His Asn Gly Arg
145 150 155 160

Trp Glu Ala Arg Ile Gly Lys Val Phe Gly Asn Lys Tyr Leu Tyr Leu
165 170 175

Gly Thr Tyr Ala Thr Gln Glu Glu Ala Ala Thr Ala Tyr Asp Leu Ala
180 185 190

Ala Ile Glu Tyr Arg Gly Leu Asn Ala Val Thr Asn Phe Asp Leu Ser
195 200 205

pctau2012001598-seql

Arg Tyr Ile Lys Trp Leu Lys Pro Asn Asn Thr Asn Ser Asn Asn Asp
210 215 220

Gln Ile Ser Ile Asn Leu Thr Asn Ile Asn Asn Cys Thr Asn Asn
225 230 235 240

Phe Ile Pro Asn Pro Asp Gln Glu Gln Glu Val Ser Phe Phe His Asn
245 250 255

Gln Asp Ser Leu Asn Asn Thr Ile Val Glu Glu Ala Thr Leu Val Pro
260 265 270

His Gln Pro Arg Pro Ala Ser Ala Thr Leu Ala Leu Glu Leu Leu Leu
275 280 285

Gln Ser Ser Lys Phe Lys Glu Met Val Glu Met Thr Ser Val Ala Asn
290 295 300

Leu Ser Thr Gln Met Glu Ser Asp Gln Leu Pro Gln Cys Thr Phe Pro
305 310 315 320

Asp His Ile Gln Thr Tyr Phe Glu Tyr Glu Asp Ser Asn Lys Tyr Glu
325 330 335

Glu Gly Asp Asp Leu Leu Phe Lys Phe Ser Glu Phe Ser Ser Ile Val
340 345 350

Pro Phe Tyr His Cys Asp Glu Phe Glu Ser
355 360

<210> 325
<211> 370
<212> PRT
<213> Glycine max

<400> 325

Met Ala Lys Lys Ser Gln Leu Arg Thr Gln Lys Asn Asn Val Thr Thr
1 5 10 15

Asn Asp Asp Asn Asn Leu Asn Val Thr Asn Thr Val Thr Thr Lys Val
20 25 30

Lys Arg Thr Arg Arg Ser Val Pro Arg Asp Ser Pro Pro Gln Arg Ser
35 40 45

Ser Ile Tyr Arg Gly Val Thr Arg His Arg Trp Thr Gly Arg Tyr Glu
50 55 60

Ala His Leu Trp Asp Lys His Cys Trp Asn Glu Ser Gln Asn Lys Lys
65 70 75 80

Gly Arg Gln Val Tyr Leu Gly Ala Tyr Asp Asn Glu Glu Ala Ala Ala
Page 403

pctau2012001598-seql

85

90

95

His Ala Tyr Asp₁₀₀ Leu Ala Ala Leu Lys₁₀₅ Tyr Trp Gly Gln Asp₁₁₀ Thr Ile

Leu Asn Phe₁₁₅ Pro Leu Ser Asn Tyr₁₂₀ Leu Asn Glu Leu Lys₁₂₅ Glu Met Glu

Gly Gln₁₃₀ Ser Arg Glu Glu Tyr₁₃₅ Ile Gly Ser Leu Arg₁₄₀ Arg Lys Ser Ser

Gly Phe Ser Arg Gly Ile₁₅₀ Ser Lys Tyr Arg Gly₁₅₅ Val Ala Arg His His₁₆₀

His Asn Gly Arg Trp₁₆₅ Glu Ala Arg Ile Gly₁₇₀ Lys Val Phe Gly Asn₁₇₅ Lys

Tyr Leu Tyr Leu₁₈₀ Gly Thr Tyr Ala Thr₁₈₅ Gln Glu Glu Ala Ala₁₉₀ Thr Ala

Tyr Asp Leu₁₉₅ Ala Ala Ile Glu Tyr₂₀₀ Arg Gly Leu Asn Ala₂₀₅ Val Thr Asn

Phe Asp₂₁₀ Leu Ser Arg Tyr Ile₂₁₅ Lys Trp Leu Lys Pro₂₂₀ Asn Asn Asn Thr

Asn Asn Val Ile Asp₂₃₀ Gln Ile Ser Ile Asn₂₃₅ Leu Thr Asn Ile Asn₂₄₀

Asn Asn Asn Asn Cys₂₄₅ Thr Asn Ser Phe Thr₂₅₀ Pro Ser Pro Asp Gln₂₅₅ Glu

Gln Glu Ala Ser₂₆₀ Phe Phe His Asn Lys₂₆₅ Asp Ser Leu Asn Asn₂₇₀ Thr Ile

Val Glu Glu₂₇₅ Val Thr Leu Val Pro₂₈₀ His Gln Pro Arg Pro₂₈₅ Ala Ser Ala

Thr Ser₂₉₀ Ala Leu Glu Leu Leu₂₉₅ Leu Gln Ser Ser Lys₃₀₀ Phe Lys Glu Met

Met₃₀₅ Glu Met Thr Ser Val₃₁₀ Ala Asn Leu Ser Ser₃₁₅ Thr Gln Met Glu Ser₃₂₀

Glu Leu Pro Gln Cys₃₂₅ Thr Phe Pro Asp His₃₃₀ Ile Gln Thr Tyr Phe₃₃₅ Glu

Tyr Glu Asp Ser₃₄₀ Asn Arg Tyr Glu Glu₃₄₅ Gly Asp Asp Leu Met₃₅₀ Phe Lys

Phe Asn Glu Phe Ser Ser Ile Val Pro Phe Tyr Gln Cys Asp Glu Phe

355

Glu Ser
370

<210> 326
<211> 356
<212> PRT
<213> Medicago truncatula

<400> 326

Met Ala Lys Lys Ser Gln Lys Gln Ile Glu Lys Asp Asp Asn Ala Ser
1 5 10 15

Asn Asp Asn Asp Asn Leu Asn Pro Ser Asn Thr Val Thr Thr Lys Ala
20 25 30

Lys Arg Thr Arg Lys Ser Val Pro Arg Thr Ser Pro Pro Gln Arg Ser
35 40 45

Ser Ile Tyr Arg Gly Val Thr Arg His Arg Trp Thr Gly Arg Tyr Glu
50 55 60

Ala His Leu Trp Asp Lys Asn Cys Trp Asn Glu Ser Gln Asn Lys Lys
65 70 75 80

Gly Arg Gln Gly Ala Tyr Asp Asn Glu Glu Thr Ala Ala His Ala Tyr
85 90 95

Asp Leu Ala Ala Leu Lys Tyr Trp Gly Gln Asp Thr Ile Ile Asn Phe
100 105 110

Pro Leu Ser Asn Tyr Gln Lys Glu Leu Ile Glu Met Glu Ser Gln Ser
115 120 125

Arg Glu Glu Tyr Ile Gly Ser Leu Arg Arg Lys Ser Ser Gly Phe Ser
130 135 140

Arg Gly Val Ser Lys Tyr Arg Gly Val Ala Arg His His His Asn Gly
145 150 155 160

Arg Trp Glu Ala Arg Ile Gly Lys Val Phe Gly Asn Lys Tyr Leu Tyr
165 170 175

Leu Gly Thr Tyr Ala Thr Gln Glu Glu Ala Ala Thr Ala Tyr Asp Met
180 185 190

Ala Ala Ile Glu Tyr Arg Gly Leu Asn Ala Val Thr Asn Phe Asp Leu
195 200 205

Ser Arg Tyr Ile Lys Trp Leu Lys Pro Asn Asn Asn Asn Asn Asp Asp
210 215 220

pctau2012001598-seql

Asn Asn Lys Ser Asn Ile Asn Leu Cys Asp Ile Asn Ser Asn Ser Ser
225 230 235 240

Ala Asn Asp Ser Asn Ser Asn Glu Glu Leu Glu Phe Ser Leu Val Asp
245 250 255

Asn Glu Ile Ser Leu Asn Asn Ser Ile Asp Glu Ala Thr Leu Val Gln
260 265 270

Pro Arg Pro Thr Ser Ala Thr Ser Ala Leu Glu Leu Leu Leu Gln Ser
275 280 285

Ser Lys Phe Lys Glu Met Val Glu Met Ala Ser Met Thr Ser Asn Val
290 295 300

Ser Thr Thr Leu Glu Ser Asp Gln Leu Ser Gln Cys Ala Phe Pro Asp
305 310 315 320

Asp Ile Gln Thr Tyr Phe Glu Tyr Glu Asn Phe Asn Asp Thr Met Leu
325 330 335

Glu Asp Leu Asn Ser Ile Met Pro Thr Phe His Tyr Asp Phe Glu Gly
340 345 350

Ala Glu Val Leu
355

<210> 327
<211> 347
<212> PRT
<213> Glycine max

<400> 327

Met Ala Lys Gln Gln Thr His Lys Ile Asn Ala Ser Thr Asn Asn Asn
1 5 10 15

Ile Ser Thr Thr Asn Thr Val Thr Ala Lys Val Lys Arg Thr Arg Arg
20 25 30

Ser Val Pro Arg Asp Ser Pro Pro Gln Arg Ser Ser Ile Tyr Arg Gly
35 40 45

Val Thr Arg His Arg Trp Thr Gly Arg Tyr Glu Ala His Leu Trp Asp
50 55 60

Lys Asn Cys Trp Asn Glu Ser Gln Asn Lys Lys Gly Arg Gln Gly Ala
65 70 75 80

Tyr Asp Asp Glu Glu Ala Ala Ala His Ala Tyr Asp Leu Ala Ala Leu
85 90 95

pctau2012001598-seql

Lys Tyr Trp Gly Gln Asp Thr Ile Leu Asn Phe Pro Leu Ser Thr Tyr
100 105 110
Gln Asn Glu Leu Lys Glu Met Glu Gly Gln Ser Arg Glu Glu Tyr Ile
115 120 125
Gly Ser Leu Arg Arg Lys Ser Ser Gly Phe Ser Arg Gly Val Ser Lys
130 135 140
Tyr Arg Gly Val Ala Arg His His His Asn Gly Arg Trp Glu Ala Arg
145 150 155 160
Ile Gly Arg Val Phe Gly Asn Lys Tyr Leu Tyr Leu Gly Thr Tyr Ala
165 170 175
Thr Gln Glu Glu Ala Ala Thr Ala Tyr Asp Met Ala Ala Ile Glu Tyr
180 185 190
Arg Gly Val Asn Ala Val Thr Asn Phe Asp Leu Ser Arg Tyr Ile Lys
195 200 205
Trp Leu Lys Pro Asn Asn Asn Asn Thr Thr Val Asn Ser Asn Leu Ile
210 215 220
Asp Ser Asn Pro Asn Cys Glu Thr Asn Phe Thr Ser Asn Ser Asn Gln
225 230 235 240
Gln Gln Gly Phe Asn Phe Phe Asn Arg Gln Glu Ser Phe Asn Asn Glu
245 250 255
Glu Ala Ala Met Thr Gln Pro Arg Pro Ala Val Ala Thr Ser Ala Leu
260 265 270
Gly Leu Leu Leu Gln Ser Ser Lys Phe Lys Glu Met Met Glu Met Thr
275 280 285
Ser Ala Thr Asp Leu Ser Thr Pro Pro Ser Glu Ser Glu Leu Pro Ser
290 295 300
Cys Thr Phe Pro Asp Asp Ile Gln Thr Tyr Phe Glu Cys Glu Asp Ser
305 310 315 320
His Arg Tyr Gly Glu Gly Asp Asp Ile Met Phe Ser Val Leu Asn Gly
325 330 335
Phe Val Pro Pro Met Phe His Cys Asp Asp Phe
340 345

<210> 328
<211> 351
<212> PRT

pctau2012001598-seql

<213> Glycine max

<400> 328

```

Met Ala Lys Gln Gln Thr His Glu Ile Asn Ala Ser Thr Asn Asn Asn
 1          5          10          15

Ile Asn Thr Thr Lys Thr Val Thr Thr Lys Val Lys Arg Thr Arg Arg
          20          25          30

Ser Val Pro Arg Asn Ser Pro Pro Gln Arg Ser Ser Ile Tyr Arg Gly
          35          40          45

Val Thr Arg His Arg Trp Thr Gly Arg Tyr Glu Ala His Leu Trp Asp
 50          55          60

Lys Asn Cys Trp Asn Glu Ser Gln Asn Lys Lys Gly Arg Gln Gly Ala
65          70          75          80

Tyr Asp Asp Glu Glu Ala Ala Ala His Ala Tyr Asp Leu Ala Ala Leu
          85          90          95

Lys Tyr Trp Gly Gln Asp Thr Ile Leu Asn Phe Pro Leu Ser Thr Tyr
100          105          110

Gln Asn Glu Leu Lys Glu Met Glu Gly Gln Ser Arg Glu Glu Tyr Ile
115          120          125

Gly Ser Leu Arg Arg Lys Ser Ser Gly Phe Ser Arg Gly Val Ser Lys
130          135          140

Tyr Arg Gly Val Ala Arg His His His Asn Gly Arg Trp Glu Ala Arg
145          150          155          160

Ile Gly Arg Val Phe Gly Asn Lys Tyr Leu Tyr Leu Gly Thr Tyr Ala
          165          170          175

Thr Gln Glu Glu Ala Ala Thr Ala Tyr Asp Met Ala Ala Ile Glu Tyr
180          185          190

Arg Gly Leu Asn Ala Val Thr Asn Phe Asp Leu Ser Arg Tyr Ile Lys
195          200          205

Trp Leu Lys Pro Asn Asn Asn Asn Asn Lys Val Asn Ser Asn Asn Leu
210          215          220

Ile Val Ser Ile Pro Asn Cys Ala Thr Asn Phe Thr Pro Asn Ser Asn
225          230          235          240

Gln Gln Gln Gly Phe Asn Phe Phe Asn Ser Gln Glu Ser Phe Asn Asn
          245          250          255

```

pctau2012001598-seql

Asn Glu Glu Ala Ala Met Thr Gln Pro Arg Pro Ala Ala Ala Thr Ser
260 265 270

Ala Leu Gly Leu Leu Leu Gln Ser Ser Lys Phe Lys Glu Met Met Glu
275 280 285

Met Thr Ser Ala Ile Asp Leu Ser Thr Pro Pro Ser Glu Ser Glu Leu
290 295 300

Pro Pro Cys Thr Phe Pro Asp Asp Ile Gln Thr Tyr Phe Glu Cys Glu
305 310 315 320

Asp Ser His Arg Tyr Gly Glu Gly Asp Asp Ile Met Phe Ser Glu Leu
325 330 335

Asn Gly Phe Val Pro Pro Met Phe His Cys Asp Asp Phe Glu Ala
340 345 350

<210> 329
<211> 353
<212> PRT
<213> Populus trichocarpa

<400> 329

Met Ala Lys Leu Ser Gln Lys Asn Thr Lys Asn Thr Ala Ser Asn Asn
1 5 10 15

Asn Asn Thr Thr Asn Gly Val Thr Lys Val Lys Arg Thr Arg Arg Ser
20 25 30

Val Pro Arg Asp Ser Pro Pro Gln Arg Ser Ser Ile Tyr Arg Gly Val
35 40 45

Thr Arg His Arg Trp Thr Gly Arg Tyr Glu Ala His Leu Trp Asp Lys
50 55 60

Asn Cys Trp Asn Glu Ser Gln Asn Lys Lys Gly Arg Gln Gly Ala Tyr
65 70 75 80

Asp Asp Glu Glu Ala Ala Ala His Ala Tyr Asp Leu Ala Ala Leu Lys
85 90 95

Tyr Trp Gly Pro Glu Thr Ile Leu Asn Phe Pro Leu Ser Thr Tyr Gln
100 105 110

Asn Glu Leu Lys Glu Met Glu Gly Gln Ser Arg Glu Glu Cys Ile Gly
115 120 125

Ser Leu Arg Arg Lys Ser Ser Gly Phe Ser Arg Gly Val Ser Lys Tyr
130 135 140

Arg Gly Val Ala Arg His His His Asn Gly Arg Trp Glu Ala Arg Ile

pctau2012001598-seq1

145 150 155 160

Gly Arg Val Phe Gly Asn Lys Tyr Leu Tyr Leu Gly Thr Tyr Ala Thr
165 170 175

Gln Glu Glu Ala Ala Thr Ala Tyr Asp Met Ala Ala Ile Glu Tyr Arg
180 185 190

Gly Leu Asn Ala Val Thr Asn Phe Asp Leu Ser Arg Tyr Ile Lys Trp
195 200 205

Leu Lys Pro Asn Gln Asn Asn Thr Asp Asn Asn Asn Gly Leu Asp Leu
210 215 220

Pro Asn Pro Ile Ile Gly Thr Asp Asn Ser Thr His Pro Asn Pro Asn
225 230 235 240

Gln Glu Leu Gly Thr Thr Phe Leu Gln Ile Asn Gln Gln Thr Tyr Gln
245 250 255

Pro Ser Glu Thr Thr Leu Thr Gln Pro Arg Pro Ala Thr Asn Pro Ser
260 265 270

Ser Ala Leu Gly Leu Leu Leu Gln Ser Ser Lys Phe Lys Glu Met Met
275 280 285

Glu Met Thr Ala Val Thr Asp Cys Pro Pro Thr Pro Pro Ser Gly Leu
290 295 300

Asp 305 Pro Thr Pro Cys 310 Ser Phe Leu Glu Asp 315 Val Gln Thr Tyr Phe Asp 320

Cys Leu Asp Ser Ser Asn Tyr Gly Asp Gln Gly Asp Asp Met Ile Phe
325 330 335

Gly Asp Leu Asn Ser Phe Val Pro Pro Met Phe Gln Cys Asp Phe Glu
340 345 350

Thr

```
<210> 330
<211> 323
<212> PRT
<213> Vitis vinifera
```

$\langle 400 \rangle$ 330

Met Ala Lys Leu Ser Gln Gln Asn His Lys Asn Ser Ala Asn Ser Asn
1 5 10 15

Ala Thr Asn Thr Thr Leu Ser Val Thr Lys Val Lys Arg Thr Arg Lys
20 25 30

pctau2012001598-seql

Thr Val Pro Arg Asp Ser Pro Pro Gln Arg Ser Ser Ile Tyr Arg Gly
 35 40 45
 Val Thr Arg His Arg Trp Thr Gly Arg Tyr Glu Ala His Leu Trp Asp
 50 55 60
 Lys Asn Cys Trp Asn Glu Ser Gln Asn Lys Lys Gly Arg Gln Val Tyr
 65 70 75 80
 Leu Gly Ala Tyr His Asp Glu Glu Ala Ala His Ala Tyr Asp Leu
 85 90 95
 Ala Ala Leu Lys Tyr Trp Gly Pro Glu Thr Ile Leu Asn Phe Pro Leu
 100 105 110
 Ser Thr Tyr Glu Lys Glu Leu Lys Glu Met Glu Gly Leu Ser Arg Glu
 115 120 125
 Glu Tyr Ile Gly Ser Leu Arg Arg Arg Ser Ser Gly Phe Ser Arg Gly
 130 135 140
 Val Ser Lys Tyr Arg Gly Val Ala Arg His His His Asn Gly Arg Trp
 145 150 155 160
 Glu Ala Arg Ile Gly Arg Val Phe Gly Asn Lys Tyr Leu Tyr Leu Gly
 165 170 175
 Thr Tyr Ala Thr Gln Glu Glu Ala Ala Thr Ala Tyr Asp Met Ala Ala
 180 185 190
 Ile Glu Tyr Arg Gly Leu Asn Ala Val Thr Asn Phe Asp Leu Ser Arg
 195 200 205
 Tyr Ile Lys Trp Leu Lys Pro Asn Gln Asn Asn Pro Cys Glu Gln Pro
 210 215 220
 Asn Asn Pro Asn Leu Asp Ser Asn Leu Thr Pro Asn Pro Asn His Asp
 225 230 235 240
 Phe Gly Ile Ser Phe Leu Asn His Pro Gln Thr Ser Gly Thr Ala Ala
 245 250 255
 Cys Lys Met Met Glu Met Thr Thr Ala Ala Asp His Leu Ser Thr Pro
 260 265 270
 Pro Glu Ser Glu Leu Pro Arg Cys Ser Phe Pro Asp Asp Ile Gln Thr
 275 280 285
 Tyr Phe Glu Cys Gln Asp Ser Gly Ser Tyr Glu Glu Gly Asp Asp Val
 290 295 300

pctau2012001598-seql

Ile Phe Ser Glu Leu Asn Ser Phe Ile Pro Pro Met Phe Gln Cys Asp
305 310 315 320

Phe Ser Ala

<210> 331
<211> 347
<212> PRT
<213> Vitis vinifera

<400> 331

Met Ala Lys Leu Ser Gln Gln Asn His Lys Asn Ser Ala Asn Ser Asn
1 5 10 15

Ala Thr Asn Thr Thr Leu Ser Val Thr Lys Val Lys Arg Thr Arg Lys
20 25 30

Thr Val Pro Arg Asp Ser Pro Pro Gln Arg Ser Ser Ile Tyr Arg Gly
35 40 45

Val Thr Arg His Arg Trp Thr Gly Arg Tyr Glu Ala His Leu Trp Asp
50 55 60

Lys Asn Cys Trp Asn Glu Ser Gln Asn Lys Lys Gly Arg Gln Gly Ala
65 70 75 80

Tyr His Asp Glu Glu Ala Ala Ala His Ala Tyr Asp Leu Ala Ala Leu
85 90 95

Lys Tyr Trp Gly Pro Glu Thr Ile Leu Asn Phe Pro Leu Ser Thr Tyr
100 105 110

Glu Lys Glu Leu Lys Glu Met Glu Gly Leu Ser Arg Glu Glu Tyr Ile
115 120 125

Gly Ser Leu Arg Arg Arg Ser Ser Gly Phe Ser Arg Gly Val Ser Lys
130 135 140

Tyr Arg Gly Val Ala Arg His His His Asn Gly Arg Trp Glu Ala Arg
145 150 155 160

Ile Gly Arg Val Phe Gly Asn Lys Tyr Leu Tyr Leu Gly Thr Tyr Ala
165 170 175

Thr Gln Glu Glu Ala Ala Thr Ala Tyr Asp Met Ala Ala Ile Glu Tyr
180 185 190

Arg Gly Leu Asn Ala Val Thr Asn Phe Asp Leu Ser Arg Tyr Ile Lys
195 200 205

pctau2012001598-seql

Trp Leu Lys Pro Asn Gln Asn Asn Pro Cys Glu Gln Pro Asn Asn Pro
210 215 220

Asn Leu Asp Ser Asn Leu Thr Pro Asn Pro Asn His Asp Phe Gly Ile
225 230 235 240

Ser Phe Leu Asn His Pro Gln Thr Ser Gly Thr Ala Ala Cys Ser Glu
245 250 255

Pro Pro Leu Thr Gln Thr Arg Pro Pro Ile Ala Ser Ser Ala Leu Gly
260 265 270

Leu Leu Leu Gln Ser Ser Lys Phe Lys Glu Met Met Glu Met Thr Thr
275 280 285

Ala Ala Asp His Leu Ser Thr Pro Pro Glu Ser Glu Leu Pro Arg Cys
290 295 300

Ser Phe Pro Asp Asp Ile Gln Thr Tyr Phe Glu Cys Gln Asp Ser Gly
305 310 315 320

Ser Tyr Glu Glu Gly Asp Asp Val Ile Phe Ser Glu Leu Asn Ser Phe
325 330 335

Ile Pro Pro Met Phe Gln Cys Asp Phe Ser Ala
340 345

<210> 332
<211> 275
<212> PRT
<213> Populus trichocarpa

<400> 332

Met Gly Lys Thr Ser Lys Gln Ser Leu Lys Asn Ser Ala Asn Thr Ser
1 5 10 15

Ile Asn Pro Ala Thr Lys Val Lys Arg Thr Arg Lys Ser Val Pro Arg
20 25 30

Asp Ser Pro Pro Gln Arg Ser Ser Ile Tyr Arg Gly Val Thr Arg His
35 40 45

Arg Trp Thr Gly Arg Tyr Glu Ala His Leu Trp Asp Lys Asn Cys Trp
50 55 60

Asn Glu Ser Gln Asn Lys Lys Gly Arg Gln Gly Ala Tyr Asp Asp Glu
65 70 75 80

Glu Ala Ala Gly His Ala Tyr Asp Leu Ala Ala Leu Lys Tyr Trp Gly
85 90 95

pctau2012001598-seql

Gln Asp Thr Ile Leu Asn Phe Pro Leu Ser Thr Tyr Glu Glu Glu Phe
100 105 110

Lys Glu Met Glu Gly His Ser Lys Glu Glu Tyr Ile Gly Ser Leu Arg
115 120 125

Arg Lys Ser Ser Gly Phe Ser Arg Gly Val Ser Lys Tyr Arg Gly Val
130 135 140

Ala Arg His His His Asn Gly Arg Trp Glu Ala Arg Ile Gly Arg Val
145 150 155 160

Phe Gly Asn Lys Tyr Leu Tyr Leu Gly Thr Tyr Ala Thr Gln Glu Glu
165 170 175

Ala Ala Thr Ala Tyr Asp Met Ala Ala Ile Glu Tyr Arg Gly Leu Asn
180 185 190

Ala Val Thr Asn Phe Asp Leu Ser Arg Tyr Ser Ser Lys Phe Lys Glu
195 200 205

Met Leu Glu Arg Thr Ser Ala Ser Asp Cys Pro Leu Thr Pro Pro Glu
210 215 220

Ser Asp Arg Asp Pro Pro Arg Arg Ser Phe Pro Asp Asp Ile Gln Thr
225 230 235 240

Tyr Phe Asp Cys Gln Asp Ser Ser Ser Tyr Thr Asp Gly Asp Asp Ile
245 250 255

Ile Phe Gly Asp Leu His Ser Phe Ala Ser Pro Ile Phe His Cys Glu
260 265 270

Leu Asp Gly
275

<210> 333
<211> 304
<212> PRT
<213> Vitis vinifera

<400> 333

Met Ala Lys Thr Ser Gln Lys Ser Gln Lys Thr Thr Gly Asn Ser Thr
1 5 10 15

Asn Asn Asn Gly Gly Ser Val Ala Lys Val Lys Arg Thr Arg Lys Ser
20 25 30

Val Pro Arg Asp Ser Pro Pro Gln Arg Ser Ser Ile Phe Arg Gly Val
35 40 45

Thr Arg His Arg Trp Thr Gly Arg Tyr Glu Ala His Leu Trp Asp Lys
Page 414

pctau2012001598-seql

50

55

60

Asn Cys Trp Asn Glu Ser Gln Asn Lys Lys Gly Arg Gln Val Tyr Leu
65 70 75 80

Gly Ala Tyr Asp Asp Glu Glu Ala Ala Ala His Ala Tyr Asp Leu Ala
85 90 95

Ala Leu Lys Tyr Trp Gly Gln Glu Thr Ile Leu Asn Phe Pro Leu Ser
100 105 110

Ala Tyr Gln Glu Glu Leu Lys Glu Met Glu Gly Gln Ser Lys Glu Glu
115 120 125

Tyr Ile Gly Ser Leu Arg Arg Lys Ser Ser Gly Phe Ser Arg Gly Val
130 135 140

Ser Lys Tyr Arg Gly Val Ala Arg His His His Asn Gly Arg Trp Glu
145 150 155 160

Ala Arg Ile Gly Arg Val Phe Gly Asn Lys Tyr Leu Tyr Leu Gly Thr
165 170 175

Tyr Ala Thr Gln Glu Glu Ala Ala Thr Ala Tyr Asp Met Ala Ala Ile
180 185 190

Glu Tyr Arg Gly Leu Asn Ala Val Thr Asn Phe Asp Leu Ser Arg Tyr
195 200 205

Ile Asn Ser Pro Ala Pro Asn Pro Asn Pro Ser Asp His Glu Leu Gly
210 215 220

Leu Ser Phe Leu Gln Gln Gln His Gly Ser Asp Ala Thr Glu Leu Pro
225 230 235 240

Leu Ser His Ala Arg Ser Asp Cys Pro Leu Thr Pro Pro Asp Gln Ile
245 250 255

Glu Met Pro Arg Ser Ser Phe Pro Asp Asp Ile Gln Thr Tyr Phe Asp
260 265 270

Cys Gln Glu Thr Asn Ser Tyr Gly Glu Ser Asp Asp Ile Ile Phe Gly
275 280 285

Asp Leu Lys Tyr Phe Ser Ser Pro Met Phe Gln Cys Glu Leu Asp Thr
290 295 300

<210> 334
<211> 393
<212> PRT
<213> Sorghum bicolor

pctau2012001598-seql

<400> 334

```

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

```

pctau2012001598-seql

Ser Ile Leu Ser Gln Ser Asp Ala Phe Lys Ser Phe Leu Glu Lys Ser
275 280 285

Thr Lys Leu Ser Glu Glu Cys Thr Leu Ser Lys Glu Ile Val Glu Gly
290 295 300

Lys Thr Val Ala Ser Val Pro Ala Thr Gly Tyr Asp Thr Gly Ala Ile
305 310 315 320

Asn Ile Asn Met Asn Glu Leu Leu Val Gln Arg Ser Thr Tyr Ser Met
325 330 335

Ala Pro Val Met Pro Thr Pro Met Lys Thr Thr Trp Ser Pro Ala Asp
340 345 350

Pro Ser Val Asp Pro Leu Phe Trp Ser Asn Phe Val Leu Pro Ser Ser
355 360 365

Gln Pro Val Thr Met Ala Thr Ile Thr Thr Thr Thr Asn Glu Val Ser
370 375 380

Ser Ser Asp Pro Phe Gln Ser Gln Glu
385 390

<210> 335
<211> 428
<212> PRT
<213> Lupinus angustifolius

<400> 335

Met Ala Ser Ser Ser Ser Asp Pro Gly Lys Ser Glu Ile Gly Gly Gly
1 5 10 15

Ala Ala Glu Thr Ser Glu Ala Ala Ala Val Ala Val Ala Val Thr Asn
20 25 30

Asp Gln Ser Leu Leu Tyr Arg Gly Leu Lys Lys Ala Lys Lys Glu Arg
35 40 45

Gly Cys Thr Ala Lys Glu Arg Ile Ser Lys Met Pro Pro Cys Ala Ala
50 55 60

Gly Lys Arg Ser Ser Ile Tyr Arg Gly Val Thr Arg His Arg Trp Thr
65 70 75 80

Gly Arg Tyr Glu Ala His Leu Arg Asp Lys Ser Thr Trp Asn Gln Asn
85 90 95

Gln Asn Lys Lys Gly Lys Gln Val Tyr Leu Gly Ala Tyr Asp Asp Glu
100 105 110

pctau2012001598-seql

Glu Ala Ala Ala Arg Ala Tyr Asp Leu Ala Ala Leu Lys Tyr Trp Gly
115 120 125
Pro Gly Thr Leu Ile Asn Phe Pro Val Thr Asp Tyr Thr Arg Asp Leu
130 135 140
Glu Glu Met Gln Asn Val Ser Arg Glu Glu Tyr Leu Ala Ser Leu Arg
145 150 155 160
Arg Lys Ser Ser Gly Phe Ser Arg Gly Ile Ser Lys Tyr Arg Ala Leu
165 170 175
Ser Ser Arg Trp Glu Pro Ser Tyr Ser Arg Phe Ala Gly Ser Asp Tyr
180 185 190
Phe Asn Ser Met His Tyr Gly Ala Gly Asp Asp Ser Ala Ala Glu Ser
195 200 205
Glu Tyr Ala Ser Gly Phe Cys Ile Glu Arg Lys Ile Asp Leu Thr Gly
210 215 220
His Ile Lys Trp Trp Gly Ser Asn Lys Ser Arg Gln Pro Asp Ala Gly
225 230 235 240
Thr Arg Leu Ser Glu Glu Lys Arg His Gly Phe Ala Gly Asp Ile Cys
245 250 255
Ser Glu Pro Lys Thr Leu Glu Gln Lys Val Gln Pro Thr Glu Pro Tyr
260 265 270
Gln Met Pro Glu Leu Gly Arg Ser His Asn Glu Lys Lys His Arg Ser
275 280 285
Ser Ala Val Ser Ala Leu Ser Ile Leu Ser Gln Ser Ala Ala Tyr Lys
290 295 300
Ser Leu Gln Glu Lys Ala Ser Lys Lys Gln Glu Asn Ser Thr Asp Asn
305 310 315 320
Asp Glu Asn Glu Asn Lys Asn Thr Val Asn Lys Leu Asp His Gly Lys
325 330 335
Ala Val Glu Lys Ser Ser Asn His Asp Gly Gly Ser Asp Arg Val Asp
340 345 350
Ile Glu Ile Gly Thr Thr Gly Ala Leu Ser Leu Gln Arg Asn Ile Tyr
355 360 365
Pro Leu Thr Pro Phe Leu Ser Ala Pro Leu Leu Thr Ala Tyr Asn Thr
370 375 380

pctau2012001598-seql

Val Asp Pro Ser Leu Val Asp Pro Val Leu Trp Thr Ser Leu Val Pro
385 390 395 400

Met Leu Ser Ala Gly Leu Ser Cys Pro Thr Gln Val Thr Lys Thr Glu
405 410 415

Thr Ser Ser Ser Tyr Thr Ile Phe Gln Pro Glu Gly
420 425

<210> 336
<211> 440
<212> PRT
<213> Ricinus communis
<400> 336

Met Ala Ser Ser Ser Ser Asp Pro Gly Leu Lys Pro Glu Leu Gly Gly
1 5 10 15

Gly Ser Gly Gly Glu Ser Ser Glu Ala Val Ile Ala Asn Asp Gln Leu
20 25 30

Leu Leu Tyr Arg Gln Leu Lys Lys Pro Lys Lys Glu Arg Gly Cys Thr
35 40 45

Ala Lys Glu Arg Ile Ser Lys Met Pro Pro Cys Thr Ala Gly Lys Arg
50 55 60

Ser Ser Ile Tyr Arg Gly Val Thr Arg His Arg Trp Thr Gly Arg Tyr
65 70 75 80

Glu Ala His Leu Trp Asp Lys Ser Thr Trp Asn Gln Asn Gln Asn Lys
85 90 95

Lys Gly Lys Gln Gly Ala Tyr Asp Asp Glu Glu Ala Ala Ala Arg Ala
100 105 110

Tyr Asp Leu Ala Ala Leu Lys Tyr Trp Gly Pro Gly Thr Leu Ile Asn
115 120 125

Phe Pro Val Thr Asp Tyr Ser Arg Asp Leu Glu Glu Met Gln Asn Val
130 135 140

Ser Arg Glu Glu Tyr Leu Ala Ser Leu Arg Arg Lys Ser Ser Gly Phe
145 150 155 160

Ser Arg Gly Ile Ser Lys Tyr Arg Gly Leu Ser Ser Gln Trp Asp Ser
165 170 175

Ser Phe Gly Arg Met Pro Gly Ser Glu Tyr Phe Ser Ser Ile Asn Tyr
180 185 190

pctau2012001598-seql

Gly Ala Ala Asp Asp Pro Ala Ala Glu Ser Glu Tyr Val Gly Ser Leu
195 200 205

Cys Phe Glu Arg Lys Ile Asp Leu Thr Ser Tyr Ile Arg Trp Trp Gly
210 215 220

Phe Asn Lys Thr Arg Glu Ser Val Ser Lys Ser Ser Asp Glu Arg Lys
225 230 235 240

His Gly Tyr Gly Glu Asp Ile Ser Glu Leu Lys Ser Ser Glu Trp Ala
245 250 255

Val Gln Ser Thr Glu Pro Tyr Gln Met Pro Arg Leu Gly Met Pro Asp
260 265 270

Asn Gly Lys Lys His Lys Cys Ser Lys Ile Ser Ala Leu Ser Ile Leu
275 280 285

Ser His Ser Ala Ala Tyr Lys Asn Leu Gln Glu Lys Ala Ser Lys Lys
290 295 300

Gln Glu Asn Cys Thr Asp Asn Asp Glu Lys Glu Asn Lys Lys Thr Asn
305 310 315 320

Lys Met Asp Tyr Gly Lys Ala Val Glu Lys Ser Thr Ser His Asp Gly
325 330 335

Ser Asn Glu Arg Leu Gly Ala Ala Leu Gly Met Ser Gly Gly Leu Ser
340 345 350

Leu Gln Arg Asn Ala Tyr Gln Leu Ala Pro Phe Leu Ser Ala Pro Leu
355 360 365

Leu Thr Asn Tyr Asn Ala Ile Asp Pro Leu Val Asp Pro Ile Leu Trp
370 375 380

Thr Ser Leu Val Pro Val Leu Pro Ala Gly Phe Ser Arg Asn Ser Glu
385 390 395 400

Val Gly Met Gly Leu Gln Ile Val Ser Cys His Lys Asp Arg Asp Lys
405 410 415

Phe Asn Leu Tyr Leu Leu Ser Ala Gly Gly Val Ser Thr Phe Leu Leu
420 425 430

Leu Val Val His Trp Arg Phe Cys
435 440

<210> 337
<211> 428
<212> PRT
<213> Lupinus angustifolius

pctau2012001598-seql

<400> 337

```

Met  Ala  Ser  Ser  Ser  Ser  Asp  Pro  Gly  Lys  Ser  Glu  Ile  Gly  Gly  Gly
 1          5          10          15

Ala  Ala  Glu  Thr  Ser  Glu  Ala  Ala  Ala  Val  Ala  Val  Ala  Val  Thr  Asn
          20          25          30

Asp  Gln  Ser  Leu  Leu  Tyr  Arg  Gly  Leu  Lys  Lys  Ala  Lys  Lys  Glu  Arg
          35          40          45

Gly  Cys  Thr  Ala  Lys  Glu  Arg  Ile  Ser  Lys  Met  Pro  Pro  Cys  Ala  Ala
          50          55          60

Gly  Lys  Arg  Ser  Ser  Ile  Tyr  Arg  Gly  Val  Thr  Arg  His  Arg  Trp  Thr
65          70          75          80

Gly  Arg  Tyr  Glu  Ala  His  Leu  Arg  Asp  Lys  Ser  Thr  Trp  Asn  Gln  Asn
          85          90          95

Gln  Asn  Lys  Lys  Gly  Lys  Gln  Val  Tyr  Leu  Gly  Ala  Tyr  Asp  Asp  Glu
          100          105          110

Glu  Ala  Ala  Ala  Arg  Ala  Tyr  Asp  Leu  Ala  Ala  Leu  Lys  Tyr  Trp  Gly
          115          120          125

Pro  Gly  Thr  Leu  Ile  Asn  Phe  Pro  Val  Thr  Asp  Tyr  Thr  Arg  Asp  Leu
          130          135          140

Glu  Glu  Met  Gln  Asn  Val  Ser  Arg  Glu  Glu  Tyr  Leu  Ala  Ser  Leu  Arg
145          150          155          160

Arg  Lys  Ser  Ser  Gly  Phe  Ser  Arg  Gly  Ile  Ser  Lys  Tyr  Arg  Ala  Leu
          165          170          175

Ser  Ser  Arg  Trp  Glu  Pro  Ser  Tyr  Ser  Arg  Phe  Ala  Gly  Ser  Asp  Tyr
          180          185          190

Phe  Asn  Ser  Met  His  Tyr  Gly  Ala  Gly  Asp  Asp  Ser  Ala  Ala  Glu  Ser
          195          200          205

Glu  Tyr  Ala  Ser  Gly  Phe  Cys  Ile  Glu  Arg  Lys  Ile  Asp  Leu  Thr  Gly
          210          215          220

His  Ile  Lys  Trp  Trp  Gly  Ser  Asn  Lys  Ser  Arg  Gln  Pro  Asp  Ala  Gly
225          230          235          240

Thr  Arg  Leu  Ser  Glu  Glu  Lys  Arg  His  Gly  Phe  Ala  Gly  Asp  Ile  Cys
          245          250          255

Ser  Glu  Pro  Lys  Thr  Leu  Glu  Gln  Lys  Val  Gln  Pro  Thr  Glu  Pro  Tyr

```

pctau2012001598-seql

260

265

270

Gln Met Pro Glu Leu Gly Arg Ser His Asn Glu Lys Lys His Arg Ser
 275 280 285

Ser Ala Val Ser Ala Leu Ser Ile Leu Ser Gln Ser Ala Ala Tyr Lys
 290 295 300

Ser Leu Gln Glu Lys Ala Ser Lys Lys Gln Glu Asn Ser Thr Asp Asn
 305 310 315 320

Asp Glu Asn Glu Asn Lys Asn Thr Val Asn Lys Leu Asp His Gly Lys
 325 330 335

Ala Val Glu Lys Ser Ser Asn His Asp Gly Gly Ser Asp Arg Val Asp
 340 345 350

Ile Glu Ile Gly Thr Thr Gly Ala Leu Ser Leu Gln Arg Asn Ile Tyr
 355 360 365

Pro Leu Thr Pro Phe Leu Ser Ala Pro Leu Leu Thr Ala Tyr Asn Thr
 370 375 380

Val Asp Pro Ser Leu Val Asp Pro Val Leu Trp Thr Ser Leu Val Pro
 385 390 395 400

Met Leu Ser Ala Gly Leu Ser Cys Pro Thr Gln Val Thr Lys Thr Glu
 405 410 415

Thr Ser Ser Ser Tyr Thr Ile Phe Gln Pro Glu Gly
 420 425

<210> 338

<211> 1527

<212> DNA

<213> Aspergillus fumigatus

<400> 338

atgaagatga gtgcttccaa aacagtaact tcgtcagcca gtgccgtag caccagctct 60

ggaaggtcga ctccgtcgaa gctcgtcaat ggagctaccc gcaatggctc cgcagcagct 120

ggcaacggct ctaccgggac agccaaaggg aagaggagga gcaaataatcg ccatgttgca 180

gcttatcatt cagagctgcg gcattcgagc ctcagtagag agaccagtgt ggtcccaagc 240

tttcttgat tccggaatct catggtgatt gttttagtag cgatgaatct tcgcctgata 300

attgaaaatt tcatgaaata tggcgtcctt atctgcatca aatgtcatga ctatcggaag 360

caggatgtag ttcttgggtc gattctcttt gcccttggtc catgccatct cttccttgcg 420

tacatcattg aattggttgc agctcagcaa tccaagaaaa ctgtcggtcg acagaaaaag 480

gatctatcga cggaggagag ggagcgcgaa caacaggcct tccgctcaac ctggaggat 540

acagccttct tccataccgt gaacgtacc ctttgctcg ccgtgaccag cttcgttgtt 600

pctau2012001598-seql

tatttctaca tcaatcatcc tggaatcggg acaatttgtg agctccacgc aatcattgtg	660
tgggtcaaaa attgctccta tgcatttacg aatcgcgatc tgcggcaggc aatgggtcgac	720
ccctctgcag agtctgctct cccagaaatc tactccacat gcccatatcc aagaaatata	780
acacttggga accttacata tttttggcta gcaccgacac ttgtatatca acccgtttat	840
cctcgatcgt cgcataattcg atgggtctttt gttgccaaac ggcttgctga gttttttggt	900
ttagctgtct tcatatggct tctctcggct cagtatgcag ccccggtctt gcgaaattcc	960
attgacaaga tagctgtgat ggatattgcg tcaatcctcg agcgggtcat gaaactctcc	1020
acaatttcgc ttgtcatttg gctcgttgga ttctttgcgc tcttccagtc gctgctgaat	1080
gctttggcgg aggtcatgcg gttcggagat cgtgagttct acacagattg gtggaacagc	1140
ccgagtctcg gcgcatactg gcggtcatgg aacaggccgg tgtacctctt catgaagagg	1200
catgttttct cgccgctggg aggaagaggc tggagccctt ttgcggcgag cttcatggtc	1260
ttttccctct ctgcagtcct gcatgagatg ttggtgggga ttcctacca caatttgatc	1320
ggtgttgcat ttgcggggat gatgttcag ctgccattga ttgctgtgac agctcccttt	1380
gaaaagggtga atgacgcctt gggcaagatt gtagggaact caatcttctg ggtaagcttt	1440
tgcctcgttg gtcaaccgct ggggtgcgctc ttgtatttct ttgcatggca ggcgaagtac	1500
ggaagtgtca gcaagataca cgtttag	1527

<210> 339
 <211> 2067
 <212> DNA
 <213> Ricinus communis

<400> 339	
ggacactgac atggactgaa ggagtagaaa cttcttcttc tgctgctggt cctctctcct	60
ccaccgccac gtcacctct ctttgcata cataatacta ttgttcttat tatcattttc	120
actctttaaa tacaacatc aattattcct tttctatcaa acacatgtat tctattctct	180
cgctgtctag attctcatct tcattgaatc ctccttctta gcgtgtcttt gtccacttct	240
tttgggcacc gacgttttta atctccatga cgattctcga aacgccagaa actcttggcg	300
tcctctctc ctccgccact tccgatctca acctctctct cgcagctaga cggacctcaa	360
atgactccga tgggtgcactt gctgatttgg cttcgaagtt tgatgatgat gacgacgtaa	420
gatcggaaga ttctgctgaa aatattatcg aagatcctgt agcagcgggt actgaattgg	480
cgacagcaaa gagtaacgga aaagactgtg ttgccaatag taataaggat aaaattgata	540
gccatggagg atcatcggtat tttaaacttg catataggcc ttcggttcca gtcaccgggt	600
cacttaagga gagtccgctt agctctgatt taatatttaa acaaagtcac gcaggctctgt	660
ttaacctttg tatagtagtg ctcgtagctg ttaacagcag gctcatcatt gagaatttaa	720
tgaagtatgg ctgggttaatt aagacgggct tttggtttag ttcaagatca ttgagagatt	780
ggccgctttt tatgtgctgt ctttctctcc cagtattccc ccttgctgcc tatctagttg	840

pctau2012001598-seq1

agaaggccgc	atatcgaaaa	tatatatctc	cgcctattgt	tattttcctt	catgtgatca	900
tcacctcagc	agctgttttg	taccagcgtt	ctgtaattct	cagttgtgaa	tctgcttttt	960
tatctggtgt	cacattgatg	gaacttgctt	gtatggtatg	gttgaaattg	gtatcctatg	1020
cacatacaaa	ctatgatatg	agagcgatcg	ctgacaccat	tcataaggaa	gatgcatcca	1080
attcttctag	tacagagtat	tgtcatgatg	tgagctttta	gactttggcg	tacttcatgg	1140
tcgcaccac	attatgttac	cagccaagtt	atcctcgcac	agcatttatt	agaaagggct	1200
gggtgttccg	tcaatttgtc	aaactaataa	tttttacagg	attcatggga	tttatcatag	1260
aacaatacat	caatcctatc	gtccagaatt	ctcaacaccc	tttaaaagg	gatctcttat	1320
atgccattga	gagggttctg	aagctctcag	ttccgaattt	atatgtgtgg	ctctgcttgt	1380
tctactgctt	ttttcacctg	tggttgaata	tagttgctga	gctccttcgc	ttcggtgacc	1440
gggagtctta	caaagattgg	tggaatgcaa	aaactgttga	ggagtactgg	aggatgtgga	1500
atatgcctgt	tcacaagtgg	atggttcgcc	atatctactt	cccatgccta	cgtcgtaaaa	1560
taccaagggg	ggtagcaata	gttattgctt	tcttcgtttc	agctgtattt	catgagttgt	1620
gcattgctgt	tccttgccac	atgttcaaac	tttgggcttt	ttttggaata	atgtttcaga	1680
ttcctttagt	tgtgatcact	aattattttc	aaaggaagtt	cagaagctca	atgggtggga	1740
atatgatctt	ctggttcttt	ttctgcattc	tcggccaacc	tatgtgtgta	ctgttgattt	1800
accatgacct	aatgaatcgc	gatgggaact	gaaccatggg	ctcagtcag	atatgggtac	1860
acctccaag	atgttatttt	cgtgagtga	gactgcacca	cagtgttggt	cttgttacac	1920
aatccccatt	gacagagtag	gttaatcgtc	agtttcagga	gataagacac	aattttgaaa	1980
gtacagcaga	ggctgctatt	aatgtatcat	gttgagtttc	tgttatgtta	tgttattctt	2040
ttttaatctc	aaaaaaaaaa	aaaaaaa				2067

<210> 340
 <211> 2084
 <212> DNA
 <213> Vernicia fordii

<400> 340	
tctctctctc	tttgctttac
gtgtacatcg	accaccacca
cagccatctt	gcgactgttc
	60
aattatccta	taagtaccac
cgcattcatc	accgccaatc
cttaactcta	atgtgctata
	120
ctaaacactt	gctttatatg
cgcttttcta	tttactcttc
actgtaattt	cttattggta
	180
ttcaaagtgt	tttcaatgac
aatccctgaa	acgccggata
attccacgga	tgctaccacc
	240
agtggcggtg	ctgagtcctc
ttccgatctt	aacctttctc
ttcgacggag	gaggactgct
	300
tcaaactccg	atggagctgt
cgcggaattg	gcttccaaga
ttgatgagtt	ggaatctgat
	360
gccggaggag	gccaggtgat
taaggatccg	ggagcagaaa
tggattcggg	gactttgaaa
	420
agtaatggaa	aagattgcgg
aaccgttaag	gataggattg
aaaatcgtga	gaatcgtgga
	480
ggatcggatg	ttaaattcac
gtatcgccg	tcggtgccgg
ctcaccgggc	gctcaaggag
	540
agtccgctta	gctctgataa
tatatttaaa	caaagtcatg
cagggtctctt	caatctctgt
	600

pctau2012001598-seql

atagtagtgc	ttgtagcgg	taacagtcgg	cttatcattg	aaaacataat	gaagtatggt	660
tggttaatta	agactgggtt	ttggtttagt	tcaagatcgt	tgagagactg	gccacttctt	720
atgtgctgtc	ttaccctccc	aatatcttct	cttgctgcct	atctagttga	gaagttggcc	780
tgtcgggaagt	atatactgc	accactgtt	gtttttcttc	atattctttt	ctcctcaaca	840
gcagttttat	accctgtttc	tgtgattctc	agttgtgaat	ctgctgtttt	gtccggtgtc	900
gcattgatgc	tctttgcttg	tatcgtgtgg	ttgaaattgg	tatcttatgc	acatacaaac	960
tttgatatga	gagcaattgc	taactcagtt	gataaggag	atgcgctatc	caatgcttcg	1020
agtgcagagt	cctctcatga	tgtagcttc	aagagtttgg	tttatttcat	ggttgctccc	1080
acattgtgtt	accagccaag	ttatcctcga	actgcatcca	ttcgaaagg	ttgggtggtt	1140
cgtcaatttg	ttaagttaat	aatatttaca	ggattcatgg	gatttatcat	agaacaatat	1200
atcaatccta	tcgttcagaa	ttcacaacat	cctttaaaag	gggatctctt	atatgccatt	1260
gagagggttt	tgaagctctc	agttccgaat	ttatatgtct	ggctttgcat	gttctactgc	1320
ttttttcacc	tatggtaaaa	tatacttgct	gagctccttc	gctttggtga	tagagagttc	1380
tataaagatt	ggtggaatgc	aaggacagtt	gaggagtatt	ggagaatgtg	gaatatgcct	1440
gttcataagt	ggatggttcg	ccatatctac	tttccatgct	tgccgcataa	aataccaagg	1500
ggggtggcct	tattaattac	tttcttcgtt	tcagcagtat	ttcatgagtt	gtgcattgct	1560
gttccttgcc	acatattcaa	gctctgggct	tttattggaa	taatgtttca	gattcctttg	1620
gtcgggatca	caaattacct	tcaaaacaag	ttcagaagct	caatggtggg	gaacatgac	1680
ttctggttca	ttttctgcat	tcttggtcaa	ccaatgtgct	tgctgttgta	ttaccatgac	1740
ctaataatc	gaaaagggac	taccgaatca	agatgacact	aactcatcgt	gtggtagact	1800
ctatatatat	acatagactt	accagagatg	ggttgcttcc	aacatattgt	gcacaagagg	1860
caattgttgt	tctcatcaga	agagtgggtt	aattaattaa	ttaatgtaca	agcaattttg	1920
aaagtataat	cactggcagg	gactagtgcc	cgactgtagt	actgagatta	tagaggtatt	1980
atcaatcgtt	agtggaaaat	tgtaaatgta	taaagttcaa	tctttgtatt	gtttcttttc	2040
taatatcata	ttttttttta	ttgctcatca	aaaaaaaaaa	aaaa		2084

<210> 341

<211> 1828

<212> DNA

<213> Vernonia galamensis

<400> 341

tctgagctca	aatcaaattt	ctgcgactca	tacaggattc	aactcaatac	tttcttgatc	60
ggttctgctg	ttcatttact	tgtaatttct	acttctgctt	tgctttcatt	tcaagctttt	120
ttccttaata	atggcgttat	tagatacgcc	tcagattgga	gaaataacga	ccaccgccac	180
cacaactata	agacggcgga	ccactgtcaa	gcctgatgct	ggaatcggag	atggattgtt	240
tgattcttcg	tcgtcttcca	aaaccaactc	atccttcgag	gatgggtgaca	gtttgaatgg	300

pctau2012001598-seq1

tgatttcaat gacaaattta aggaacagat cggagctggt gatgaatcca aggacgactc	360
caaggggaac ggacagaaga tagatcacgg aggagttaaa aagggacgtg aaacgactgt	420
ggtgcattat gcttatcggc cttcttctcc ggctcatcgg agaattaaag aatctccgct	480
tagctctgac gccatcttca agcagagtca tgcaggcctc tttaaccttt gcatagtggg	540
gcttggtgca gtaaatggta ggctcatcat tgagaatctg atgaagtatg gactattgat	600
caattccaac ttttggttca gttcgagatc attgagagac tggccacttc tgatgtgctg	660
cctcactcct tctgactttc cacttgctgc ctacattggt gagaaattgg catggaaaaa	720
acgtatatcc gaccctgttg taatcacact ccatgttata ataactacaa ctgcaattct	780
ttatccggtc ttcatgattc tgaggttcga ttcagttggt ctatcaggcg tctcgttgat	840
gctgtgtgct tgcattaatt ggttgaagtt ggtatctttt gtgcatacaa attatgacat	900
gcggtcgctt ttgaactcaa ctgataaggg agaagtggaa cccatgtctt caaatatgga	960
ttatttttat gatgtcaact tcaaaagctt ggtttatttc atggttgctc caactttgtg	1020
ttaccagata agctatcctc gcactgcatt tattcgaaag ggttggtgtg tacggcaact	1080
gatcaagcta gtaatattta cagggttcat gggattcatc attgaacaat atatcaatcc	1140
gattgtcaaa aattctcgtc atccattgaa aggagacttt ttatatgca ttgagcgggt	1200
tttaaagctt tcagttccga atttatatgt gtggctctgt atgttctact gcttttttca	1260
cctttggtta aatatacttg ctgagcttct ttgttttggg gatcgtgaat ttataaaga	1320
ttggtggaat gcacaaacta ttgaagagta ttggaggcta tggaatatgc ctgttcataa	1380
atggattgtt aggcaccttt attttccatg cttgcgtaat gggataccta aggggtgctgc	1440
catattgggtt gcatttttca tgtctgccgt gttccatgag ctttgtattg ctgttccctg	1500
ccacattttc aagttttggg cttttatcgg gatcatgttt cagggtcccgt tggtcctact	1560
cacaaattac ttgcagcaca agtttcaaaa ctcgatgggt ggaaatatga tcttctgggtg	1620
ctttttcagc atttttggtc aacctatgtg tgtattactt tactaccatg atgtcatgaa	1680
tcaaaagggg aaaagcaaat aaaaagatgt gattgtgttg ctccatttga tctcatagca	1740
tgactggact aaacaaacc aagggacaca ttttagtcct taaaggaaaa tttttgtagg	1800
aaaaaaaaa aaaaaaaaaa aaaaaaaaa	1828

<210> 342
<211> 1738

<212> DNA
<213> Vernonia galamensis

<400> 342	
gttcgtaatt cggctgtggt ttcctttcca acatttctac gtaatcatgg cgttgttaga	60
tacgcctcaa attggagaaa taacgacgac cgcaacaacg accattaggc agcaccctt	120
gggcaagcct gatgctggaa ttggagatgg attgttttct tcgtcgtctt ccaaaaccaa	180
ctcatccttc gaggatgggt acagtttgaa tggtgatttc aatgacaaat ttaaggaaca	240
gatcggagct ggtgatgaat ccaagaaggg gaacggaaag atagatcacg gaggagttaa	300

pctau2012001598-seql

aaagggacgt gaaacgactg tgggtgatta tgcttatcgg ccttcttctc cggctcatcg	360
gagaattaaa gaatctccgc ttagctctga cgccatcttc aagcagagtc atgcaggcct	420
ctttaacctt tgcatagtgg tgcttggtgc agtaaattgg aggctcatca tcgagaatct	480
gatgaagtat ggactattga ttaattccaa attttggttc agttcgagat cattgagaga	540
ctggccgctt ctgatgtgtt ggctgacccc ctccgacttc cccctcgccg cctacattgt	600
cgagaaattg gcatggaaaa aacgtatatc cgaccctggt gtaatcacac tccatgttgt	660
aataactaca actgcaattc tctatccgat cttcatgatt ctgaggttcg actcggtcgt	720
tctattaggc gtctcgttga tgctgtgtgc ttgcattaat tggttgaagt tggatatctt	780
tgtgcataca aattatgaca tgcggtcgct attgaactca actggtaagg gagaagtgga	840
gcccattgtc tcaaatatgg actactttta tgatatcaac ttcaaaagct tggtttattt	900
catggttgct ccaactttgt gttaccagat aagctatcct cgcaccgcct ttattcgaaa	960
gggctgggtg ttccggcaac tgatcaagct agtaatatct acagggttca tgggattcat	1020
cattgaacaa tatatcaatc cgattgtcaa aaattctcgg catccattga acggagactt	1080
tttatatgcg attgaacgag tattaaggt ttcagttccg aatttatatg tgtggctctg	1140
tatgttctat tgcttttttc acctttggtt aaatatactt gctgagcttc tttggtttgg	1200
ggatcgtgaa tttataaag attggtggaa tacacaaact attgaagagt attggaggct	1260
atggaatatg cctgttcata agtggattgt taggcacctc tattttccat gcttgcgtaa	1320
tgggatatct aagggtgctg ccatattggt tgcttttttc atgtctgccg tgttccacga	1380
gctttgcata gctgttcctt gccacatttt aaagttttgg gctttcatcg ggatcatggt	1440
ccaggctccg ttggtactac tcacaaatta cttgcagcac aagtttcaaa actcgatggt	1500
gggaaacatg atcttttggg gcttcttcag ctttttcggt caacccatgt gtgtatttct	1560
ttactaccat gaagtcaatc aaaaggggaa aagcaaatga aaggacgtta tcgtatttcc	1620
ccaatctttc ttatatcgtg aatctaatat ccataacaaa gcaaaacaat taagtcactg	1680
gagaatacta ttagcaggta ataaagaacc aaacaaaaaa aaaaaaaaaa aaaaaaaa	1738

<210> 343
 <211> 1524
 <212> DNA
 <213> Euonymus alatus

<400> 343	
atggctgcta acttgaacga agcctcggat cttaatTTTT cgcttcggag gagaactggt	60
ggcatctcaa gtacgactgt gcctgattct agttccgaga caagttcgtc ggaggcggac	120
tatttgacg gaggcaaagg tgccgcggac gtcaaagatc gtggggatgg tgcggtggag	180
tttcagaatt cgatgaagaa cgtggagagg attgagaaac atgagagccg agtaggattg	240
gattcgagat tcacgtatag gccatcggtc ccggctcatc gcacaataaa ggagagcccg	300
cttagctcgg acgcaatatt caaacagagt cacgcaggtc tcttcaatct ctgtatagta	360

pctau2012001598-seq1

gttctggttg	ctgtgaacag	caggctgata	attgaaaatc	taatgaagta	tggatggtta	420
ataaggagtg	ggttttggtt	cagctcaaga	tcattgagag	actggcccct	ttttatgtgt	480
tgtctcacac	taccagtatt	ccctcttgct	gcttttctgt	ttgagaagtt	ggctcaaaaa	540
aatttaatat	ctgaacctgt	tgttgttttg	cttcatatag	taaacactac	agctgccggt	600
ttataccctg	ttttggtgat	tctaagggtg	gattctgcct	ttatgtctgg	ggttacgttg	660
atgctctttg	cttgtattgt	gtgggttaaag	ttggatatctt	atgctcatac	caactatgat	720
atgagagccc	tcaccaagtc	tgttgaaaag	ggggacacgc	cgttgagctc	tcagaacatg	780
gattactcgt	ttgatgtcaa	tatcaagagt	ttggcatatt	ttatggttgc	tcccacatta	840
tgttaccaga	ttagctatcc	tcgtaccca	tatgttcgca	agggttgggt	ggttcgtcaa	900
tttgtcaagt	taataatatt	tactggactt	atgggattta	taattgaaca	atatatcaat	960
cctattgtcc	agaattcaca	acaccctttg	aaaggaaact	ttttgtatgc	cattgagaga	1020
gttttgaagc	tttcagtcct	aaacctttat	gtttggctct	gcatgttcta	ctgccttttt	1080
catctctggt	taaacatact	tgctgaactt	ctttgttttg	gtgatcgtga	gttctacaag	1140
gattggtgga	acgctaaaac	tgttgaggag	tactggagaa	tgtggaacat	gcctgttcat	1200
aagtggatgg	ttcgtcatat	ctacttccca	tgcttgagga	acgggatacc	caagggtggt	1260
gcttttgtca	tttccttctt	agtttctgcc	gtcttccatg	agctatgcat	tgctgttccc	1320
tgccacatct	tcaagttatg	ggctttcttt	ggaataatgc	ttcaggttcc	cttgggtgtg	1380
atcacaagtt	atctgcaaaa	taagttcaga	agctcaatgg	tgggaaatat	gatgttttgg	1440
ttctctttct	gcatttttgg	tcaacctatg	tgcttacttc	tatattacca	tgatttgatg	1500
aatcgcaatg	ggaagatgga	gtag				1524

<210> 344
 <211> 1497
 <212> DNA
 <213> Caenorhabditis elegans

<400> 344						
atgcaaatgc	gtcaacaaac	gggacgacgg	cggcgtcagc	cttcggaaac	atctaattggt	60
tctttggctt	ccagtagacg	ctcctcattt	gcacaaaatg	gtaattcgtc	aaggaaaagt	120
tcagaaatga	gagggccttg	cgagaaagtg	gtacatactg	ctcaagattc	attgttttcg	180
acgagttctg	gatggacaaa	tttccgtgga	ttcttcaatt	tgtctatttt	acttttggtg	240
ctttcaaatg	gacgcgtggc	acttgaaaat	gtgatcaaat	atggatattt	gataacaccc	300
cttcagtgga	tctcaacgtt	tgttgagcat	cactactcaa	tttgagctg	gccaaatctt	360
gctctcatcc	tatgctcaaa	tattcagatt	ctctcggttt	ttggaatgga	aaaaattctt	420
gaacgtggat	ggcttggaaa	cggattcgct	gcagtgttct	acacctcgct	tgtgattgca	480
catctgacaa	ttccagttgt	ggtcactctt	accacaaaat	ggaagaatcc	tttgtggtca	540
gtcgtaatga	tgggtgttta	tgttattgaa	gctctcaaat	tcatttcata	tggccacgtc	600
aactactggg	ctcgtgatgc	tcggcgaaaa	atcacagagc	tcaaaacaca	agtcaccgat	660

pctau2012001598-seq1

ttggcaaaga aaacatgtga tccgaaacaa ttttgggatt tgaaagatga attatcaatg	720
catcagatgg ctgctcaata tcctgccaat ttgacacttt ccaatatcta ctacttcatg	780
gctgcaccaa cattgtgcta cgaattcaaa tttccaagat tgttgcgaat tcggaagcac	840
tttttgatta aaagaaccgt ggagcttata tttctatcgt ttttgatagc tgcacttggt	900
caacaatggg ttgttccgac tgtccgaaat agtatgaaac ctttaagtga aatggaatac	960
tctagatggt tggaacgact cttgaaactt gcaattccaa atcatctcat ctggcttcta	1020
ttcttctaca cattcttcca ttcatttttg aacttgatcg ccgagctgct tcgatttgcc	1080
gatcgtgagt tctacagaga cttttggaat gcagagacga taggatattt ctggaaatca	1140
tggaacatcc cagttcaccg atttgctggt cgccacatct acagtccaat gatgcgtaac	1200
aatttctcaa aaatgagcgc attcttcggt gtgttcttcg tgtcggcatt cttccatgaa	1260
tatctgggtt ctgtgccatt aaagattttc cgattgtggt cctactatgg aatgatggga	1320
caaattcctc tatccattat cactgataaa gtggtgagag gtggacgtac aggaaacatc	1380
atcgtctggc tctcactgat tgttggccaa cctcttgcaa ttctcatgta cggacatgat	1440
tggtacattt tgaactttgg tgtttcagca gttcaaaacc aaaccgttgg tatttga	1497

<210> 345
 <211> 1751
 <212> DNA
 <213> Rattus norvegicus

<400> 345	
gaagattggg ccgcgacgag gtgcggggccg aagccatggg cgaccgcgga ggcgcgggaa	60
gctctcggcg tcgcaggacc ggctcgcggg tttccgtcca gggaggtagt gggcccaagg	120
tagaagagga cgaggcgca gaagcggctg tgagccccga cttgggcgcc gggggtgacg	180
cgccggctcc ggctccggct ccagcccata cccgggacaa agaccggcag accagcgtgg	240
gcgacggcca ctgggagctg aggtgccatc gtctgcaaga ctctttgttc agctcagaca	300
gcggtttcag caattaccgt ggtatcctga attggtgcgt ggtgatgctg atcctgagta	360
atgcaagggt atctttagag aatcttatca agtatggcat cctggtggat cccatccagg	420
tggtgtctct gtttctgaag gaccctaca gctggcctgc cccatgcttg atcattgcat	480
ccaatatctt tattgtggct acatttcaga ttgagaagcg cctgtcagtg ggtgccctga	540
cagagcagat ggggctgctg ctacatgtgg ttaacctggc cacaattatc tgcttcccag	600
cagctgtggc ctactgggt gagtctatca ctccagtggg ttccctgttt gctctggcat	660
catactccat catcttcctc aagctttctt cctaccggga tgtcaatctg tggtgccgcc	720
agcgaagggt caaggccaaa gctgtgtctg cagggaagaa ggtcagtggg gctgctgccc	780
agaacactgt aagctatccg gacaacctga cctaccgaga tctctattac ttcatctttg	840
ctcctacttt gtgttatgaa ctcaactttc ctcgatcccc ccgaatacga aagcgctttc	900
tgctacggcg ggttcttgag atgctctttt tcaccagct tcaagtggg ctgatccagc	960

pctau2012001598-seq1

agtggatggt ccctactatc cagaactcca tgaagccctt caaggacatg gactattcac	1020
gaatcattga gcgtctctta aagctggcgg tccccaacca tctgatatgg ctcatcttct	1080
tctattggct tttccactca tgtctcaatg ctgtggcaga gctcctgcag tttggagacc	1140
gcgagtctta cagggactgg tggaatgctg agtctgtcac ctacttttgg cagaactgga	1200
atatccccgt gcacaagtgg tgcacagac actttttacaa gcctatgctc agactgggca	1260
gcaacaaatg gatggccagg actgggggtct tttgggcgtc agctttcttt catgagtacc	1320
tagtgagcat tcccctgagg atgttccgcc tttgggcatt cacagcaatg atggctcagg	1380
tcccactggc ctggattgtg aaccgcttct tccaaggga ctatggcaat gcagctgtgt	1440
gggtgacact catcattggg caaccggtgg ctgtgctcat gtatgtccac gactactacg	1500
tgctcaacta tgatgcccc gtgggggcat gagctactgc caaaggccag cctccctaac	1560
ctgggcctgg agttctggag ggcttcctgc tgctgcacac tcccctagtt tggaggcctt	1620
tctgccccta tggggcctac tcctgctctt ggggatggcc ctgagccagc tggttgagcc	1680
agtgtggga gtttgctg accaggggct taggatatca ataaagagct atctaaaaa	1740
aaaaaaaaa a	1751

<210> 346
 <211> 3718
 <212> DNA
 <213> Homo sapiens

<400> 346	
cgcctcaggc cggccgcgga ctacaaatgg acgagagagg cggccgtcca ttagttagcg	60
gctccggagc aacgcagccg ttgtccttga ggccgacggg cctgacgcgg gcgggttgaa	120
cgcgctggtg aggcggtcac ccgggctacg gcggccggca gggggcagtg gcggccgttg	180
tctagggccc ggaggtgggg ccgcgcgcct cgggcgctac gaaccggca ggcccacgct	240
tggctgcggc cgggtgcggg ctgaggccat gggcgaccgc ggcagctccc ggcgccggag	300
gacagggctg cggccctcga gccacggcgg cggcgggcct gcggcggcgg aagaggaggt	360
gcgggacgcc gctgcgggcc ccgacgtggg agccgcgggg gacgcgccag ccccgcccc	420
caacaaggac ggagacgccg gcgtgggcag cggccactgg gagctgaggt gccatcgctt	480
gcaggattct ttattcagct ctgacagtgg ctacagcaac taccgtggca tcctgaactg	540
gtgtgtggtg atgctgatct tgagcaatgc ccggttattt ctggagaacc tcatcaagta	600
tggcatcctg gtggaccca tccaggtggg ttctctgttc ctgaaggatc cctatagctg	660
gcccgcacca tgcctgggta ttgcggccaa tgtctttgct gtggctgcat tccaggttga	720
gaagcgctg gcggtgggtg ccctgacgga gcaggcggga ctgctgctgc acgtggccaa	780
cctggccacc attctgtgtt tcccagcggc tgtggtctta ctggttgagt ctatcactcc	840
agtgggctcc ctgctggcgc tgatggcgca caccatcctc ttcctcaagc tcttctccta	900
ccgcgacgtc aactcatggt gccgcagggc cagggccaaag gctgcctctg cagggaagaa	960
ggccagcagt gctgctgccc cgcacaccgt gagctaccgc gacaatctga cctaccgcga	1020

pctau2012001598-seq1

tctctactac	ttcctcttcg	ccccacctt	gtgctacgag	ctcaactttc	cccgtcttcc	1080
ccgcatccgg	aagcgctttc	tgctgcgacg	gatccttgag	atgctgttct	tcaccagct	1140
ccaggtgggg	ctgatccagc	agtggatggg	ccccaccatc	cagaactcca	tgaagccctt	1200
caaggacatg	gactactcac	gcatcatcga	gcgcctcctg	aagctggcgg	tccccaatca	1260
cctcatctgg	ctcatcttct	tctactggct	cttccactcc	tgctgaatg	ccgtggctga	1320
gctcatgcag	tttggagacc	gggagttcta	ccgggactgg	tggaactccg	agtctgtcac	1380
ctacttctgg	cagaactgga	acatccctgt	gcacaagtgg	tgcatcagac	acttctacaa	1440
gccccatgctt	cgacggggca	gcagcaagtg	gatggccagg	acaggggtgt	tcctggcctc	1500
ggccttcttc	cacgagtacc	tggtgagcgt	ccctctgcga	atgttccgcc	tctgggcgtt	1560
cacgggcatg	atggctcaga	tcccactggc	ctggttcgtg	ggccgctttt	tccagggcaa	1620
ctatggcaac	gcagctgtgt	ggctgtcgct	catcatcgga	cagccaatag	ccgtcctcat	1680
gtacgtccac	gactactacg	tgctcaacta	tgaggcccca	gcggcagagg	cctgagctgc	1740
acctgagggc	ctggcttctc	actgccacct	cacaccgct	gccagagccc	acctctcctc	1800
ctaggcctcg	agtgtggggg	atgggcctgg	ctgcacagca	tcctcctctg	gtcccaggga	1860
ggcctctctg	cccctatggg	gctctgtcct	gcaccctca	gggatggcga	cagcaggcca	1920
gacacagtct	gatgccagct	gggagtcttg	ctgaccctgc	cccgggtccg	agggtgtcaa	1980
taaagtgtg	tccagtgacc	tcttcagcct	gccaggggcc	tggggcctgg	tggggggtat	2040
ggccacaccc	acaagggcga	gtgccagagc	tgtgtggaca	gctgtcccag	gacctgccgg	2100
ggagcagcag	ctccactgca	gcagggcggg	catggccggt	agggggagtg	caaggccagg	2160
cagacgcccc	cattccccac	actcccctac	ctagaaaagc	tcagctcagg	cgctccttcc	2220
tggtgctact	cctgctgggt	gtgcgggaat	cggcactctg	gccctcagtg	caggatggtc	2280
aacttgacc	caccgtgctc	agcttcagct	tgggaagagg	tcggccctgg	gtttgcggga	2340
agtgggctcc	agccaggatg	gccagggat	gggcagtcca	cagggccttt	ctggtgtggg	2400
tctggtgagc	cgaggggcta	cggtaacaaa	tggctatgtc	catggctatt	cgccttgggg	2460
cagggctcgg	tttggcagcc	tctttccttc	acttgaacc	tgggcagcca	ggtgggctgg	2520
tggctctcag	ctgccctcca	cgaggttggg	gcaccctga	tccccaggcc	tgcttgccc	2580
tctgtcttct	tggcaccagg	ctgggcagtg	cagacagcgc	agccctgcag	ccttcaggac	2640
acctgtttcc	caggccctgc	aatccaccaa	gctggataga	tggggccccc	cttgatggta	2700
gcgggaggtc	aggggtaacc	ccagggccct	gctgtgcct	taccttgctg	gggtgactgc	2760
ttgattaaag	acccaagac	tggaccaggc	cctggtgtcc	ccagacctgc	agctatggcc	2820
caacaaggca	ggtgttttag	gaggccctgc	acaggggtccc	aggactgcac	caggggcctc	2880
atcagcaacc	tgctgaactg	gtggtgtgag	gggaggggtg	gctcccttgg	tggtatcctc	2940
cctctacccc	agtgggggct	ggtgggccct	gagatggagg	catctccttg	ggccctgctg	3000
aggatggcag	tagtgggtga	gctcctgggt	gggggatggg	tgtggggccg	actcctccat	3060

pctau2012001598-seql

```

cccacatggg ctcctgtgcc tctccctatc agttggccag tggagctcga gaacaaactg      3120
ttgctgccag ataaagcagc aaagcctgct ccactgccca ggcgaggccc ccttccttc      3180
ctccaccagg ctgagtggag aaccctaaag ctacgggaca ggctcccgtt ccccgtcct      3240
acacaggggtg ctggctctgg cagaaagaca gtgtgcccgc cccgcatcac ctgtccgagt      3300
gcgaacacct tgggaagggg gttcctggtg gtctgtgac aaccagagac cccaggggc      3360
aaggtgaggc gcccaccctt ggtccctcaa taaagaactg caagcacatg ctggctctgg      3420
ggagtccggc aggcacggag ctagctgcaa gcatccgtct tccccagga ggccactgca      3480
gcccgccctt tctgtgccct gggcccctcc tcattgcagc accaccacc gcaccacct      3540
gggacagagg aaacacatgg aggccatagc atagagcctg tctgtttata gatctctgcc      3600
tgtcttgtcc gtccatccac tgtgtgtata tatctgtgta tctctgtgga gcggggaacg      3660
ggacagttgt gtaaaaatcc aaaatacaat tctgactatg aacaaaaaaaa aaaaaaaaa      3718

```

<210> 347
 <211> 508
 <212> PRT
 <213> *Aspergillus fumigatus*

<400> 347

Met Lys Met Ser Ala Ser Lys Thr Val Thr Ser Ser Ala Ser Ala Val
 1 5 10 15

Ser Thr Ser Ser Gly Arg Ser Thr Pro Ser Lys Leu Val Asn Gly Ala
 20 25 30

Thr Arg Asn Gly Ser Ala Ala Ala Gly Asn Gly Ser Thr Gly Thr Ala
 35 40 45

Lys Gly Lys Arg Arg Ser Lys Tyr Arg His Val Ala Ala Tyr His Ser
 50 55 60

Glu Leu Arg His Ser Ser Leu Ser Arg Glu Thr Ser Val Val Pro Ser
 65 70 75 80

Phe Leu Gly Phe Arg Asn Leu Met Val Ile Val Leu Val Ala Met Asn
 85 90 95

Leu Arg Leu Ile Ile Glu Asn Phe Met Lys Tyr Gly Val Leu Ile Cys
 100 105 110

Ile Lys Cys His Asp Tyr Arg Lys Gln Asp Val Val Leu Gly Ser Ile
 115 120 125

Leu Phe Ala Leu Val Pro Cys His Leu Phe Leu Ala Tyr Ile Ile Glu
 130 135 140

Leu Val Ala Ala Gln Gln Ser Lys Lys Thr Val Gly Arg Gln Lys Lys
 Page 432

pctau2012001598-seql															
145				150				155				160			
Asp	Leu	Ser	Thr	Glu 165	Glu	Arg	Glu	Arg	Glu 170	Gln	Gln	Ala	Phe	Arg 175	Ser
Thr	Trp	Arg	Tyr 180	Thr	Ala	Phe	Phe	His 185	Thr	Val	Asn	Ala	Thr 190	Leu	Cys
Leu	Ala	Val 195	Thr	Ser	Phe	Val	Val 200	Tyr	Phe	Tyr	Ile	Asn 205	His	Pro	Gly
Ile	Gly 210	Thr	Ile	Cys	Glu	Leu 215	His	Ala	Ile	Ile	Val 220	Trp	Leu	Lys	Asn
Cys 225	Ser	Tyr	Ala	Phe	Thr 230	Asn	Arg	Asp	Leu	Arg 235	Gln	Ala	Met	Val	Asp 240
Pro	Ser	Ala	Glu	Ser 245	Ala	Leu	Pro	Glu	Ile 250	Tyr	Ser	Thr	Cys	Pro 255	Tyr
Pro	Arg	Asn	Ile 260	Thr	Leu	Gly	Asn	Leu 265	Thr	Tyr	Phe	Trp	Leu 270	Ala	Pro
Thr	Leu	Val 275	Tyr	Gln	Pro	Val	Tyr 280	Pro	Arg	Ser	Ser	His 285	Ile	Arg	Trp
Ser	Phe 290	Val	Ala	Lys	Arg	Leu 295	Ala	Glu	Phe	Phe	Gly 300	Leu	Ala	Val	Phe
Ile 305	Trp	Leu	Leu	Ser	Ala 310	Gln	Tyr	Ala	Ala	Pro 315	Val	Leu	Arg	Asn	Ser 320
Ile	Asp	Lys	Ile	Ala 325	Val	Met	Asp	Ile	Ala 330	Ser	Ile	Leu	Glu	Arg 335	Val
Met	Lys	Leu	Ser 340	Thr	Ile	Ser	Leu	Val 345	Ile	Trp	Leu	Ala	Gly 350	Phe	Phe
Ala	Leu	Phe 355	Gln	Ser	Leu	Leu	Asn 360	Ala	Leu	Ala	Glu	Val 365	Met	Arg	Phe
Gly	Asp 370	Arg	Glu	Phe	Tyr	Thr 375	Asp	Trp	Trp	Asn	Ser 380	Pro	Ser	Leu	Gly
Ala 385	Tyr	Trp	Arg	Ser	Trp 390	Asn	Arg	Pro	Val	Tyr 395	Leu	Phe	Met	Lys	Arg 400
His	Val	Phe	Ser	Pro 405	Leu	Val	Gly	Arg	Gly 410	Trp	Ser	Pro	Phe	Ala 415	Ala
Ser	Phe	Met	Val	Phe	Ser	Leu	Ser	Ala	Val	Leu	His	Glu	Met	Leu	Val

pctau2012001598-seql

420

425

430

Gly Ile Pro Thr His Asn Leu Ile Gly Val Ala Phe Ala Gly Met Met
 435 440 445

Phe Gln Leu Pro Leu Ile Ala Val Thr Ala Pro Phe Glu Lys Val Asn
 450 455 460

Asp Ala Leu Gly Lys Ile Val Gly Asn Ser Ile Phe Trp Val Ser Phe
 465 470 475 480

Cys Leu Val Gly Gln Pro Leu Gly Ala Leu Leu Tyr Phe Phe Ala Trp
 485 490 495

Gln Ala Lys Tyr Gly Ser Val Ser Lys Ile His Val
 500 505

<210> 348
 <211> 521
 <212> PRT
 <213> Ricinus communis

<400> 348

Met Thr Ile Leu Glu Thr Pro Glu Thr Leu Gly Val Ile Ser Ser Ser
 1 5 10 15

Ala Thr Ser Asp Leu Asn Leu Ser Leu Arg Arg Arg Arg Thr Ser Asn
 20 25 30

Asp Ser Asp Gly Ala Leu Ala Asp Leu Ala Ser Lys Phe Asp Asp Asp
 35 40 45

Asp Asp Val Arg Ser Glu Asp Ser Ala Glu Asn Ile Ile Glu Asp Pro
 50 55 60

Val Ala Ala Val Thr Glu Leu Ala Thr Ala Lys Ser Asn Gly Lys Asp
 65 70 75 80

Cys Val Ala Asn Ser Asn Lys Asp Lys Ile Asp Ser His Gly Gly Ser
 85 90 95

Ser Asp Phe Lys Leu Ala Tyr Arg Pro Ser Val Pro Ala His Arg Ser
 100 105 110

Leu Lys Glu Ser Pro Leu Ser Ser Asp Leu Ile Phe Lys Gln Ser His
 115 120 125

Ala Gly Leu Phe Asn Leu Cys Ile Val Val Leu Val Ala Val Asn Ser
 130 135 140

Arg Leu Ile Ile Glu Asn Leu Met Lys Tyr Gly Trp Leu Ile Lys Thr
 145 150 155 160

pctau2012001598-seql

Gly Phe Trp Phe Ser Ser Arg Ser Leu Arg Asp Trp Pro Leu Phe Met
 165 170 175
 Cys Cys Leu Ser Leu Pro Val Phe Pro Leu Ala Ala Tyr Leu Val Glu
 180 185 190
 Lys Ala Ala Tyr Arg Lys Tyr Ile Ser Pro Pro Ile Val Ile Phe Leu
 195 200 205
 His Val Ile Ile Thr Ser Ala Ala Val Leu Tyr Pro Ala Ser Val Ile
 210 215 220
 Leu Ser Cys Glu Ser Ala Phe Leu Ser Gly Val Thr Leu Met Glu Leu
 225 230 235 240
 Ala Cys Met Val Trp Leu Lys Leu Val Ser Tyr Ala His Thr Asn Tyr
 245 250 255
 Asp Met Arg Ala Ile Ala Asp Thr Ile His Lys Glu Asp Ala Ser Asn
 260 265 270
 Ser Ser Ser Thr Glu Tyr Cys His Asp Val Ser Phe Lys Thr Leu Ala
 275 280 285
 Tyr Phe Met Val Ala Pro Thr Leu Cys Tyr Gln Pro Ser Tyr Pro Arg
 290 295 300
 Thr Ala Phe Ile Arg Lys Gly Trp Val Phe Arg Gln Phe Val Lys Leu
 305 310 315 320
 Ile Ile Phe Thr Gly Phe Met Gly Phe Ile Ile Glu Gln Tyr Ile Asn
 325 330 335
 Pro Ile Val Gln Asn Ser Gln His Pro Leu Lys Gly Asp Leu Leu Tyr
 340 345 350
 Ala Ile Glu Arg Val Leu Lys Leu Ser Val Pro Asn Leu Tyr Val Trp
 355 360 365
 Leu Cys Leu Phe Tyr Cys Phe Phe His Leu Trp Leu Asn Ile Val Ala
 370 375 380
 Glu Leu Leu Arg Phe Gly Asp Arg Glu Phe Tyr Lys Asp Trp Trp Asn
 385 390 395 400
 Ala Lys Thr Val Glu Glu Tyr Trp Arg Met Trp Asn Met Pro Val His
 405 410 415
 Lys Trp Met Val Arg His Ile Tyr Phe Pro Cys Leu Arg Arg Lys Ile
 420 425 430

pctau2012001598-seql

Pro Arg Gly Val Ala Ile Val Ile Ala Phe Phe Val Ser Ala Val Phe
435 440 445

His Glu Leu Cys Ile Ala Val Pro Cys His Met Phe Lys Leu Trp Ala
450 455 460

Phe Phe Gly Ile Met Phe Gln Ile Pro Leu Val Val Ile Thr Asn Tyr
465 470 475 480

Phe Gln Arg Lys Phe Arg Ser Ser Met Val Gly Asn Met Ile Phe Trp
485 490 495

Phe Phe Phe Cys Ile Leu Gly Gln Pro Met Cys Val Leu Leu Tyr Tyr
500 505 510

His Asp Leu Met Asn Arg Asp Gly Asn
515 520

<210> 349
<211> 526
<212> PRT
<213> Vernicia fordii

<400> 349

Met Thr Ile Pro Glu Thr Pro Asp Asn Ser Thr Asp Ala Thr Thr Ser
1 5 10 15

Gly Gly Ala Glu Ser Ser Ser Asp Leu Asn Leu Ser Leu Arg Arg Arg
20 25 30

Arg Thr Ala Ser Asn Ser Asp Gly Ala Val Ala Glu Leu Ala Ser Lys
35 40 45

Ile Asp Glu Leu Glu Ser Asp Ala Gly Gly Gly Gln Val Ile Lys Asp
50 55 60

Pro Gly Ala Glu Met Asp Ser Gly Thr Leu Lys Ser Asn Gly Lys Asp
65 70 75 80

Cys Gly Thr Val Lys Asp Arg Ile Glu Asn Arg Glu Asn Arg Gly Gly
85 90 95

Ser Asp Val Lys Phe Thr Tyr Arg Pro Ser Val Pro Ala His Arg Ala
100 105 110

Leu Lys Glu Ser Pro Leu Ser Ser Asp Asn Ile Phe Lys Gln Ser His
115 120 125

Ala Gly Leu Phe Asn Leu Cys Ile Val Val Leu Val Ala Val Asn Ser
130 135 140

pctau2012001598-seql

Arg Leu Ile Ile Glu Asn Ile Met Lys Tyr Gly Trp Leu Ile Lys Thr
 145 150 155 160
 Gly Phe Trp Phe Ser Ser Arg Ser Leu Arg Asp Trp Pro Leu Leu Met
 165 170 175
 Cys Cys Leu Thr Leu Pro Ile Phe Ser Leu Ala Ala Tyr Leu Val Glu
 180 185 190
 Lys Leu Ala Cys Arg Lys Tyr Ile Ser Ala Pro Thr Val Val Phe Leu
 195 200 205
 His Ile Leu Phe Ser Ser Thr Ala Val Leu Tyr Pro Val Ser Val Ile
 210 215 220
 Leu Ser Cys Glu Ser Ala Val Leu Ser Gly Val Ala Leu Met Leu Phe
 225 230 235 240
 Ala Cys Ile Val Trp Leu Lys Leu Val Ser Tyr Ala His Thr Asn Phe
 245 250 255
 Asp Met Arg Ala Ile Ala Asn Ser Val Asp Lys Gly Asp Ala Leu Ser
 260 265 270
 Asn Ala Ser Ser Ala Glu Ser Ser His Asp Val Ser Phe Lys Ser Leu
 275 280 285
 Val Tyr Phe Met Val Ala Pro Thr Leu Cys Tyr Gln Pro Ser Tyr Pro
 290 295 300
 Arg Thr Ala Ser Ile Arg Lys Gly Trp Val Val Arg Gln Phe Val Lys
 305 310 315 320
 Leu Ile Ile Phe Thr Gly Phe Met Gly Phe Ile Ile Glu Gln Tyr Ile
 325 330 335
 Asn Pro Ile Val Gln Asn Ser Gln His Pro Leu Lys Gly Asp Leu Leu
 340 345 350
 Tyr Ala Ile Glu Arg Val Leu Lys Leu Ser Val Pro Asn Leu Tyr Val
 355 360 365
 Trp Leu Cys Met Phe Tyr Cys Phe Phe His Leu Trp Leu Asn Ile Leu
 370 375 380
 Ala Glu Leu Leu Arg Phe Gly Asp Arg Glu Phe Tyr Lys Asp Trp Trp
 385 390 395 400
 Asn Ala Arg Thr Val Glu Glu Tyr Trp Arg Met Trp Asn Met Pro Val
 405 410 415

pctau2012001598-seql

His Lys Trp Met Val Arg His Ile Tyr Phe Pro Cys Leu Arg His Lys
420 425 430

Ile Pro Arg Gly Val Ala Leu Leu Ile Thr Phe Phe Val Ser Ala Val
435 440 445

Phe His Glu Leu Cys Ile Ala Val Pro Cys His Ile Phe Lys Leu Trp
450 455 460

Ala Phe Ile Gly Ile Met Phe Gln Ile Pro Leu Val Gly Ile Thr Asn
465 470 475 480

Tyr Leu Gln Asn Lys Phe Arg Ser Ser Met Val Gly Asn Met Ile Phe
485 490 495

Trp Phe Ile Phe Cys Ile Leu Gly Gln Pro Met Cys Leu Leu Leu Tyr
500 505 510

Tyr His Asp Leu Met Asn Arg Lys Gly Thr Thr Glu Ser Arg
515 520 525

<210> 350
<211> 523
<212> PRT
<213> Vernonia galamensis

<400> 350

Met Ala Leu Leu Asp Thr Pro Gln Ile Gly Glu Ile Thr Thr Thr Ala
1 5 10 15

Thr Thr Thr Ile Arg Arg Arg Thr Thr Val Lys Pro Asp Ala Gly Ile
20 25 30

Gly Asp Gly Leu Phe Asp Ser Ser Ser Ser Lys Thr Asn Ser Ser
35 40 45

Phe Glu Asp Gly Asp Ser Leu Asn Gly Asp Phe Asn Asp Lys Phe Lys
50 55 60

Glu Gln Ile Gly Ala Gly Asp Glu Ser Lys Asp Asp Ser Lys Gly Asn
65 70 75 80

Gly Gln Lys Ile Asp His Gly Gly Val Lys Lys Gly Arg Glu Thr Thr
85 90 95

Val Val His Tyr Ala Tyr Arg Pro Ser Ser Pro Ala His Arg Arg Ile
100 105 110

Lys Glu Ser Pro Leu Ser Ser Asp Ala Ile Phe Lys Gln Ser His Ala
115 120 125

pctau2012001598-seql

Gly Leu Phe Asn Leu Cys Ile Val Val Leu Val Ala Val Asn Gly Arg
 130 135 140
 Leu Ile Ile Glu Asn Leu Met Lys Tyr Gly Leu Leu Ile Asn Ser Asn
 145 150 155 160
 Phe Trp Phe Ser Ser Arg Ser Leu Arg Asp Trp Pro Leu Leu Met Cys
 165 170 175
 Cys Leu Thr Pro Ser Asp Phe Pro Leu Ala Ala Tyr Ile Val Glu Lys
 180 185 190
 Leu Ala Trp Lys Lys Arg Ile Ser Asp Pro Val Val Ile Thr Leu His
 195 200 205
 Val Ile Ile Thr Thr Thr Ala Ile Leu Tyr Pro Val Phe Met Ile Leu
 210 215 220
 Arg Phe Asp Ser Val Val Leu Ser Gly Val Ser Leu Met Leu Cys Ala
 225 230 235 240
 Cys Ile Asn Trp Leu Lys Leu Val Ser Phe Val His Thr Asn Tyr Asp
 245 250 255
 Met Arg Ser Leu Leu Asn Ser Thr Asp Lys Gly Glu Val Glu Pro Met
 260 265 270
 Ser Ser Asn Met Asp Tyr Phe Tyr Asp Val Asn Phe Lys Ser Leu Val
 275 280 285
 Tyr Phe Met Val Ala Pro Thr Leu Cys Tyr Gln Ile Ser Tyr Pro Arg
 290 295 300
 Thr Ala Phe Ile Arg Lys Gly Trp Val Leu Arg Gln Leu Ile Lys Leu
 305 310 315 320
 Val Ile Phe Thr Gly Phe Met Gly Phe Ile Ile Glu Gln Tyr Ile Asn
 325 330 335
 Pro Ile Val Lys Asn Ser Arg His Pro Leu Lys Gly Asp Phe Leu Tyr
 340 345 350
 Ala Ile Glu Arg Val Leu Lys Leu Ser Val Pro Asn Leu Tyr Val Trp
 355 360 365
 Leu Cys Met Phe Tyr Cys Phe Phe His Leu Trp Leu Asn Ile Leu Ala
 370 375 380
 Glu Leu Leu Cys Phe Gly Asp Arg Glu Phe Tyr Lys Asp Trp Trp Asn
 385 390 395 400

pctau2012001598-seql

Ala Gln Thr Ile Glu Glu Tyr Trp Arg Leu Trp Asn Met Pro Val His
405 410 415

Lys Trp Ile Val Arg His Leu Tyr Phe Pro Cys Leu Arg Asn Gly Ile
420 425 430

Pro Lys Gly Ala Ala Ile Leu Val Ala Phe Phe Met Ser Ala Val Phe
435 440 445

His Glu Leu Cys Ile Ala Val Pro Cys His Ile Phe Lys Phe Trp Ala
450 455 460

Phe Ile Gly Ile Met Phe Gln Val Pro Leu Val Leu Leu Thr Asn Tyr
465 470 475 480

Leu Gln His Lys Phe Gln Asn Ser Met Val Gly Asn Met Ile Phe Trp
485 490 495

Cys Phe Phe Ser Ile Phe Gly Gln Pro Met Cys Val Leu Leu Tyr Tyr
500 505 510

His Asp Val Met Asn Gln Lys Gly Lys Ser Lys
515 520

<210> 351
<211> 517
<212> PRT
<213> Vernonia galamensis

<400> 351

Met Ala Leu Leu Asp Thr Pro Gln Ile Gly Glu Ile Thr Thr Thr Ala
1 5 10 15

Thr Thr Thr Ile Arg Gln His Pro Leu Gly Lys Pro Asp Ala Gly Ile
20 25 30

Gly Asp Gly Leu Phe Ser Ser Ser Ser Lys Thr Asn Ser Ser Phe
35 40 45

Glu Asp Gly Asp Ser Leu Asn Gly Asp Phe Asn Asp Lys Phe Lys Glu
50 55 60

Gln Ile Gly Ala Gly Asp Glu Ser Lys Lys Gly Asn Gly Lys Ile Asp
65 70 75 80

His Gly Gly Val Lys Lys Gly Arg Glu Thr Thr Val Val His Tyr Ala
85 90 95

Tyr Arg Pro Ser Ser Pro Ala His Arg Arg Ile Lys Glu Ser Pro Leu
100 105 110

Ser Ser Asp Ala Ile Phe Lys Gln Ser His Ala Gly Leu Phe Asn Leu
Page 440

115

120

125

Cys Ile Val Val Leu Val Ala Val Asn Gly Arg Leu Ile Ile Glu Asn
 130 135 140

Leu Met Lys Tyr Gly Leu Leu Ile Asn Ser Lys Phe Trp Phe Ser Ser
 145 150 155 160

Arg Ser Leu Arg Asp Trp Pro Leu Leu Met Cys Trp Leu Thr Pro Ser
 165 170 175

Asp Phe Pro Leu Ala Ala Tyr Ile Val Glu Lys Leu Ala Trp Lys Lys
 180 185 190

Arg Ile Ser Asp Pro Val Val Ile Thr Leu His Val Val Ile Thr Thr
 195 200 205

Thr Ala Ile Leu Tyr Pro Ile Phe Met Ile Leu Arg Phe Asp Ser Val
 210 215 220

Val Leu Leu Gly Val Ser Leu Met Leu Cys Ala Cys Ile Asn Trp Leu
 225 230 235 240

Lys Leu Val Ser Phe Val His Thr Asn Tyr Asp Met Arg Ser Leu Leu
 245 250 255

Asn Ser Thr Gly Lys Gly Glu Val Glu Pro Met Ser Ser Asn Met Asp
 260 265 270

Tyr Phe Tyr Asp Ile Asn Phe Lys Ser Leu Val Tyr Phe Met Val Ala
 275 280 285

Pro Thr Leu Cys Tyr Gln Ile Ser Tyr Pro Arg Thr Ala Phe Ile Arg
 290 295 300

Lys Gly Trp Val Phe Arg Gln Leu Ile Lys Leu Val Ile Phe Thr Gly
 305 310 315 320

Phe Met Gly Phe Ile Ile Glu Gln Tyr Ile Asn Pro Ile Val Lys Asn
 325 330 335

Ser Arg His Pro Leu Asn Gly Asp Phe Leu Tyr Ala Ile Glu Arg Val
 340 345 350

Leu Lys Val Ser Val Pro Asn Leu Tyr Val Trp Leu Cys Met Phe Tyr
 355 360 365

Cys Phe Phe His Leu Trp Leu Asn Ile Leu Ala Glu Leu Leu Trp Phe
 370 375 380

Gly Asp Arg Glu Phe Tyr Lys Asp Trp Trp Asn Thr Gln Thr Ile Glu
 Page 441

pctau2012001598-seq1

385 390 395 400

Glu Tyr Trp Arg Leu Trp Asn Met Pro Val His Lys Trp Ile Val Arg
405 410 415

His Leu Tyr Phe Pro Cys Leu Arg Asn Gly Ile Ser Lys Gly Ala Ala
420 425 430

Ile Leu Val Ala Phe Phe Met Ser Ala Val Phe His Glu Leu Cys Ile
435 440 445

Ala Val Pro Cys His Ile Leu Lys Phe Trp Ala Phe Ile Gly Ile Met
450 455 460

Phe Gln Val Pro Leu Val Leu Leu Thr Asn Tyr Leu Gln His Lys Phe
465 470 475 480

Gln Asn Ser Met Val Gly Asn Met Ile Phe Trp Cys Phe Phe Ser Ile
485 490 495

Phe Gly Gln Pro Met Cys Val Phe Leu Tyr Tyr His Glu Val Asn Gln
500 505 510

Lys Gly Lys Ser Lys
515

<210>	352
<211>	507
<212>	PRT
<213>	Euonymus alatus

<400> 352

Met Ala Ala Asn Leu Asn Glu Ala Ser Asp Leu Asn Phe Ser Leu Arg
1 5 10 15

Arg Arg Thr Gly Gly Ile Ser Ser Thr Thr Val Pro Asp Ser Ser Ser
20 25 30

Glu Thr Ser Ser Ser Glu Ala Asp Tyr Leu Asp Gly Gly Lys Gly Ala
35 40 45

Ala Asp Val Lys Asp Arg Gly Asp Gly Ala Val Glu Phe Gln Asn Ser
50 55 60

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

Asp Ser Arg Phe Thr Tyr Arg Pro Ser Val Pro Ala His Arg Thr Ile
85 90 95

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

pctau2012001598-seql

Gly Leu Phe Asn Leu Cys Ile Val Val Leu Val Ala Val Asn Ser Arg
 115 120 125
 Leu Ile Ile Glu Asn Leu Met Lys Tyr Gly Trp Leu Ile Arg Ser Gly
 130 135 140
 Phe Trp Phe Ser Ser Arg Ser Leu Arg Asp Trp Pro Leu Phe Met Cys
 145 150 155 160
 Cys Leu Thr Leu Pro Val Phe Pro Leu Ala Ala Phe Leu Phe Glu Lys
 165 170 175
 Leu Ala Gln Lys Asn Leu Ile Ser Glu Pro Val Val Val Leu Leu His
 180 185 190
 Ile Val Asn Thr Thr Ala Ala Val Leu Tyr Pro Val Leu Val Ile Leu
 195 200 205
 Arg Cys Asp Ser Ala Phe Met Ser Gly Val Thr Leu Met Leu Phe Ala
 210 215 220
 Cys Ile Val Trp Leu Lys Leu Val Ser Tyr Ala His Thr Asn Tyr Asp
 225 230 235 240
 Met Arg Ala Leu Thr Lys Ser Val Glu Lys Gly Asp Thr Pro Leu Ser
 245 250 255
 Ser Gln Asn Met Asp Tyr Ser Phe Asp Val Asn Ile Lys Ser Leu Ala
 260 265 270
 Tyr Phe Met Val Ala Pro Thr Leu Cys Tyr Gln Ile Ser Tyr Pro Arg
 275 280 285
 Thr Pro Tyr Val Arg Lys Gly Trp Val Val Arg Gln Phe Val Lys Leu
 290 295 300
 Ile Ile Phe Thr Gly Leu Met Gly Phe Ile Ile Glu Gln Tyr Ile Asn
 305 310 315 320
 Pro Ile Val Gln Asn Ser Gln His Pro Leu Lys Gly Asn Phe Leu Tyr
 325 330 335
 Ala Ile Glu Arg Val Leu Lys Leu Ser Val Pro Asn Leu Tyr Val Trp
 340 345 350
 Leu Cys Met Phe Tyr Cys Leu Phe His Leu Trp Leu Asn Ile Leu Ala
 355 360 365
 Glu Leu Leu Cys Phe Gly Asp Arg Glu Phe Tyr Lys Asp Trp Trp Asn
 370 375 380

pctau2012001598-seql

Ala Lys Thr Val Glu Glu Tyr Trp Arg Met Trp Asn Met Pro Val His
385 390 395 400

Lys Trp Met Val Arg His Ile Tyr Phe Pro Cys Leu Arg Asn Gly Ile
405 410 415

Pro Lys Gly Val Ala Phe Val Ile Ser Phe Leu Val Ser Ala Val Phe
420 425 430

His Glu Leu Cys Ile Ala Val Pro Cys His Ile Phe Lys Leu Trp Ala
435 440 445

Phe Phe Gly Ile Met Leu Gln Val Pro Leu Val Leu Ile Thr Ser Tyr
450 455 460

Leu Gln Asn Lys Phe Arg Ser Ser Met Val Gly Asn Met Met Phe Trp
465 470 475 480

Phe Ser Phe Cys Ile Phe Gly Gln Pro Met Cys Leu Leu Leu Tyr Tyr
485 490 495

His Asp Leu Met Asn Arg Asn Gly Lys Met Glu
500 505

<210> 353
<211> 498
<212> PRT
<213> Caenorhabditis elegans

<400> 353

Met Gln Met Arg Gln Gln Thr Gly Arg Arg Arg Arg Gln Pro Ser Glu
1 5 10 15

Thr Ser Asn Gly Ser Leu Ala Ser Ser Arg Arg Ser Ser Phe Ala Gln
20 25 30

Asn Gly Asn Ser Ser Arg Lys Ser Ser Glu Met Arg Gly Pro Cys Glu
35 40 45

Lys Val Val His Thr Ala Gln Asp Ser Leu Phe Ser Thr Ser Ser Gly
50 55 60

Trp Thr Asn Phe Arg Gly Phe Phe Asn Leu Ser Ile Leu Leu Leu Val
65 70 75 80

Leu Ser Asn Gly Arg Val Ala Leu Glu Asn Val Ile Lys Tyr Gly Ile
85 90 95

Leu Ile Thr Pro Leu Gln Trp Ile Ser Thr Phe Val Glu His His Tyr
100 105 110

pctau2012001598-seql

```

Ser Ile Trp Ser Trp Pro Asn Leu Ala Leu Ile Leu Cys Ser Asn Ile
    115                      120                      125

Gln Ile Leu Ser Val Phe Gly Met Glu Lys Ile Leu Glu Arg Gly Trp
    130                      135                      140

Leu Gly Asn Gly Phe Ala Ala Val Phe Tyr Thr Ser Leu Val Ile Ala
    145                      150                      155                      160

His Leu Thr Ile Pro Val Val Val Thr Leu Thr His Lys Trp Lys Asn
    165                      170                      175

Pro Leu Trp Ser Val Val Met Met Gly Val Tyr Val Ile Glu Ala Leu
    180                      185                      190

Lys Phe Ile Ser Tyr Gly His Val Asn Tyr Trp Ala Arg Asp Ala Arg
    195                      200                      205

Arg Lys Ile Thr Glu Leu Lys Thr Gln Val Thr Asp Leu Ala Lys Lys
    210                      215                      220

Thr Cys Asp Pro Lys Gln Phe Trp Asp Leu Lys Asp Glu Leu Ser Met
    225                      230                      235                      240

His Gln Met Ala Ala Gln Tyr Pro Ala Asn Leu Thr Leu Ser Asn Ile
    245                      250                      255

Tyr Tyr Phe Met Ala Ala Pro Thr Leu Cys Tyr Glu Phe Lys Phe Pro
    260                      265                      270

Arg Leu Leu Arg Ile Arg Lys His Phe Leu Ile Lys Arg Thr Val Glu
    275                      280                      285

Leu Ile Phe Leu Ser Phe Leu Ile Ala Ala Leu Val Gln Gln Trp Val
    290                      295                      300

Val Pro Thr Val Arg Asn Ser Met Lys Pro Leu Ser Glu Met Glu Tyr
    305                      310                      315                      320

Ser Arg Cys Leu Glu Arg Leu Leu Lys Leu Ala Ile Pro Asn His Leu
    325                      330                      335

Ile Trp Leu Leu Phe Phe Tyr Thr Phe Phe His Ser Phe Leu Asn Leu
    340                      345                      350

Ile Ala Glu Leu Leu Arg Phe Ala Asp Arg Glu Phe Tyr Arg Asp Phe
    355                      360                      365

Trp Asn Ala Glu Thr Ile Gly Tyr Phe Trp Lys Ser Trp Asn Ile Pro
    370                      375                      380

```

pctau2012001598-seql

Val His Arg Phe Ala Val Arg His Ile Tyr Ser Pro Met Met Arg Asn
385 390 395 400

Asn Phe Ser Lys Met Ser Ala Phe Phe Val Val Phe Phe Val Ser Ala
405 410 415

Phe Phe His Glu Tyr Leu Val Ser Val Pro Leu Lys Ile Phe Arg Leu
420 425 430

Trp Ser Tyr Tyr Gly Met Met Gly Gln Ile Pro Leu Ser Ile Ile Thr
435 440 445

Asp Lys Val Val Arg Gly Gly Arg Thr Gly Asn Ile Ile Val Trp Leu
450 455 460

Ser Leu Ile Val Gly Gln Pro Leu Ala Ile Leu Met Tyr Gly His Asp
465 470 475 480

Trp Tyr Ile Leu Asn Phe Gly Val Ser Ala Val Gln Asn Gln Thr Val
485 490 495

Gly Ile

<210> 354
<211> 498
<212> PRT
<213> Rattus norvegicus
<400> 354

Met Gly Asp Arg Gly Gly Ala Gly Ser Ser Arg Arg Arg Arg Thr Gly
1 5 10 15

Ser Arg Val Ser Val Gln Gly Gly Ser Gly Pro Lys Val Glu Glu Asp
20 25 30

Glu Val Arg Glu Ala Ala Val Ser Pro Asp Leu Gly Ala Gly Gly Asp
35 40 45

Ala Pro Ala Pro Ala Pro Ala Pro Ala His Thr Arg Asp Lys Asp Arg
50 55 60

Gln Thr Ser Val Gly Asp Gly His Trp Glu Leu Arg Cys His Arg Leu
65 70 75 80

Gln Asp Ser Leu Phe Ser Ser Asp Ser Gly Phe Ser Asn Tyr Arg Gly
85 90 95

Ile Leu Asn Trp Cys Val Val Met Leu Ile Leu Ser Asn Ala Arg Leu
100 105 110

pctau2012001598-seql

Ser Leu Glu Asn Leu Ile Lys Tyr Gly Ile Leu Val Asp Pro Ile Gln
115 120 125

Val Val Ser Leu Phe Leu Lys Asp Pro Tyr Ser Trp Pro Ala Pro Cys
130 135 140

Leu Ile Ile Ala Ser Asn Ile Phe Ile Val Ala Thr Phe Gln Ile Glu
145 150 155 160

Lys Arg Leu Ser Val Gly Ala Leu Thr Glu Gln Met Gly Leu Leu Leu
165 170 175

His Val Val Asn Leu Ala Thr Ile Ile Cys Phe Pro Ala Ala Val Ala
180 185 190

Leu Leu Val Glu Ser Ile Thr Pro Val Gly Ser Leu Phe Ala Leu Ala
195 200 205

Ser Tyr Ser Ile Ile Phe Leu Lys Leu Ser Ser Tyr Arg Asp Val Asn
210 215 220

Leu Trp Cys Arg Gln Arg Arg Val Lys Ala Lys Ala Val Ser Ala Gly
225 230 235 240

Lys Lys Val Ser Gly Ala Ala Ala Gln Asn Thr Val Ser Tyr Pro Asp
245 250 255

Asn Leu Thr Tyr Arg Asp Leu Tyr Tyr Phe Ile Phe Ala Pro Thr Leu
260 265 270

Cys Tyr Glu Leu Asn Phe Pro Arg Ser Pro Arg Ile Arg Lys Arg Phe
275 280 285

Leu Leu Arg Arg Val Leu Glu Met Leu Phe Phe Thr Gln Leu Gln Val
290 295 300

Gly Leu Ile Gln Gln Trp Met Val Pro Thr Ile Gln Asn Ser Met Lys
305 310 315 320

Pro Phe Lys Asp Met Asp Tyr Ser Arg Ile Ile Glu Arg Leu Leu Lys
325 330 335

Leu Ala Val Pro Asn His Leu Ile Trp Leu Ile Phe Phe Tyr Trp Leu
340 345 350

Phe His Ser Cys Leu Asn Ala Val Ala Glu Leu Leu Gln Phe Gly Asp
355 360 365

Arg Glu Phe Tyr Arg Asp Trp Trp Asn Ala Glu Ser Val Thr Tyr Phe
370 375 380

pctau2012001598-seql

Trp Gln Asn Trp Asn Ile Pro Val His Lys Trp Cys Ile Arg His Phe
385 390 395 400

Tyr Lys Pro Met Leu Arg Leu Gly Ser Asn Lys Trp Met Ala Arg Thr
405 410 415

Gly Val Phe Trp Ala Ser Ala Phe Phe His Glu Tyr Leu Val Ser Ile
420 425 430

Pro Leu Arg Met Phe Arg Leu Trp Ala Phe Thr Ala Met Met Ala Gln
435 440 445

Val Pro Leu Ala Trp Ile Val Asn Arg Phe Phe Gln Gly Asn Tyr Gly
450 455 460

Asn Ala Ala Val Trp Val Thr Leu Ile Ile Gly Gln Pro Val Ala Val
465 470 475 480

Leu Met Tyr Val His Asp Tyr Tyr Val Leu Asn Tyr Asp Ala Pro Val
485 490 495

Gly Ala

<210> 355
<211> 488
<212> PRT
<213> Homo sapiens

<400> 355

Met Gly Asp Arg Gly Ser Ser Arg Arg Arg Arg Thr Gly Ser Arg Pro
1 5 10 15

Ser Ser His Gly Gly Gly Gly Pro Ala Ala Ala Glu Glu Glu Val Arg
20 25 30

Asp Ala Ala Ala Gly Pro Asp Val Gly Ala Ala Gly Asp Ala Pro Ala
35 40 45

Pro Ala Pro Asn Lys Asp Gly Asp Ala Gly Val Gly Ser Gly His Trp
50 55 60

Glu Leu Arg Cys His Arg Leu Gln Asp Ser Leu Phe Ser Ser Asp Ser
65 70 75 80

Gly Phe Ser Asn Tyr Arg Gly Ile Leu Asn Trp Cys Val Val Met Leu
85 90 95

Ile Leu Ser Asn Ala Arg Leu Phe Leu Glu Asn Leu Ile Lys Tyr Gly
100 105 110

Ile Leu Val Asp Pro Ile Gln Val Val Ser Leu Phe Leu Lys Asp Pro
Page 448

115

120

125

Tyr Ser Trp Pro Ala Pro Cys Leu Val Ile Ala Ala Asn Val Phe Ala
 130 135 140

Val Ala Ala Phe Gln Val Glu Lys Arg Leu Ala Val Gly Ala Leu Thr
 145 150 155 160

Glu Gln Ala Gly Leu Leu Leu His Val Ala Asn Leu Ala Thr Ile Leu
 165 170 175

Cys Phe Pro Ala Ala Val Val Leu Leu Val Glu Ser Ile Thr Pro Val
 180 185 190

Gly Ser Leu Leu Ala Leu Met Ala His Thr Ile Leu Phe Leu Lys Leu
 195 200 205

Phe Ser Tyr Arg Asp Val Asn Ser Trp Cys Arg Arg Ala Arg Ala Lys
 210 215 220

Ala Ala Ser Ala Gly Lys Lys Ala Ser Ser Ala Ala Ala Pro His Thr
 225 230 235 240

Val Ser Tyr Pro Asp Asn Leu Thr Tyr Arg Asp Leu Tyr Tyr Phe Leu
 245 250 255

Phe Ala Pro Thr Leu Cys Tyr Glu Leu Asn Phe Pro Arg Ser Pro Arg
 260 265 270

Ile Arg Lys Arg Phe Leu Leu Arg Arg Ile Leu Glu Met Leu Phe Phe
 275 280 285

Thr Gln Leu Gln Val Gly Leu Ile Gln Gln Trp Met Val Pro Thr Ile
 290 295 300

Gln Asn Ser Met Lys Pro Phe Lys Asp Met Asp Tyr Ser Arg Ile Ile
 305 310 315 320

Glu Arg Leu Leu Lys Leu Ala Val Pro Asn His Leu Ile Trp Leu Ile
 325 330 335

Phe Phe Tyr Trp Leu Phe His Ser Cys Leu Asn Ala Val Ala Glu Leu
 340 345 350

Met Gln Phe Gly Asp Arg Glu Phe Tyr Arg Asp Trp Trp Asn Ser Glu
 355 360 365

Ser Val Thr Tyr Phe Trp Gln Asn Trp Asn Ile Pro Val His Lys Trp
 370 375 380

Cys Ile Arg His Phe Tyr Lys Pro Met Leu Arg Arg Gly Ser Ser Lys
 Page 449

pctau2012001598-seql

385 390 395 400

Trp Met Ala Arg Thr Gly Val Phe Leu Ala Ser Ala Phe Phe His Glu
405 410 415

Tyr Leu Val Ser Val Pro Leu Arg Met Phe Arg Leu Trp Ala Phe Thr
420 425 430

Gly Met Met Ala Gln Ile Pro Leu Ala Trp Phe Val Gly Arg Phe Phe
435 440 445

Gln Gly Asn Tyr Gly Asn Ala Ala Val Trp Leu Ser Leu Ile Ile Gly
450 455 460

Gln Pro Ile Ala Val Leu Met Tyr Val His Asp Tyr Tyr Val Leu Asn
465 470 475 480

Tyr Glu Ala Pro Ala Ala Glu Ala
485

<210> 356
<211> 11
<212> PRT
<213> Artificial Sequence

<220>
<223> conserved sequence

<220>
<221> X
<222> (4)..(4)
<223> Threonine (T) or Serine (S)

<400> 356

Arg Gly Val Xaa Arg His Arg Trp Thr Gly Arg
1 5 10

<210> 357
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> conserved sequence

<220>
<221> X
<222> (1)..(1)
<223> Phenylalanine (F) or Tyrosine (Y)

<400> 357

Xaa Glu Ala His Leu Trp Asp Lys
1 5

<210> 358

<211> 9
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> conserved sequence

<400> 358

Asp Leu Ala Ala Leu Lys Tyr Trp Gly
 1 5

<210> 359
 <211> 8
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> conserved sequence

<220>
 <221> misc_feature
 <222> (2)..(2)
 <223> Xaa can be any naturally occurring amino acid

<220>
 <221> X
 <222> (5)..(5)
 <223> Serine (S) or Alanine (A)

<220>
 <221> X
 <222> (8)..(8)
 <223> any amino acid

<400> 359

Ser Xaa Gly Phe Xaa Arg Gly Xaa
 1 5

<210> 360
 <211> 14
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> conserved sequence

<220>
 <221> X
 <222> (3)..(3)
 <223> Histidine (H) or Gutamine (Q)

<220>
 <221> X
 <222> (6)..(6)
 <223> Arginine (R) or Lysine (K)

<220>
 <221> X
 <222> (12)..(12)
 <223> Arginine (R) or Lysine (K)

<220>

pctau2012001598-seql

<221> misc_feature
 <222> (13)..(13)
 <223> Xaa can be any naturally occurring amino acid

<400> 360

His His Xaa Asn Gly Xaa Trp Glu Ala Arg Ile Gly Xaa Val
 1 5 10

<210> 361
 <211> 9
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> conserved sequence

<220>
 <221> X
 <222> (7)..(7)
 <223> any amino acid

<400> 361

Gln Glu Glu Ala Ala Ala Xaa Tyr Asp
 1 5

<210> 362
 <211> 165
 <212> PRT
 <213> Brassica napus

<400> 362

Met Gly Ile Leu Arg Lys Lys Lys His Glu Arg Lys Pro Ser Phe Lys
 1 5 10 15

Ser Val Leu Thr Ala Ile Leu Ala Thr His Ala Ala Thr Phe Leu Leu
 20 25 30

Leu Ile Ala Gly Val Ser Leu Ala Gly Thr Ala Ala Ala Phe Ile Ala
 35 40 45

Thr Met Pro Leu Phe Val Val Phe Ser Pro Ile Leu Val Pro Ala Gly
 50 55 60

Ile Thr Thr Gly Leu Leu Thr Thr Gly Leu Ala Ala Ala Gly Gly Ala
 65 70 75 80

Gly Ala Thr Ala Val Thr Ile Ile Leu Trp Leu Tyr Lys Arg Ala Thr
 85 90 95

Gly Lys Ala Pro Pro Lys Val Leu Glu Lys Val Leu Lys Lys Ile Ile
 100 105 110

Pro Gly Ala Ala Ala Ala Pro Ala Ala Ala Pro Gly Ala Ala Pro Ala
 115 120 125

pctau2012001598-seql

Ala Ala Pro Ala Ala Ala Pro Ala Val Ala Pro Ala Ala Ala Pro Ala
130 135 140

Ala Ala Pro Ala Pro Lys Pro Ala Ala Pro Pro Ala Pro Lys Pro Ala
145 150 155 160

Ala Ala Pro Ser Ile
165

<210> 363
<211> 193
<212> PRT
<213> Brassica napus

<400> 363

Met Ala Asp Val Arg Thr His Ala His Gln Val Gln Val His Pro Leu
1 5 10 15

Arg Gln Gln Glu Gly Gly Ile Lys Val Val Tyr Pro Gln Ser Gly Pro
20 25 30

Ser Ser Thr Gln Val Leu Ala Val Ile Ala Gly Val Pro Val Gly Gly
35 40 45

Thr Leu Leu Thr Leu Ala Gly Leu Thr Leu Ala Gly Ser Val Ile Gly
50 55 60

Leu Met Leu Ala Phe Pro Leu Phe Leu Ile Phe Ser Pro Val Ile Val
65 70 75 80

Pro Ala Ala Phe Val Ile Gly Leu Ala Met Thr Gly Phe Met Ala Ser
85 90 95

Gly Ala Ile Gly Leu Thr Gly Leu Ser Ser Met Ser Trp Val Leu Asn
100 105 110

His Ile Arg Arg Val Arg Glu Arg Met Pro Asp Glu Leu Glu Glu Ala
115 120 125

Lys Gln Arg Leu Ala Asp Met Ala Glu Tyr Val Gly Gln Arg Thr Lys
130 135 140

Asp Ala Gly Gln Thr Ile Glu Glu Lys Ala His Asp Val Arg Glu Ser
145 150 155 160

Lys Thr Tyr Asp Val Arg Asp Arg Asp Thr Lys Gly His Thr Ala Thr
165 170 175

Gly Gly Asp Arg Asp Thr Lys Thr Thr Arg Glu Val Arg Val Ala Thr
180 185 190

pctau2012001598-seql

Thr

<210> 364
 <211> 188
 <212> PRT
 <213> Brassica napus

<400> 364

Met Ala Asn Val Asp Arg Arg Val Asn Val Asp Arg Thr Asp Lys Gly
 1 5 10 15

Leu Gln Leu Gln Pro Gln Tyr Glu Asp Arg Val Gly Tyr Gly Tyr Gly
 20 25 30

Tyr Gly Gly Asn Thr Asp Tyr Lys Ser Cys Gly Pro Ser Thr Asn Gln
 35 40 45

Ile Val Ala Leu Ile Ala Gly Val Pro Ile Gly Gly Ser Leu Leu Ala
 50 55 60

Leu Ala Gly Leu Thr Leu Ala Gly Ser Val Ile Gly Phe Met Leu Ser
 65 70 75 80

Ile Pro Leu Phe Leu Leu Phe Ser Pro Val Ile Val Pro Ala Ala Leu
 85 90 95

Thr Ile Gly Leu Ala Val Thr Gly Ile Leu Ala Ser Gly Leu Phe Gly
 100 105 110

Leu Thr Gly Leu Ser Ser Val Ser Trp Val Leu Asn Tyr Ile Arg Gly
 115 120 125

Arg Ser Asp Thr Val Pro Glu Gln Leu Asp Tyr Ala Lys Arg Arg Met
 130 135 140

Ala Asp Ala Val Gly Tyr Ala Gly Gln Lys Gly Lys Glu Met Gly Gln
 145 150 155 160

Tyr Val Gln Asp Lys Ala His Glu Ala His Asp Thr Ser Leu Thr Thr
 165 170 175

Glu Thr Asn Gly Lys Thr Arg Arg Ala His Ile Ala
 180 185

<210> 365
 <211> 180
 <212> PRT
 <213> Brassica napus

<400> 365

Met Ala Asp Thr Ala Arg Thr His His Asp Ile Thr Ser Arg Asp Gln
 1 5 10 15

pctau2012001598-seql

Tyr Pro Ile Leu Gly Arg Asp Arg Asp Gln Tyr Pro Tyr Gly Arg Ser
20 25 30

Asp Tyr Gln Thr Ser Gly Gln Asp Tyr Ser Lys Thr Arg Gln Ile Ala
35 40 45

Lys Ala Ala Thr Ala Val Thr Ala Gly Gly Ser Leu Leu Val Leu Ser
50 55 60

Ser Leu Thr Leu Val Gly Thr Val Ile Ala Leu Thr Val Ala Thr Thr
65 70 75 80

Leu Leu Val Ile Phe Ser Pro Ile Leu Val Pro Ala Leu Ile Thr Val
85 90 95

Ala Leu Leu Ile Thr Gly Phe Leu Ser Ser Gly Gly Phe Gly Ile Ala
100 105 110

Asp Ile Thr Val Phe Ser Trp Ile Tyr Lys Tyr Ala Thr Gly Glu His
115 120 125

Pro Gln Gly Ser Asp Lys Leu Asp Ser Ala Arg Met Lys Leu Gly Thr
130 135 140

Lys Ala Gln Asp Ile Lys Asp Arg Ala Gln Tyr Tyr Gly Gln Gln His
145 150 155 160

Thr Gly Gly Glu His Asp Arg Asp Arg Thr Arg Gly Thr His His Thr
165 170 175

Thr Thr Thr Thr
180

<210> 366
<211> 210
<212> PRT
<213> Brassica napus

<400> 366

Met Ala Asp Thr His Arg Val Asp Arg Thr Asp Arg His Leu Gln Phe
1 5 10 15

Gln Ser Pro Tyr Glu Gly Gly Arg Val Ser Ile Gln Tyr Glu Gly Gly
20 25 30

Gly Gly Ala Gly Gly Tyr Gly Gly Arg Gly Gly Gly Tyr Gly Ala Glu
35 40 45

Gly Tyr Lys Ser Met Met Pro Glu Arg Gly Pro Ser Ser Thr Gln Val
50 55 60

pctau2012001598-seql

Leu Ser Phe Leu Val Gly Val Pro Ile Val Gly Ser Leu Leu Ala Ile
65 70 75 80

Ala Gly Leu Leu Leu Ala Gly Ser Val Ile Gly Leu Leu Ile Ser Ile
85 90 95

Pro Leu Phe Leu Leu Phe Ser Pro Val Ile Val Pro Ala Ala Leu Thr
100 105 110

Ile Gly Leu Ala Ala Thr Gly Phe Leu Ala Ser Gly Met Phe Gly Leu
115 120 125

Thr Gly Leu Ser Ser Val Ser Trp Val Leu Asn Tyr Leu Arg Gly Thr
130 135 140

Arg Lys Ser Ser Val Pro Glu Gln Leu Glu Tyr Ala Lys Lys Arg Met
145 150 155 160

Ala Asp Ala Val Gly Tyr Ala Gly Gln Lys Gly Lys Gly Met Gly Gln
165 170 175

His Val Gln Asn Lys Ala Gln Glu Ala Lys Gln Tyr Asp Ile Ser Lys
180 185 190

Thr His Asp Thr Thr Thr Lys Gly His Glu Thr Thr Gln Arg Thr Ala
195 200 205

Ala Ala
210

<210> 367
<211> 149
<212> PRT
<213> Brassica napus

<400> 367

Met Ala Asn Gln Thr Arg Thr His Gln Asp Ile Ile Val Arg Asp Ser
1 5 10 15

Arg Ile Thr Leu Asp Arg Asp His Pro Lys Thr Gly Ala Gln Met Val
20 25 30

Lys Val Ala Thr Gly Val Ala Ala Gly Gly Ser Leu Leu Val Leu Ser
35 40 45

Gly Leu Thr Leu Ala Gly Thr Val Ile Ala Phe Ala Val Ala Thr Pro
50 55 60

Leu Leu Ile Ile Phe Ser Pro Val Leu Val Pro Ala Val Ile Thr Val
65 70 75 80

pctau2012001598-seql

Val Leu Ile Ile Thr Gly Phe Leu Ala Ser Gly Gly Phe Gly Ile Ala
85 90 95

Ala Ile Thr Ala Phe Ser Trp Leu Tyr Arg His Met Thr Gly Ser Gly
100 105 110

Ser Asp Gln Lys Ile Glu Ser Ala Arg Met Lys Val Gly Ser Arg Gly
115 120 125

Tyr Asp Thr Lys Tyr Gly Gln His Asn Ile Gly Val His Gln Gln His
130 135 140

Gln Gln Ala Ala Ser
145

<210> 368
<211> 137
<212> PRT
<213> Arachis hypogaea

<400> 368

Met Ala Glu Ala Leu Tyr Tyr Gly Gly Arg Gln Arg Gln Glu Gln Pro
1 5 10 15

Arg Ser Thr Gln Leu Val Lys Ala Thr Thr Ala Val Val Ala Gly Gly
20 25 30

Ser Leu Leu Ile Leu Ala Gly Leu Val Leu Ala Gly Thr Val Ile Gly
35 40 45

Leu Thr Thr Ile Thr Pro Leu Phe Val Ile Phe Ser Pro Val Leu Val
50 55 60

Pro Ala Val Ile Thr Val Ala Leu Leu Gly Leu Gly Phe Leu Ala Ser
65 70 75 80

Gly Gly Phe Gly Val Ala Ala Ile Thr Val Leu Thr Trp Ile Tyr Arg
85 90 95

Tyr Val Thr Gly Lys His Pro Pro Gly Ala Asn Gln Leu Asp Thr Ala
100 105 110

Arg His Lys Leu Met Gly Lys Ala Arg Glu Ile Lys Asp Phe Gly Gln
115 120 125

Gln Gln Thr Ser Gly Ala Gln Ala Ser
130 135

<210> 369
<211> 150
<212> PRT
<213> Arachis hypogaea

pctau2012001598-seql

<400> 369

Met Thr Asp Arg Thr Gln Pro His Ala Val Gln Val His Thr Thr Ala
1 5 10 15

Gly Arg Phe Gly Asp Thr Ala Ala Gly Thr Asn Arg Tyr Ala Asp Arg
20 25 30

Gly Pro Ser Thr Ser Lys Val Ile Ala Val Ile Thr Gly Leu Pro Ile
35 40 45

Gly Gly Thr Leu Leu Leu Phe Ala Gly Leu Ala Leu Ala Gly Thr Leu
50 55 60

Leu Gly Leu Ala Val Thr Thr Pro Leu Phe Ile Leu Phe Ser Pro Val
65 70 75 80

Ile Val Pro Ala Thr Ile Val Val Gly Leu Ser Val Ala Gly Phe Leu
85 90 95

Thr Ser Gly Ala Cys Gly Leu Thr Gly Leu Ser Ser Phe Ser Trp Val
100 105 110

Met Asn Tyr Ile Arg Gln Thr His Gly Ser Val Pro Glu Gln Leu Glu
115 120 125

Met Ala Lys His Arg Met Ala Asp Val Ala Gly Tyr Val Gly Gln Lys
130 135 140

Thr Lys Asp Val Gly Gln
145 150

<210> 370

<211> 166

<212> PRT

<213> Arachis hypogaea

<400> 370

Met Ser Asp Gln Thr Arg Thr Gly Tyr Gly Gly Gly Gly Ser Tyr Gly
1 5 10 15

Ser Ser Tyr Gly Gly Gly Gly Thr Tyr Gly Ser Ser Tyr Gly Thr Ser
20 25 30

Tyr Asp Pro Ser Thr Asn Gln Pro Ile Arg Gln Ala Ile Lys Phe Met
35 40 45

Thr Ala Ser Thr Ile Gly Val Ser Phe Leu Ile Leu Ser Gly Leu Ile
50 55 60

Leu Thr Gly Thr Val Ile Gly Leu Ile Ile Ala Thr Pro Leu Leu Val
65 70 75 80

pctau2012001598-seql

Ile Phe Ser Pro Ile Leu Val Pro Ala Ala Ile Thr Leu Ala Leu Ala
85 90 95

Ala Gly Gly Phe Leu Phe Ser Gly Gly Cys Gly Val Ala Ala Ile Ala
100 105 110

Ala Leu Ser Trp Leu Tyr Ser Tyr Val Thr Gly Lys His Pro Ala Gly
115 120 125

Ser Asp Arg Leu Asp Tyr Ala Lys Gly Val Ile Ala Asp Lys Ala Arg
130 135 140

Asp Val Lys Asp Arg Ala Lys Asp Tyr Ala Gly Ala Gly Arg Ala Gln
145 150 155 160

Glu Gly Thr Pro Gly Tyr
165

<210> 371
<211> 176
<212> PRT
<213> Arachis hypogaea

<400> 371

Met Ala Thr Ala Thr Asp Arg Ala Pro His Gln Val Gln Val His Thr
1 5 10 15

Pro Thr Thr Gln Arg Val Asp Val Pro Arg Arg Gly Tyr Asp Val Ser
20 25 30

Gly Gly Gly Ile Lys Thr Leu Leu Pro Glu Arg Gly Pro Ser Thr Ser
35 40 45

Gln Ile Ile Ala Val Leu Val Gly Val Pro Thr Gly Gly Thr Leu Leu
50 55 60

Leu Leu Ser Gly Leu Ser Leu Leu Gly Thr Ile Ile Gly Leu Ala Ile
65 70 75 80

Ala Thr Pro Val Phe Ile Phe Phe Ser Pro Val Ile Val Pro Ala Val
85 90 95

Val Thr Ile Gly Leu Ala Val Thr Gly Ile Leu Thr Ala Gly Ala Cys
100 105 110

Gly Leu Thr Gly Leu Met Ser Leu Ser Trp Met Ile Asn Phe Ile Arg
115 120 125

Gln Val His Gly Thr Thr Val Pro Asp Gln Leu Asp Ser Val Lys Arg
130 135 140

pctau2012001598-seql

Arg Met Ala Asp Met Ala Asp Tyr Val Gly Gln Lys Thr Lys Asp Ala
145 150 155 160

Gly Gln Glu Ile Gln Thr Lys Ala Gln Asp Val Lys Arg Ser Ser Ser
165 170 175

<210> 372
<211> 153
<212> PRT
<213> Ricinus communis

<400> 372

Met Ala Asp Arg Pro Gln Pro His Gln Val Gln Val His Arg Tyr Asp
1 5 10 15

Pro Thr Thr Gly Tyr Lys Gly Gln Gln Lys Gly Pro Ser Ala Ser Lys
20 25 30

Val Leu Ala Val Leu Thr Phe Leu Pro Val Gly Gly Gly Leu Leu Ser
35 40 45

Leu Ser Gly Ile Thr Leu Thr Asn Thr Leu Ile Gly Met Ala Ile Ala
50 55 60

Thr Pro Leu Phe Ile Leu Phe Gly Pro Ile Ile Leu Pro Ala Ala Val
65 70 75 80

Val Ile Gly Leu Ala Met Met Ala Phe Met Val Ala Gly Ala Leu Gly
85 90 95

Leu Ser Gly Leu Thr Ser Gln Ser Trp Ala Leu Lys Tyr Phe Arg Glu
100 105 110

Gly Thr Ala Met Pro Glu Ser Leu Asp Gln Ala Lys Lys Arg Met Gln
115 120 125

Asp Met Ala Gly Tyr Val Gly Met Lys Thr Lys Glu Val Gly Gln Asp
130 135 140

Ile Gln Arg Lys Ala Gln Glu Gly Lys
145 150

<210> 373
<211> 138
<212> PRT
<213> Ricinus communis

<400> 373

Met Ala Glu His Gln Gln Ser Pro Val Val Ser His Arg Pro Arg Val
1 5 10 15

Asn Gln Leu Val Lys Ala Gly Thr Ala Ala Thr Ala Gly Ser Ser Leu
20 25 30

pctau2012001598-seql

Leu Phe Leu Ser Gly Leu Thr Leu Thr Gly Thr Val Ile Ala Leu Ala
35 40 45

Leu Ala Thr Pro Leu Met Val Leu Phe Ser Pro Val Leu Leu Pro Ala
50 55 60

Val Ile Ile Ile Ser Leu Ile Gly Ala Gly Phe Leu Thr Ser Gly Gly
65 70 75 80

Phe Gly Phe Gly Ala Ile Leu Val Leu Ser Trp Ile Tyr Arg Tyr Val
85 90 95

Thr Gly Lys Gln Pro Pro Gly Ala Glu Ser Leu Asp Gln Ala Arg Leu
100 105 110

Lys Leu Ala Gly Lys Ala Arg Glu Met Lys Asp Arg Ala Glu Gln Phe
115 120 125

Gly Gln His Val Thr Gly Gln Gln Thr Ser
130 135

<210> 374
<211> 226
<212> PRT
<213> Glycine max
<400> 374

Met Thr Thr Gln Val Pro Pro His Ser Val Gln Val His Thr Thr Thr
1 5 10 15

Thr His Arg Tyr Glu Ala Gly Val Val Pro Pro Gly Ala Arg Phe Glu
20 25 30

Thr Ser Tyr Glu Ala Gly Val Lys Ala Ala Ser Ile Tyr His Ser Glu
35 40 45

Arg Gly Pro Thr Thr Ser Gln Val Leu Ala Val Leu Ala Gly Leu Pro
50 55 60

Val Gly Gly Ile Leu Leu Leu Ala Gly Leu Thr Leu Ala Gly Thr
65 70 75 80

Leu Thr Gly Leu Ala Val Ala Thr Pro Leu Phe Val Leu Phe Ser Pro
85 90 95

Val Leu Val Pro Ala Thr Val Ala Ile Gly Leu Ala Val Ala Gly Phe
100 105 110

Leu Thr Ser Gly Ala Phe Gly Leu Thr Ala Leu Ser Ser Phe Ser Trp
115 120 125

pctau2012001598-seql

Ile Leu Asn Tyr Ile Arg Glu Thr Gln Pro Ala Ser Glu Asn Leu Ala
130 135 140

Ala Ala Ala Lys His His Leu Ala Glu Ala Ala Glu Tyr Val Gly Gln
145 150 155 160

Lys Thr Lys Glu Val Gly Gln Lys Thr Lys Glu Val Gly Gln Asp Ile
165 170 175

Gln Ser Lys Ala Gln Asp Thr Arg Glu Ala Ala Ala Arg Asp Ala Arg
180 185 190

Glu Ala Ala Ala Arg Asp Ala Arg Glu Ala Ala Ala Arg Asp Ala Lys
195 200 205

Val Glu Ala Arg Asp Val Lys Arg Thr Thr Val Thr Ala Thr Thr Ala
210 215 220

Thr Ala
225

<210> 375
<211> 223
<212> PRT
<213> Glycine max

<400> 375

Met Thr Thr Val Pro Pro His Ser Val Gln Val His Thr Thr Thr His
1 5 10 15

Arg Tyr Glu Ala Gly Val Val Pro Pro Ala Arg Phe Glu Ala Pro Arg
20 25 30

Tyr Glu Ala Gly Ile Lys Ala Pro Ser Ser Ile Tyr His Ser Glu Arg
35 40 45

Gly Pro Thr Thr Ser Gln Val Leu Ala Val Val Ala Gly Leu Pro Val
50 55 60

Gly Gly Ile Leu Leu Leu Leu Ala Gly Leu Thr Leu Ala Gly Thr Leu
65 70 75 80

Thr Gly Leu Val Val Ala Thr Pro Leu Phe Ile Ile Phe Ser Pro Val
85 90 95

Leu Ile Pro Ala Thr Val Ala Ile Gly Leu Ala Val Ala Gly Phe Leu
100 105 110

Thr Ser Gly Val Phe Gly Leu Thr Ala Leu Ser Ser Phe Ser Trp Ile
115 120 125

pctau2012001598-seql

Leu Asn Tyr Ile Arg Glu Thr Gln Pro Ala Ser Glu Asn Leu Ala Ala
130 135 140

Ala Ala Lys His His Leu Ala Glu Ala Ala Glu Tyr Val Gly Gln Lys
145 150 155 160

Thr Lys Glu Val Gly Gln Lys Thr Lys Glu Val Gly Gln Asp Ile Gln
165 170 175

Ser Lys Ala Gln Asp Thr Arg Glu Ala Ala Ala Arg Asp Ala Arg Asp
180 185 190

Ala Arg Glu Ala Ala Ala Arg Asp Ala Arg Asp Ala Lys Val Glu Ala
195 200 205

Arg Asp Val Lys Arg Thr Thr Val Thr Ala Thr Thr Ala Thr Ala
210 215 220

<210> 376
<211> 155
<212> PRT
<213> Linum usitatissimum

<400> 376

Met Asp Gln Thr His Gln Thr Tyr Ala Gly Thr Thr Gln Asn Pro Ser
1 5 10 15

Tyr Gly Gly Gly Gly Thr Met Tyr Gln Gln Gln Gln Pro Arg Ser Tyr
20 25 30

Gln Ala Val Lys Ala Ala Thr Ala Ala Thr Ala Gly Gly Ser Leu Ile
35 40 45

Val Leu Ser Gly Leu Ile Leu Thr Ala Thr Val Ile Ser Leu Ile Ile
50 55 60

Ala Thr Pro Leu Leu Val Ile Phe Ser Pro Val Leu Val Pro Ala Leu
65 70 75 80

Ile Thr Val Gly Leu Leu Ile Thr Gly Phe Leu Ala Ser Gly Gly Phe
85 90 95

Gly Val Ala Ala Val Thr Val Leu Ser Trp Ile Tyr Arg Tyr Val Thr
100 105 110

Gly Gly His Pro Ala Gly Gly Asp Ser Leu Asp Gln Ala Arg Ser Lys
115 120 125

Leu Ala Gly Lys Ala Arg Glu Val Lys Asp Arg Ala Ser Glu Phe Ala
130 135 140

Gln Gln His Val Thr Gly Gly Gln Gln Thr Ser
Page 463

pctau2012001598-seql

145

150

155

<210> 377

<211> 180

<212> PRT

<213> Linum usitatissimum

<400> 377

Met Ala Asp Arg Thr Thr Gln Pro His Gln Val Gln Val His Thr Gln
 1 5 10 15

His His Tyr Pro Thr Gly Gly Ala Phe Gly Arg Tyr Glu Gly Val Leu
 20 25 30

Lys Gly Gly Pro Tyr His Gln Gln Gly Thr Gly Ser Gly Pro Ser Ala
 35 40 45

Ser Lys Val Leu Ala Val Met Thr Ala Leu Pro Ile Gly Gly Thr Leu
 50 55 60

Leu Ala Leu Ala Gly Ile Thr Leu Ala Gly Thr Met Ile Gly Leu Ala
 65 70 75 80

Ile Thr Thr Pro Ile Phe Val Ile Cys Ser Pro Val Leu Val Pro Ala
 85 90 95

Ala Leu Leu Ile Gly Phe Ala Val Ser Ala Phe Leu Ala Ser Gly Met
 100 105 110

Ala Gly Leu Thr Gly Leu Thr Ser Leu Ser Trp Phe Ala Arg Tyr Leu
 115 120 125

Gln Gln Ala Gly Gln Gly Val Gly Val Gly Val Pro Asp Ser Phe Asp
 130 135 140

Gln Ala Lys Arg Arg Met Gln Asp Ala Ala Gly Tyr Met Gly Gln Lys
 145 150 155 160

Thr Lys Glu Val Gly Gln Glu Ile Gln Arg Lys Ser Gln Asp Val Lys
 165 170 175

Ala Ser Asp Lys
 180

<210> 378

<211> 181

<212> PRT

<213> Helianthus annuus

<400> 378

Thr Thr Thr Thr Tyr Asp Arg His Phe Thr Thr Thr Gln Pro His Tyr
 1 5 10 15

pctau2012001598-seql

Arg Gln Asp Asp Arg Ser Arg Tyr Asp Gln Gln Thr His Ser Gln Ser
20 25 30

Thr Ser Arg Thr Leu Ala Ile Ile Ala Leu Leu Pro Val Gly Gly Ile
35 40 45

Leu Leu Gly Leu Ala Ala Leu Thr Phe Ile Gly Thr Leu Ile Gly Leu
50 55 60

Ala Leu Ala Thr Pro Leu Phe Val Ile Phe Ser Pro Ile Ile Val Pro
65 70 75 80

Ala Val Leu Thr Ile Gly Leu Ala Val Thr Gly Phe Leu Ala Ser Gly
85 90 95

Thr Phe Gly Leu Thr Gly Leu Ser Ser Leu Ser Tyr Leu Phe Asn Met
100 105 110

Val Arg Gln Thr Ala Gly Ser Val Pro Glu Ser Leu Asp Tyr Val Lys
115 120 125

Gly Thr Leu Gln Asp Ala Gly Glu Tyr Ala Gly Gln Lys Thr Lys Asp
130 135 140

Phe Gly Gln Lys Ile Gln Ser Thr Ala His Glu Met Gly Asp Gln Gly
145 150 155 160

Gln Val Gly Val His Ala Gln Val Gly Gly Lys Glu Gly Arg Lys
165 170 175

Ser Gly Asp Arg Thr
180

<210> 379
<211> 156
<212> PRT
<213> Zea mays

<400> 379

Met Ala Asp His His Arg Gly Ala Thr Gly Gly Gly Gly Gly Tyr Gly
1 5 10 15

Asp Leu Gln Arg Gly Gly Gly Met His Gly Glu Ala Gln Gln Gln Gln
20 25 30

Lys Gln Gly Ala Met Met Thr Ala Leu Lys Ala Ala Thr Ala Ala Thr
35 40 45

Phe Gly Gly Ser Met Leu Val Leu Ser Gly Leu Ile Leu Ala Gly Thr
50 55 60

pctau2012001598-seql

Val Ile Ala Leu Thr Val Ala Thr Pro Val Leu Val Ile Phe Ser Pro
65 70 75 80

Val Leu Val Pro Ala Ala Ile Ala Leu Ala Leu Met Ala Ala Gly Phe
85 90 95

Val Thr Ser Gly Gly Leu Gly Val Ala Ala Leu Ser Val Phe Ser Trp
100 105 110

Met Tyr Lys Tyr Leu Thr Gly Lys His Pro Pro Ala Ala Asp Gln Leu
115 120 125

Asp His Ala Lys Ala Arg Leu Ala Ser Lys Ala Arg Asp Val Lys Asp
130 135 140

Ala Ala Gln His Arg Ile Asp Gln Ala Gln Gly Ser
145 150 155

<210> 380
<211> 244
<212> PRT
<213> Brassica napus

<400> 380

Val Ser Lys Pro Asp Asp Cys Arg Arg Ile Val Asp Glu Thr Ile Ser
1 5 10 15

His Phe Gly Arg Leu Asp His Leu Val Asn Asn Ala Gly Ile Met Gln
20 25 30

Ile Ser Met Phe Glu Asn Ile Glu Glu Ile Thr Arg Thr Arg Ala Val
35 40 45

Met Asp Thr Asn Phe Trp Gly Ser Val Tyr Thr Thr Arg Ala Ala Leu
50 55 60

Pro Tyr Leu Arg Gln Ser Asn Gly Lys Ile Val Ala Met Ser Ser Ser
65 70 75 80

Ala Ala Trp Leu Thr Ala Pro Arg Met Ser Phe Tyr Asn Ala Ser Lys
85 90 95

Ala Ala Leu Leu Asn Phe Phe Glu Thr Leu Arg Ile Glu Leu Gly Ser
100 105 110

Asp Val His Ile Thr Ile Val Thr Pro Gly Tyr Ile Glu Ser Glu Leu
115 120 125

Thr Gln Gly Lys Tyr Phe Ser Gly Glu Gly Glu Leu Val Val Asn Gln
130 135 140

Asp Ile Arg Asp Val Gln Ile Gly Ala Phe Pro Val Thr Ser Val Ser
Page 466

pctau2012001598-seql

Glu Glu Ile Thr Arg Thr Arg Ala Val Met Asp Thr Asn Phe Trp Gly
145 150 155 160

Ala Val Tyr Thr Thr Arg Ala Ala Leu Pro Tyr Leu Arg Gln Ser Asn
165 170 175

Gly Lys Ile Val Ala Met Ser Ser Ser Ala Ala Trp Leu Thr Ala Pro
180 185 190

Arg Met Ser Phe Tyr Asn Ala Ser Lys Ala Ala Leu Leu Asn Phe Phe
195 200 205

Glu Thr Leu Arg Ile Glu Leu Gly Ser Asp Val His Ile Thr Ile Val
210 215 220

Thr Pro Gly Tyr Ile Glu Ser Glu Leu Thr Gln Gly Lys Tyr Val Ser
225 230 235 240

Gly Glu Gly Glu Leu Val Val Asn Gln Asp Ile Arg Asp Val Gln Ile
245 250 255

Gly Ala Phe Pro Val Thr Ser Val Ser Gly Arg Ala Lys Gly Ile Val
260 265 270

Lys Gly Val Cys Arg Lys Glu Arg Tyr Val Thr Glu Pro Ser Trp Phe
275 280 285

Lys Val Thr Tyr Leu Trp Lys Val Phe Cys Pro Glu Leu Ile Glu Trp
290 295 300

Gly Cys Arg Leu Met Phe Leu Ser Gly His Gly Thr Pro Glu Glu Asn
305 310 315 320

Ala Leu Asn Lys Lys Ile Leu Asp Ile Pro Gly Val Arg Ser Ala Leu
325 330 335

Tyr Pro Glu Pro Ile Arg Thr Pro Glu Ile Lys Ser Glu
340 345

<210> 382
<211> 456
<212> PRT
<213> Brassica napus

<400> 382

Met Val Asp Leu Leu Asn Ser Val Met Asn Leu Val Ala Pro Pro Ala
1 5 10 15

Thr Met Val Val Met Ala Phe Ser Trp Pro Leu Leu Cys Phe Ile Thr
20 25 30

pctau2012001598-seql

Phe Ser Glu Arg Leu Tyr Asn Ser Tyr Phe Val Thr Glu Asp Met Glu
35 40 45
Asp Lys Val Val Val Ile Thr Gly Ala Ser Pro Ala Ile Gly Glu Gln
50 55 60
Ile Ala Tyr Glu Tyr Ala Lys Arg Gly Ala Asn Leu Val Leu Val Ala
65 70 75 80
Arg Arg Glu Gln Arg Leu Arg Val Val Ser Asn Asn Ala Arg Gln Ile
85 90 95
Gly Ala Asn His Val Ile Ile Ile Ala Ala Asp Val Val Lys Glu Asp
100 105 110
Asp Cys Arg Arg Phe Ile Thr Gln Ala Val Asn Tyr Tyr Gly Arg Val
115 120 125
Asp His Leu Val Asn Ser Ala Ser Leu Gly His Thr Phe Tyr Phe Asp
130 135 140
Glu Val Ser Asp Thr Thr Val Phe Pro His Leu Leu Asp Ile Asn Phe
145 150 155 160
Trp Gly Asn Val Tyr Pro Thr Tyr Val Ala Leu Pro His Leu Gln Lys
165 170 175
Thr Asn Gly Arg Ile Val Val Asn Ala Ser Val Glu Asn Trp Leu Pro
180 185 190
Leu Pro Arg Met Ser Leu Tyr Ser Ala Ala Lys Ala Ala Leu Val Asn
195 200 205
Phe Tyr Glu Thr Leu Arg Phe Glu Leu Asn Gly Asp Val Gly Ile Thr
210 215 220
Ile Ala Thr His Gly Trp Ile Gly Ser Glu Met Ser Arg Gly Lys Phe
225 230 235 240
Met Leu Glu Glu Gly Ala Glu Met Gln Trp Lys Glu Glu Arg Glu Val
245 250 255
Pro Ala Asn Gly Gly Pro Leu Glu Glu Phe Ala Lys Met Ile Val Ala
260 265 270
Gly Ala Cys Arg Gly Asp Ala Tyr Val Lys Phe Pro Asn Trp Tyr Asp
275 280 285
Val Phe Leu Leu Tyr Arg Val Phe Thr Pro Asn Val Leu Arg Trp Thr
290 295 300

pctau2012001598-seql

Phe Lys Leu Leu Leu Ser Ser Glu Gly Ser Arg Gln Ser Ser Leu Val
305 310 315 320

Gly Val Gly Gln Gly Leu Pro Pro Glu Glu Ser Ser Ser Gln Met Lys
325 330 335

Leu Met Leu Glu Gly Gly Ser Pro Arg Val Thr Ala Ser Pro Pro Arg
340 345 350

Tyr Thr Pro Ser Pro Ser Pro Pro His His Thr Ala Ser Pro Pro Arg
355 360 365

Tyr Thr Pro Ser Pro Ser Pro Pro His His Thr Ser Ser Pro Gln Arg
370 375 380

Tyr Thr Pro Ser Pro Ser Pro Pro His Tyr Thr Ser Ser Arg His Arg
385 390 395 400

Tyr Thr Pro Ser Pro Ser Pro Pro His Tyr Thr Glu Ser Pro Pro Leu
405 410 415

Tyr Thr Glu Ser Pro Pro His Tyr Thr Thr Ser Pro Asn Trp Tyr Thr
420 425 430

Glu Ser Pro Pro Arg Tyr Thr Pro Ser Pro Ser Pro Pro Arg Phe Ser
435 440 445

Arg Phe Asn Ile Gln Glu Leu Pro
450 455

<210> 383
<211> 348
<212> PRT
<213> Sesamum indicum

<400> 383

Met Asp Leu Ile His Thr Phe Leu Asn Leu Ile Ala Pro Pro Phe Thr
1 5 10 15

Phe Phe Phe Leu Leu Phe Phe Leu Pro Pro Phe Gln Ile Phe Lys Phe
20 25 30

Phe Leu Ser Ile Leu Gly Thr Leu Phe Ser Glu Asp Val Ala Gly Lys
35 40 45

Val Val Val Ile Thr Gly Ala Ser Ser Gly Ile Gly Glu Ser Leu Ala
50 55 60

Tyr Glu Tyr Ala Lys Arg Gly Ala Cys Leu Val Leu Ala Ala Arg Arg
65 70 75 80

pctau2012001598-seql

Glu Arg Ser Leu Gln Glu Val Ala Glu Arg Ala Arg Asp Leu Gly Ser
85 90 95

Pro Asp Val Val Val Val Arg Ala Asp Val Ser Lys Ala Glu Asp Cys
100 105 110

Arg Lys Val Val Asp Gln Thr Met Asn Arg Phe Gly Arg Leu Asp His
115 120 125

Leu Val Asn Asn Ala Gly Ile Met Ser Val Ser Met Leu Glu Glu Val
130 135 140

Glu Asp Ile Thr Gly Tyr Arg Glu Thr Met Asp Ile Asn Phe Trp Gly
145 150 155 160

Tyr Val Tyr Met Thr Arg Phe Ala Ala Pro Tyr Leu Arg Asn Ser Arg
165 170 175

Gly Arg Ile Val Val Leu Ser Ser Ser Ser Trp Met Pro Thr Pro
180 185 190

Arg Met Ser Phe Tyr Asn Ala Ser Lys Ala Ala Ile Ser Gln Phe Phe
195 200 205

Glu Thr Leu Arg Val Glu Phe Gly Pro Asp Ile Gly Ile Thr Leu Val
210 215 220

Thr Pro Gly Phe Ile Glu Ser Glu Leu Thr Gln Gly Lys Phe Tyr Asn
225 230 235 240

Ala Gly Glu Arg Val Ile Asp Gln Asp Met Arg Asp Val Gln Val Ser
245 250 255

Thr Thr Pro Ile Leu Arg Val Glu Ser Ala Ala Arg Ser Ile Val Arg
260 265 270

Ser Ala Ile Arg Gly Glu Arg Tyr Val Thr Glu Pro Ala Trp Phe Arg
275 280 285

Val Thr Tyr Trp Trp Lys Leu Phe Cys Pro Glu Val Met Glu Trp Val
290 295 300

Phe Arg Leu Met Tyr Leu Ala Ser Pro Gly Glu Pro Glu Lys Glu Thr
305 310 315 320

Phe Gly Lys Lys Val Leu Asp Tyr Thr Gly Val Lys Ser Leu Leu Tyr
325 330 335

Pro Glu Thr Val Gln Val Pro Glu Pro Lys Asn Asp
340 345

pctau2012001598-seql

<210> 384
 <211> 350
 <212> PRT
 <213> Zea mays

<400> 384

Met Leu Gly Met Ser Arg Thr Gly Leu Ala Gly Ala Ala Leu Arg Val
 1 5 10 15

Ala Leu Thr Ala Leu Leu Pro Leu Val Leu Pro Ala Tyr Tyr Val Tyr
 20 25 30

Lys Leu Thr Thr Tyr Leu Leu Gly Ala Val Phe Pro Glu Asp Val Ala
 35 40 45

Gly Lys Val Val Leu Ile Thr Gly Ala Ser Ser Gly Ile Gly Glu His
 50 55 60

Leu Ala Tyr Glu Tyr Ala Lys Arg Gly Ala Tyr Leu Ala Leu Val Ala
 65 70 75 80

Arg Arg Glu Ala Ser Leu Arg Glu Val Gly Asp Val Ala Leu Gly Leu
 85 90 95

Gly Ser Pro Gly Val Leu Val Leu Pro Ala Asp Val Ser Lys Pro Arg
 100 105 110

Asp Cys Glu Gly Phe Ile Asp Asp Thr Ile Ser Tyr Phe Gly Arg Leu
 115 120 125

Asp His Leu Val Asn Asn Ala Ser Ile Trp Gln Val Cys Lys Phe Glu
 130 135 140

Glu Ile Gln Asp Val Arg His Leu Arg Ala Leu Met Asp Ile Asn Phe
 145 150 155 160

Trp Gly His Val Tyr Pro Thr Arg Leu Ala Ile Pro His Leu Arg Arg
 165 170 175

Ser Arg Gly Arg Ile Val Gly Val Thr Ser Asn Ser Ser Tyr Ile Phe
 180 185 190

Ile Gly Arg Asn Thr Phe Tyr Asn Ala Ser Lys Ala Ala Ala Leu Ser
 195 200 205

Phe Tyr Asp Thr Leu Arg Met Glu Leu Gly Ser Asp Ile Arg Ile Thr
 210 215 220

Glu Val Val Pro Gly Val Val Glu Ser Glu Ile Thr Lys Gly Lys Met
 225 230 235 240

Leu Thr Lys Gly Gly Glu Met Lys Val Asp Gln Asp Glu Arg Asp Ala
 Page 472

pctau2012001598-seql

245

250

255

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

Val Val Arg Asp Val Cys Arg Gly Ala Arg Tyr Val Phe Glu Pro Arg
 275 280 285

Trp Tyr Met Gly Val Tyr Leu Leu Arg Ala Cys Leu Pro Glu Val Leu
 290 295 300

Ala Trp Asn Ser Arg Leu Leu Thr Val Asp Thr Val Gly Ala Ser Thr
 305 310 315 320

Thr Asp Thr Leu Gly Lys Trp Leu Val Glu Leu Pro Gly Val Arg Arg
 325 330 335

Val Val Gln Pro Pro Ser Leu Arg Ser Pro Glu Ile Lys Asp
 340 345 350

<210> 385

<211> 245

<212> PRT

<213> Brassica napus

<400> 385

Met Gly Thr Ala Thr Glu Ile Met Glu Arg Asp Ala Met Ala Thr Val
 1 5 10 15

Ala Pro Tyr Ala Pro Val Thr Phe His Arg Arg Ala Arg Val Asp Leu
 20 25 30

Asp Asp Arg Leu Pro Lys Pro Tyr Met Pro Arg Ala Leu Gln Ala Pro
 35 40 45

Asp Arg Glu His Pro Tyr Gly Thr Pro Gly His Lys Asn Tyr Gly Leu
 50 55 60

Ser Val Leu Gln Gln His Val Ala Phe Phe Asp Ile Asp Asp Asn Gly
 65 70 75 80

Ile Ile Tyr Pro Trp Glu Thr Tyr Ser Gly Leu Arg Met Ile Gly Phe
 85 90 95

Asn Ile Ile Gly Ser Leu Ile Ile Ala Ala Val Ile Asn Leu Ala Leu
 100 105 110

Ser Tyr Ala Thr Leu Pro Gly Trp Leu Pro Ser Pro Phe Phe Pro Ile
 115 120 125

Tyr Ile His Asn Ile His Lys Ser Lys His Gly Ser Asp Ser Arg Thr
 130 135 140

pctau2012001598-seql

Tyr Asp Asn Glu Gly Arg Phe Met Pro Val Asn Leu Glu Leu Ile Phe
145 150 155 160

Ser Lys Tyr Ala Lys Thr Leu Pro Asp Lys Leu Ser Leu Gly Glu Leu
165 170 175

Trp Asp Met Thr Glu Gly Gln Arg Asp Ala Trp Asp Ile Phe Gly Trp
180 185 190

Phe Ala Ser Lys Ile Glu Trp Gly Leu Leu Tyr Leu Leu Ala Arg Asp
195 200 205

Glu Glu Gly Phe Leu Ser Lys Glu Ala Ile Arg Arg Cys Phe Asp Gly
210 215 220

Ser Leu Phe Glu Tyr Cys Ala Lys Ile Tyr Val Gly Ile Asn Glu Asp
225 230 235 240

Lys Thr Ala Tyr Tyr
245

<210> 386
<211> 244
<212> PRT
<213> Brassica napus

<400> 386

Met Val Arg Glu Ser Met Gly Glu Glu Ser Glu Ala Phe Ala Thr Thr
1 5 10 15

Ala Pro Leu Ala Pro Val Thr Gly Glu Arg Lys Val Arg Asn Asp Leu
20 25 30

Glu Glu Thr Leu Pro Lys Pro Tyr Leu Ala Arg Ala Leu Val Ala Pro
35 40 45

Asp Thr Glu His Pro Asn Gly Ser Glu Gly His Asp Ser Lys Gly Met
50 55 60

Ser Val Thr Gln Gln His Val Ala Phe Phe Asp Gln Asn Gly Asp Gly
65 70 75 80

Ile Val Tyr Pro Trp Glu Thr Tyr Ala Gly Phe Arg Asp Leu Gly Phe
85 90 95

Asn Pro Ile Ser Ser Val Phe Trp Ala Ile Phe Ile Asn Phe Ala Phe
100 105 110

Ser Tyr Val Thr Leu Pro Ser Trp Leu Pro Ser Pro Leu Leu Pro Val
115 120 125

pctau2012001598-seql

Tyr Ile Asp Asn Ile His Lys Ala Lys His Gly Ser Asp Ser Ser Thr
130 135 140

Tyr Asp Thr Glu Gly Arg Tyr Val Pro Val Asn Leu Glu Asn Ile Phe
145 150 155 160

Ser Lys Tyr Ala Leu Thr Ala Pro Asn Lys Ile Thr Leu Lys Glu Leu
165 170 175

Trp Asn Leu Thr Glu Gly Asn Arg Met Ala Ile Asp Pro Phe Gly Trp
180 185 190

Leu Ala Asn Lys Val Glu Trp Leu Leu Val Tyr Leu Leu Ala Lys Asp
195 200 205

Glu Glu Gly Phe Val Ser Lys Glu Ala Val Arg Gly Val Phe Asp Ala
210 215 220

Ser Phe Phe Glu Tyr Cys Ala Lys Lys Asn Lys Glu Lys Ala Asp Ser
225 230 235 240

Arg Lys Gln Asp

<210> 387

<211> 245

<212> PRT

<213> Sesamum indicum

<400> 387

Met Ala Thr His Val Leu Ala Ala Ala Ala Glu Arg Asn Ala Ala Leu
1 5 10 15

Ala Pro Asp Ala Pro Leu Ala Pro Val Thr Met Glu Arg Pro Val Arg
20 25 30

Thr Asp Leu Glu Thr Ser Ile Pro Lys Pro Tyr Met Ala Arg Gly Leu
35 40 45

Val Ala Pro Asp Met Asp His Pro Asn Gly Thr Pro Gly His Val His
50 55 60

Asp Asn Leu Ser Val Leu Gln Gln His Cys Ala Phe Phe Asp Gln Asp
65 70 75 80

Asp Asn Gly Ile Ile Tyr Pro Trp Glu Thr Tyr Ser Gly Leu Arg Gln
85 90 95

Ile Gly Phe Asn Val Ile Ala Ser Leu Ile Met Ala Ile Val Ile Asn
100 105 110

pctau2012001598-seql

Val Ala Leu Ser Tyr Pro Thr Leu Pro Gly Trp Ile Pro Ser Pro Phe
115 120 125

Phe Pro Ile Tyr Leu Tyr Asn Ile His Lys Ala Lys His Gly Ser Asp
130 135 140

Ser Gly Thr Tyr Asp Thr Glu Gly Arg Tyr Leu Pro Met Asn Phe Glu
145 150 155 160

Asn Leu Phe Ser Lys His Ala Arg Thr Met Pro Asp Arg Leu Thr Leu
165 170 175

Gly Glu Leu Trp Ser Met Thr Glu Ala Asn Arg Glu Ala Phe Asp Ile
180 185 190

Phe Gly Trp Ile Ala Ser Lys Met Glu Trp Thr Leu Leu Tyr Ile Leu
195 200 205

Ala Arg Asp Gln Asp Gly Phe Leu Ser Lys Glu Ala Ile Arg Arg Cys
210 215 220

Tyr Asp Gly Ser Leu Phe Glu Tyr Cys Ala Lys Met Gln Arg Gly Ala
225 230 235 240

Glu Asp Lys Met Lys
245

<210> 388
<211> 243
<212> PRT
<213> Zea mays

<400> 388

Met Ser Ser Tyr Ser Pro Pro Pro Pro Pro Pro Arg Asp Gln Ser Met
1 5 10 15

Asp Thr Glu Ala Pro Asn Ala Pro Ile Thr Arg Glu Arg Arg Leu Asn
20 25 30

Pro Asp Leu Gln Glu Gln Leu Pro Lys Pro Tyr Leu Ala Arg Ala Leu
35 40 45

Glu Ala Val Asp Pro Ser His Pro Gln Gly Thr Lys Gly Arg Asp Pro
50 55 60

Arg Gly Met Ser Val Leu Gln Gln His Ala Ala Phe Phe Asp Arg Asn
65 70 75 80

Gly Asp Gly Val Ile Tyr Pro Trp Glu Thr Phe Gln Gly Leu Arg Ala
85 90 95

Ile Gly Cys Gly Leu Thr Val Ser Phe Ala Phe Ser Ile Leu Ile Asn
Page 476

pctau2012001598-seql

100

105

110

Leu Phe Leu Ser Tyr Pro Thr Gln Pro Gly Trp Leu Pro Ser Pro Leu
 115 120 125

Leu Ser Ile Arg Ile Asp Asn Ile His Lys Gly Lys His Gly Ser Asp
 130 135 140

Ser Glu Thr Tyr Asp Thr Glu Gly Arg Phe Asp Pro Ser Lys Phe Asp
 145 150 155 160

Ala Ile Phe Ser Lys Tyr Gly Arg Thr His Pro Asn Ala Ile Thr Arg
 165 170 175

Asp Glu Leu Ser Ser Met Leu Gln Gly Asn Arg Asn Thr Tyr Asp Phe
 180 185 190

Leu Gly Trp Leu Ala Ala Ala Gly Glu Trp Leu Leu Leu Tyr Ser Leu
 195 200 205

Ala Lys Asp Lys Asp Gly Leu Leu Gln Arg Glu Thr Val Arg Gly Leu
 210 215 220

Phe Asp Gly Ser Leu Phe Glu Arg Leu Glu Asp Asp Asn Asn Lys Lys
 225 230 235 240

Lys Ser Ser

<210> 389

<211> 737

<212> DNA

<213> Brassica napus

<400> 389

gaattccatt caaactagta aacaatgggg atactcagga agaaaaaaca cgagcgaaag 60

ccgtcgttta agagtgtttt aaccgcaata ttagctacac atgccgcaac attcctcttg 120

ttgatcgccg gtgtatccct cgccggcaca gccgccgcat ttatcgctac catgccacta 180

ttcgtagtat tcagtccgat tctcgtacca gctggtatta ccaactggttt actgactacg 240

ggtttagcag ccgccggtgg cgccggcgcg actgctgtca ccatcatcct gtggctctac 300

aagcgagcaa cgggcaaggc gccgccaaaa gtcctagaaa aagtcttgaa aaagataata 360

ccaggtgctg cagctgcacc agcagccgct ccaggagccg ctccagcagc ggcgccagca 420

gccgcaccag ctgtggcgcc agcagccgca ccagctgctg cgccagcacc taagccagca 480

gccccaccag cacctaagcc agcagccgca ccgagtatat gaaaagaagt ggtgggcatg 540

agtaaagggt gatatggaaa actggataca tagaaaaaag agtaatccaa cttttaaaaa 600

ataaataaca acttcacgtg gggatagaaa aattttcaaa tattatttta ctaatggatg 660

tcgcggtaca aaataataac aaatgtaagc ctttttattg tatagtattt taagaacgaa 720

pctau2012001598-seql

gctatgtagc gttgaca

737

<210> 390
 <211> 837
 <212> DNA
 <213> Brassica napus

<400> 390
 tcttgctttt agtttcttgc gtaacaaaga tttagcttgg aggattaatg gcggacgttc 60
 gcacacatgc tcaccagggt caagtgcacc ctctacgccg gcaggaagga ggcaccaaag 120
 tggtttatcc ccagagcggg ccttcttcca ctcaggttct agcagtgatc gccggtgtac 180
 cggttgagg gacgtgcta actctggccg gtttaacttt agccggttcg gttataggac 240
 taatgctggc attcccgtg ttcctcatct tcagtccggt tatcgtacca gcggcctttg 300
 tgatcggttt agctatgaca ggattcatgg cgtcaggggc aatagggttc acgggactat 360
 cgtcgatgtc gtgggtactg aaccacatcc gaagagtaag ggaacgtatg ccggatgagc 420
 tggaggaagc aaagcagcgt ttggctgaca tggccgagta tgtggggcag aggacaaaag 480
 atgctggaca gaccatagaa gaaaaagctc acgatgtacg agagagcaag acttatgatg 540
 tccgagacag agacacaaag ggtcactact ccacaggagg agacagggac accaagacaa 600
 ctgcgcaggt ccgagtggcg acaacatgag gcagcgtgtg ttttgtggtg aataaaagag 660
 agtcgcctgt tgtgggttctg gtttcttggg agatagggtg tgtgtgtttt aagcttaaaa 720
 tgggtgttctg ctctgtaagc gttctttgag tgttgagtca attttctttt tctttttgta 780
 ataataataa gtgaaaatgc atcttgttta gtagaaaaag aaaaaaaca aaaaaaa 837

<210> 391
 <211> 809
 <212> DNA
 <213> Brassica napus

<400> 391
 gtttagactt tgcttttttag tagttctaga tcatatcaat cgctatggca aatgtggaca 60
 ggcgtgtgaa tgtcgaccgt actgacaaag gacttcaact tcagccgcag tacgaagatc 120
 gtgtcggtta tggatatggt tatggtggta atactgatta caagagttgc ggcccttcca 180
 ctaaccaa at cgtggcactt atagcaggag ttccaattgg tggctcattg ctagccttag 240
 ctggactcac tctagccggt tcggtgatag gttttatgct ctccattccg ctcttcttac 300
 tcttcagtcc ggtgattggt ccggcagctc tcaccatcgg gctagctgta accggaattt 360
 tggcgtcagg gttatttggg ctgacgggtc tcagctcggg ctcgtgggtc ctaaactaca 420
 tccgtgggag gagtgatacg gtgccggagc aattggacta tgctaagagg cgtatggctg 480
 atgcagtagg ctatgctggt cagaagggaa aagagatggg gcagtatgtg caggataagg 540
 ctcatgaagc acatgatact agtctgacca ctgagactaa tggtaagaca aggagggccc 600
 atatagcgta agtatgaaag caccctctgt tgtgggcttt tgttggaag acaattatgt 660
 aattttactt ttgtgtccca gtttttttct cttagaatat attgtcaatg tgtaagttcg 720

pctau2012001598-seql

ttttgtcgtg tccaaataaa atattttgtg ggtgaaataa gaaaagaact gttgtctgca 780
 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 809

<210> 392
 <211> 891
 <212> DNA
 <213> Brassica napus

<400> 392
 gccggggact tcaaaacata aactcatctc tctcttcaac aagatcataa tggcgggatac 60
 tgctagaact catcacgaca taaccagccg tgaccagtac ccgatattgg gccgagaccg 120
 agaccagtat ccttacggac gatcagacta ccagacgtcg ggccaagact actccaagac 180
 taggcagatt gctaaagctg ctaccgcagt caccgccgga gggtccttctc ttgtcctctc 240
 cagtctcacc cttgtcggaa cagtcattgc tttgactggt gccactactc tgcttggttat 300
 ctttagccca atcctcgtgc cggctctcat caccgtagca cttctcatca ctggctttct 360
 ctctcttggt gggtttgga ttgcagatat aaccgtcttc tcctggatct ataagtacgc 420
 aacgggagag caccacagg ggtcagataa gttggacagt gcaaggatga agctgggaac 480
 caaagctcag gatattaaag acagagctca atactacgga cagcaacata cagggtggtga 540
 gcatgaccgt gaccgtaccc gtggaacca tcacaccacc actactacat aaatcacgca 600
 acagcccca gtaatgtcga tgtcagggga gtctagtcca tgaggaataa ggtgtttaga 660
 atttgatcag ggtgggtaat aaaaggggaa gcatatatca tcagtttttt ttcttttttt 720
 gttataaata atgtgcaagt gtgtttctct tgtcaaatgg taccatgttt tgttggttta 780
 ttctctttgt ggtaaattgt aaaccttttg ttgtactttc tttgccttgt ttgtaagttg 840
 taacgtctat atggagaagt cttgttcgtc caataaacia tagaaciaaa a 891

<210> 393
 <211> 890
 <212> DNA
 <213> Brassica napus

<400> 393
 acgatcaaa gaaaaactaga cactactcaa aattatctta atggcggata cacatcgctg 60
 cgaccgtact gataggcact tgcaatttca gtcgccgtat gaaggcggac gagttagtat 120
 tcagtatgaa ggcggtggtg gcgcaggcgg ctatggtggc aggggtggag gatacgggtg 180
 tgaaggttac aaaagcatga tgcctgaacg tggcccatct agcacccaag tgttatcctt 240
 tttggtgggg gtccctatcg ttggttcgct gcttgccata gccggtttac ttttgccggg 300
 ttcggtgata ggtttactga tctctatacc acttttcctc ctcttcagcc cggtcatagt 360
 cccagcggct ctaaccatcg ggcttcgacg gacaggattc ttggcctctg gaatgttcgg 420
 tctaacgggg cttagctcag tctcttggtt cctgaactac cttcgtggga ccaggaagag 480
 tagtgtgcct gagcagttgg agtatgctaa gaagagaatg gcggatgctg ttggctacgc 540
 tggacagaag ggaaaaggaa tggggcagca cgtgcagaat aaggcacagg aagctaaaca 600

pctau2012001598-seql

gtatgatatc tctaagaccc atgatactac tactaagggc catgagacta ctcagaggac	660
ggccgcagca tgatgaagtc ctgtttttgag ggcaagttat gtaactttca ccgtgatgtc	720
cttttacttc tcctggacat ttcactacta caatgttaag ctcttttttg tttctaagtc	780
aaaatcatga aatgtagcgt ctgtgtatct tcatgcaaga tcgcaagtgg cgtacttgaa	840
tattcaataa tttcttcgta caaaaaaaaa aaaaaaaaaa aaaaaaaaaa	890

<210> 394
 <211> 631
 <212> DNA
 <213> Brassica napus

<400> 394	
ggtgtcggag agatcataac aagaatcaat ggcgaatcaa acaagaaccc atcaagacat	60
aatagtccga gacagtagaa ttaccctaga cagagaccat ccgaagacag gggcgcagat	120
ggtgaaggta gcgactgggt tcgcagccgg tggatccctc ctagtccctc caggcttaac	180
actcgccggg acagtgatag cattcgctgt agccactcct ctgcttatta tatttagccc	240
tgtcttggtc ccagcgggtg tcaccgtggg tctcatcatc actgggttcc ttgcctccgg	300
tggctttggg atagccgcca taaccgcctt ctcttggtc tacaggcaca tgacgggggc	360
aggatcggat caaaagatag agagtgcccg aatgaagggt ggaagtagag ggtatgatac	420
caagtatggg cagcacaaca ttggagtcca tcagcaacac caacaagcag cttcttaaaa	480
ttaatgtcag atagaactat ttccacacaa attgtcatcc atcgtttgat ctcttttctt	540
cgttttctca atgcttggtg tgtaacgtaa cgtgtaataa taacatgtgt gtatgctgtt	600
aaaaaaaaa aaaaaaaaaa aaaaaaaaaa a	631

<210> 395
 <211> 655
 <212> DNA
 <213> Arachis hypogaea

<400> 395	
ttatggccgg ggattcgcat tagcattagc cttttcattt cacttcataa ttaattaata	60
acatggctga agcactctac tacggcggcc gccaacgcc aagagcaacca aggtccaccc	120
agcttgctca ggccaccacc gctgttgctg ccggaggctc cctcttgatc ctgcgccggc	180
ttgtgctggc cggcaccgtc attggcctca caacgatcac accgctcttc gtgatcttca	240
gcccgggtgct tgtgccagct gtcactactg tggcactctt aggcttgggg ttcttggcct	300
ctggaggctt cggcgtggcg gcaataacag tgctgacgtg gatctatagg tacgtaacag	360
gtaagcatcc acctggcgcc aaccaattgg acacagcccg ccacaagctg atgggcaagg	420
cgctgtgagat taaggacttt ggtcaacaac aaaccagtgg ggcccaggct tcttgagcat	480
accatcttcg tttgcatctt tgtttgcacg cacgtccacg ccatcattta tctttttcga	540
attgttatgg tttattttat tttatttaat ttttttatga gtctgggggt tccttgaaat	600
taaccgttgg tttaaaatat tttccctggg ttatccaatc ccattcaaat tttta	655

pctau2012001598-seql

<210> 396
 <211> 578
 <212> DNA
 <213> Arachis hypogaea

<400> 396
 ggtcacatgc atacacaaac caaattaaat atcattcctt ctctttacct cttctccctg 60
 aaacctgctt cattcagagt aaaagaaaca taaggaagaa aaggagctt cctgcaacca 120
 tcagccatga ctgaccgtac ccaaccacac gctgtccaag tccacaccac agctggccgt 180
 ttcggcgaca ccgctgccgg aactaaccgc tatgccgaca gaggcccgtc aacatctaag 240
 gttatcgccg tcatcactgg actccctatc ggcggcacgt tgctattgtt cgcggggctt 300
 gcccttgccg gaacctgct tgggctggcg gtgaccaccc cgcttttcat cctcttcagc 360
 cctgtcatag ttccagccac cattgtcgtt gggctctcgg tggcgggggtt cttgacgtca 420
 ggtgcatgtg ggctgacggg gctgtcttcg ttctcgtggg tcatgaatta catccggcag 480
 acccacggat cggtgccgga gcagctggaa atggcaaagc accgcatggc tgacgtggcc 540
 ggttacgttg gacagaagac gaaggatgta ggacagaa 578

<210> 397
 <211> 647
 <212> DNA
 <213> Arachis hypogaea

<400> 397
 gaaaccccat cacttcttgt ctaaaaattc tcaaaagtca ccagccacca aaaaccatt 60
 taccattatg tctgatcaaa caaggacagg ctatggagga ggagggctct atggatcatc 120
 ctatggtgga ggaggcacct atggttcac tttatggaacc tcctatgacc ccagtactaa 180
 ccaacctata cgccaagcca tcaagttcat gacagcatca accattggtg tctcattctt 240
 gatcctgtct gggttgatcc tcaactggaac tgtcataggt ttgatcattg caacaccact 300
 tcttgttatc ttcagtccta tccttgctcc tgctgccata acccttgac tggtctgtgg 360
 tggatttttg ttctctggtg gctgtggtgt tgctgccatt gctgcattgt catggttgta 420
 cagctatgtc actgggaaac accctgctgg ctctgatagg cttgattatg ctaaaggggt 480
 gattgctgat aaggctaggg atgttaagga cagggccaa gattatgctg gtgctggtag 540
 ggctcaggag ggcacccag ggtattgatc ctcatgtga tgaaaaaaaa tggaagcttt 600
 tgtgtgtaat gtgtgggtga agtgaaggct tgaaagggtg caccctc 647

<210> 398
 <211> 817
 <212> DNA
 <213> Arachis hypogaea

<400> 398
 tttttcggga tttggggggg tttttttttt ttttctggga aagcgccccc ttgttttggc 60
 ccccggaat tggccatta gggccggggg aacaaaagga atcccaaaaa gggccttttc 120

pctau2012001598-seq1

ttttctctta acaatcaaac aacagcaatg gctactgcta ctgatcgtgc acctcaccag	180
gttcaagttc acacccccac cacacaacgc gtcgacgttc cacgccgcgg ctacgatgtt	240
agtgggtgtg gtattaagac tcttctcccc gagagaggtc cgtccacctc tcaaatacat	300
gccgtcctcg tcggcgtccc cactgggggc actctgtttg tcctctccgg cctttcactt	360
ctcggaaacca taatcgggct ggcaattgcc accccggttt ttatcttctt cagcccgttt	420
atagttcccc cggtcgttac cattggactt gcagtcactg gtattctcac ggcgggagca	480
tgtggactaa ccgggctgat gtctttgtca tggatgatta acttcatccg acaggtacat	540
gggacgacgg tgccggatca gctggactca gtgaagcggc gcatggcgga catggcggat	600
tacgtggggc agaagacaaa ggatgctggc caagagatac agactaaggc ccaggatgtt	660
aagaggtcat catcataaag taaaaaaaaa aaaaaaaaaa aaaaaaaagg gggggcgccc	720
cccccccccc cctatctaaa aaacaaaaaa ttttttgggg gggggcccca aatgggttcc	780
tgggggccta cccccaagg agagggggga aaaaaa	817

<210> 399
 <211> 625
 <212> DNA
 <213> Helianthus annuus

<400> 399	
actaccacca cttacgaccg tcatttcacc accaccaac ccactaccg ccaagacgac	60
aggtcccgct acgaccagca aaccattcc cagtcacca gcaggacact cgccatcatc	120
gccctacttc ctgtcggcgg aatcttactc ggcttagccg ctctcacatt catcgggacg	180
cttatcgggc tcgccctcgc caccgcgtt ttcgtcatat tcagcccgat catcgtgccg	240
gccgttctaa caatcgggct tgctgttaca ggcttttttg cgtcggggac gttcgggttg	300
acgggtttga gtcattgtc gtatttgttc aatatggtta ggcagacggc tgggtcgggtg	360
cccgagtcct tggattatgt taaggggacg ttgcaggatg ccggtgagta tgccgggcag	420
aagacgaagg atttcgggca gaagattcag agcacggctc atgagatggg tgatcagggg	480
caggttggtg ttcattgtca agttggtggc gggaaagaag ggcgaaaaag cggatgatcg	540
acttgaggat tcaaggttga tattgtggaa taataatgtt gatgtaagtt tttagtgtta	600
tcaaagcttt gtttgtttgt ttgta	625

<210> 400
 <211> 471
 <212> DNA
 <213> Zea mays

<400> 400	
atggctgatc accaccgggg tgcgacggga ggtggcgggg gctacggcga cctccagcgc	60
gggggcggca tgcacggcga ggcgcagcag cagcagaagc agggcgccat gatgacggcg	120
ctcaaggccg cgacggccgc gaccttcggc gggtcgatgc tgggtgctgtc cgggctgatc	180
ctggccggca ccgtgatcgc gtcacggtg gccacccccg tgctggtgat cttcagcccc	240

pctau2012001598-seq1

gtgctggtgc	cggccgccat	cgcgctggcg	ctcatggcgg	ccgggttcgt	cacctccggc	300
ggcctcggcg	tcgctgcgct	gtccgtgttc	tcctggatgt	acaagtacct	gacgggcaag	360
cacccgccgg	ccgccgacca	gctggaccac	gccaaaggcg	ggctggcgtc	caaggcccg	420
gacgtcaagg	atgcagcaca	gcaccgcata	gaccaggcgc	aggggtcttg	a	471

<210> 401
 <211> 890
 <212> DNA
 <213> Brassica napus

<400> 401		
gtttccaaac	ctgatgattg	tagacgaatc gtcgatgaga ccatctccca ttttggcaga 60
ttggatcatc	ttgtaaacia	tgctggaata atgcagattt caatgttcga aaacattgaa 120
gaaataacta	ggacaagagc	agttatggat actaactttt ggggatcggg ttatacaact 180
cgtgctgcgc	ttccgtacct	tcgacaaagc aatggtaaga ttgtggccat gtcgtcctct 240
gcggcatggc	taaccgcccc	aaggatgagc ttttataatg ctagcaaagc cgcattgttg 300
aacttcttcg	agacgttgag	gattgagctt ggcagcgatg tacacattac aatcgtcaca 360
cctgggtata	ttgaatctga	actcacacia ggcaagtact tctctgggtga aggcgagcta 420
gtagtcaacc	aagacattag	agatgttcaa attggagcat ttccggtaac gtcggtatca 480
ggttggtcca	aggggatagt	gaaagggtgtg tgtaggaaac agagatacgt gaccgaacca 540
tcgtggttta	aggtgacgta	cctttggaaa gtgttttgtc cggaactgat cgagtggggg 600
tgcaggttgc	tgttcttgtc	cggacatggg acgtcggaga aaaatgcact caacaagaag 660
atcttggaca	tacctggtgt	acgtagtgtc ctataccctg aatctatcag aacgccagaa 720
atcaagtcgg	agtagagtga	ggttgatact taataagtgt ctcatatagt ggagccatgt 780
tttgtaaata	gactttctat	tatgcacatg ttactatgat gtatccgttt gtttatgtgt 840
ataagaataa	gtgaactttg	gagctccaaa aaaaaaaaaa aaaaaaaaaa 890

<210> 402
 <211> 1266
 <212> DNA
 <213> Brassica napus

<400> 402		
ggtcgacggc	cattacggcc	gggatcccgt taatgtgtat tgtgttttga tcgaagatgg 60
aattgataaa	cgatttcctc	aatctaactg cacctttctt tactttcttt ggtctttgct 120
tcttcttgcc	ccctttttat	ttcttcaagt tcgtgcagtc tatcttctcg acaattttct 180
ctgaaaacgt	atatgggaaa	gtggttctca tcaactggtg ttcctccggg atcggcgagc 240
aattggcata	tgagtacgca	agtaaagggt catgtttggc tctgactgcc cgaaggaaga 300
accgtctaga	ggaagtggca	gagattgtct gtgaagttgg atcccctaatt gttgtcacag 360
ttcacgctga	tgtttccaaa	cctgatgatt gtagacgaat cgtcgatgag accatctccc 420
attttggcag	attggatcat	cttgtaaaca atgctggaat aatgcaaatt tcaatgttcg 480

pctau2012001598-seq1

aaaacattga	agaaataact	aggacaagag	cagttatgga	tactaacttt	tggggagcgg	540
tttatacaac	tcgtgctgcg	cttccgtacc	ttcgacaaag	caatggtaag	attgtggcta	600
tgctgcctc	tgcgcatg	ctaaccgccc	caaggatgag	cttttataat	gctagcaaag	660
cagctttgtt	gaacttcttc	gagacgttga	ggattgagct	tggtagcgat	gtacacatta	720
caatcgtcac	acctggttat	attgaatctg	aactcacaca	aggcaagtac	gtctctggtg	780
aaggcgagct	agtagtcaac	caagacatta	gagatgttca	aattggagca	tttccggtaa	840
cgtcagtatc	aggtcgtgcc	aaggggatag	tgaaggtgt	gtgtaggaaa	gagagatacg	900
tgaccgaacc	atcgtggttt	aaggtgacgt	acctttggaa	agtgttttgt	ccggaactga	960
tcgagtgggg	ttgcagattg	atgttcttgt	ccggacatgg	tacgccggag	gaaaatgcac	1020
tcaacaagaa	gatcctggac	atacctggtg	tacgtagtgc	tctataccct	gaacctatca	1080
gaacgccaga	aatcaagtcg	gagtagagtg	aggttgatac	ttaataagtg	tctcataaag	1140
tggagccatg	ttttgtaaat	ggactttcta	ttatgcacat	gttactatga	tgtatctgtt	1200
tgtttatgtg	tataagaata	agtgaacttt	ggagctcaaa	aaaaaaaaaa	aaaaaaaaaa	1260
aaaaaa						1266

<210> 403
 <211> 1575
 <212> DNA
 <213> Brassica napus

<220>
 <221> misc_feature
 <222> (1486)..(1488)
 <223> n is a, c, g, or t

<400> 403	
atagcctagt	ccccaaaaaa actctgattg tgtcccaagt aaagagaaat ggtggatcta 60
ttgaactcgg	tgatgaatct ggtggctccg ccggccacga tgggtggtgat ggccttctca 120
tggccattac	tgtgtttcat taccttctcc gagcggcttt ataactctta ttttgtgact 180
gaagacatgg	aggataaagt cgtcgtcatc accggagctt ccccgccat cggagagcaa 240
atagcatacg	aatatgcaaa gaggggagca aacttagtgt tgggtggcgag gagagagcaa 300
agactgagag	ttgtgagcaa caacgccaga cagattggag ccaaccacgt catcatcatc 360
gctgctgatg	tcgtcaaaga agatgattgc cgccgtttca tcaactcaagc cgtcaactat 420
tacggtcgcg	tggatcacct agtgaattca gcgagtcttg gacacacgtt ttacttcgac 480
gaagtgagtg	acacgaccgt ctttccgcat ttgctggaca taaacttctg ggggaatgtc 540
tatccgacgt	atgtagcgtt gccacacctt caaaagacca atggccggat cgttgtgaac 600
gcatcggtcg	agaattggct gcctctaccg cggtatgagtc ttatttcgc tgcaaaagcg 660
gcactagtca	acttctacga gacgctgaga ttcgagctaa atggagatgt cggaataact 720
atcgcgactc	acgggtggat agggagtgag atgagtagag gaaagtcat gctagaagaa 780
ggtgccgaga	tgcagtggaa ggaagaaaga gaagtacctg caaatggtgg acctctagag 840

pctau2012001598-seql

gagtttgcaa	agatgattgt	ggcaggagct	tgcaggggag	acgcatacgt	gaagttccca	900
aactgggtacg	acgtgtttct	cctctacaga	gtcttcacgc	cgaatgtgct	gagatggaca	960
ttcaagctgc	tgctctctag	tgaggggttca	cgtcaaagct	cccttggttg	ggtcgggcaa	1020
ggctgcctc	cagaggaatc	ctcttcacaa	atgaaactta	tgcttgaagg	aggctcacct	1080
cgggtgactg	ctagcccacc	tcggtatact	cccagtccea	gcccaccta	tcacactgct	1140
agcccacctc	ggtatactcc	cagtccaagc	ccacctcatc	atacctaag	cccacaacgg	1200
tatactccca	gccaagccc	acccattat	acctaagcc	gacatcgcta	tacccaagc	1260
ccgagcccac	ctcattatac	cgaagccct	cctctgtata	ccgaaagtcc	acctaactat	1320
accacaagcc	caaattggta	taccgaaagc	cctcctcgat	acacccaag	ccaagccca	1380
cctcggtttt	cacgttttaa	tatccaagag	ttaccatagg	aagtgaatgg	agaatcaaga	1440
gtaagagtaa	ggatggtttg	gcgtgtagag	tgagcttgag	tttcannnct	ttctccaaat	1500
taggaagtca	acaaaagata	aatgcaatgt	aaaggacata	gaatcaagaa	taatgcagct	1560
atgttggtac	ttgtc					1575

<210> 404
 <211> 1188
 <212> DNA
 <213> zea mays

<400> 404						
attgcgca	ccttgatgct	gggcatgagc	aggacggggg	tggccggagc	ggcgctgcgg	60
gtggcgctga	cggcgctgct	gccgctggtc	ctgccggcgt	actacgtgta	caagctgacc	120
acgtacctcc	tcggcgccgt	cttccccgag	gacgtcgccg	gcaaggtggg	actcatcacc	180
ggcgccctcct	ccggcatcgg	cgagcacctg	gcctatgagt	acgcgaagcg	gggagcctac	240
ctggcgctgg	tggcgaggag	ggaggcgagc	ctccgcgagg	tcggcgacgt	cgcgctgggg	300
ctgggctcgc	cgggcgctct	cgtcctcccg	gctgacgtct	ccaagcctcg	ggactgcgag	360
ggcttcatcg	acgacacgat	tagctacttc	ggtagactgg	atcacctggg	gaacaacgcg	420
tccatctggc	aagtgtgcaa	gttcgaagag	atccaggacg	tcaggcactt	gagagccctg	480
atggacatca	acttctgggg	ccacgtgtac	ccaacccggc	tcgccatccc	tcacctcagg	540
agaagccgtg	gccggatcgt	gggcgtcacc	tccaactcgt	cctacatatt	catcgggagg	600
aacaccttct	acaatgccag	caaggcggcg	gcgctcagct	tctacgacac	cctgaggatg	660
gagctgggca	gcgacatccg	catcaccgag	gtggtgccag	gcgtgggtgga	gtctgagatc	720
accaagggga	agatgctcac	caagggaggc	gagatgaagg	tggaccagga	cgaagagac	780
gccatcctgg	ggccgacgcc	ggccgagccc	gtgggcgact	tcgccaggac	cgtggtgcgc	840
gacgtgtgcc	ggggcgcgag	gtacgtgttc	gagcccaggt	ggtacatggg	cgtctacttg	900
ctgcgggcct	gcctcccgga	agtcctggcc	tggaactccc	gcctgctcac	tgctgacacg	960
gtcggcgcgct	ccaccacgga	cacgctcgga	aagtggctgg	tcgagctgcc	cggcgtgcgc	1020

pctau2012001598-seq1

cgcgctcgtgc agccgccgctc gctccgctcg ccggagatca aggactagtg acgggtgatcg	1080
tgtacgttct gtggccatgg atagcactag ctgtatgaga ccgaagttcc tttagacatc	1140
gacacgatta ataaaggact caattatctt aaaaaaaaaa aaaaaaaaaa	1188

<210> 405
 <211> 1022
 <212> DNA
 <213> Brassica napus

<400> 405	
attacggccg gggattgcat tcaatctaga gaacaagaat aacgagatgg gtacggcgac	60
ggagataatg gagagagacg caatggctac ggtggctccc tacgctccgg tcacttttca	120
tcgccgtgct cgtgttgact tggatgatag acttcctaaa ccatatatgc caagagccct	180
gcaagcacca gatagagagc atccatatgg aacccagggc cataagaatt atggactaag	240
tgttcttcaa caacatgtcg ctttcttcga tatagatgat aatggcatta tctacccttg	300
ggagacctac tctggactgc gaatgattgg tttcaatatc attgggtcgc ttatcatagc	360
tgctgtaatc aacctggccc ttagctatgc tactcttccg ggatgggttac cttcgccgtt	420
cttcctata tacatacaca acatacaca gtcaaagcat ggaagcgact caagaacata	480
tgacaatgaa gggaggttta tgccggtgaa tcttgagttg atatttagca aatatgcaa	540
aaccttgcca gataaattga gtcttgaga attatgggac atgacagaag gacaacgtga	600
cgcattgggac attttcggat ggtttgcaag caaaatagag tggggactgt tgtacttgct	660
agcgagggat gaagaaggtt ttttatcaaa agaagctatt aggcggtgtt tcgatgggag	720
cttggtcgag tactgcgcca agatctacgt tggatcaat gaagacaaga cagcgacta	780
ctaaagtaat tgatcgagct gctttaggct tataatcgat ccatatgtgt tgttttgtta	840
tgtccaaata ttttaaggtaa aatataagtt gcagctctat taaaaagctt caatgtggaa	900
taatagtatg ttttggtgtt tttctttttg taccagcagt ctctttttcc atttatttcc	960
aatttaaatg caattatgca gtgttggtga gcaaaaaaaaa aaaaaaaaaa aaaaaaaaaa	1020
aa	1022

<210> 406
 <211> 1069
 <212> DNA
 <213> Brassica napus

<400> 406	
tcaattatta acacttacag aggagagaga aggagagcca aatggtaaga gagagtatgg	60
gagaagagtc cgaggctttt gccacgacgg cgccgttagc tccagtgacc ggtgagcgaa	120
aagtaaggaa cgacttgag gaaacattac caaaaccata tttggcaaga gcactagtag	180
ctccagatac agagcatccg aatggatcag aaggctatga cagcaaaggc atgagtgtta	240
cgcaacaaca tgttgctttt ttcgaccaa acggcgatgg aatcgtctat ccttgggaga	300
cttatgcggg attccgtgac ctcggtttca acccaatctc ttctgttttt tgggcatat	360

pctau2012001598-seq1

tcataaactt	tgcgttttagc	tacgtttacac	ttccgagttg	gttgccatca	ccactattgc	420
ctgtttatat	cgacaacatt	cacaaagcca	agcatgggag	tgattcaagc	acctatgata	480
ccgaaggaag	gtacgtcccc	gttaatctcg	agaacatctt	tagcaaatat	gcgttaacgg	540
ctccaaataa	aataacatta	aaagagcttt	ggaacttaac	cgagggaaac	cgaatggcaa	600
tcgatccttt	tggatggctt	gcgaataaag	ttgaatggct	actagtctat	cttcttgcaa	660
aggatgagga	agggttcgtg	tctaaagaag	ctgtgagagg	tgtctttgat	gcaagtttct	720
ttgaatactg	tgctaaaaag	aataaagaga	aggccgattc	tcgcaagcaa	gactaagcat	780
ttaatttgct	tatcttttaa	atttgtccgt	ttgtattatt	atgctctcaa	tcttaattag	840
gaaaaggaca	gaaaacaaaa	ctctatgtgg	aataaaacgt	ttgtgttctt	ggatatattg	900
tgtgtatatt	gtgaataata	aatgagatt	tcatcattcg	gaaagtgaag	tgatataatt	960
aaaacgagga	gagagcatga	aacaatgtgt	gcgcgcgctc	tctctagaga	gacaaatatt	1020
agcgcggggg	gcgctcgtct	tttttttctt	taaaaataaa	aaaaaaaaaa		1069

<210> 407
 <211> 1030
 <212> DNA
 <213> Sorghum indicum

<400> 407

ggcacgagag	agaaaaaagg	tgattttgtc	aagggaaata	tggcaactca	tgttttggct	60
gctgcggcgg	agagaaatgc	tgcgttggcg	ccggacgccc	cgcttgctcc	ggtgactatg	120
gagcgcgccag	tgcgcactga	cttgagact	tcgatcccg	agccctatat	ggcaagagga	180
ttggttgcac	ctgatatgga	tcacccaac	ggaacaccag	gccatgtgca	tgataatttg	240
agtgtgctgc	aacagcattg	tgctttcttt	gatcaggatg	ataacggaat	catctatcca	300
tgggagactt	actctggact	tcgccaaatt	ggtttcaatg	tgatagcttc	ccttataatg	360
gctatcgtca	ttaatgtggc	gctgagttat	cctactctcc	cgggttggat	tccttctcct	420
tttttcccca	tatatttgta	caacatacac	aaggccaaac	atggaagcga	ctccggaacc	480
tatgatactg	aaggaaggta	cctacctatg	aattttgaga	acctgttcag	caagcatgcc	540
cggacaatgc	ccgataggct	cactctaggg	gagctatgga	gcatgactga	agctaacaga	600
gaagcatttg	acattttcgg	ctggatcgca	agcaaaatgg	agtggactct	cctctacatt	660
cttgcaagag	accaggacgg	tttcctgtcg	aaagaagcca	tcaggcggtg	ttacgatggc	720
agtttgctcg	agtactgtgc	aaagatgcaa	aggggagccg	aggacaagat	gaaatgaagg	780
aaatcggcta	tcgcggtagg	tgtaagttat	gatgtggtgt	gtatgatgga	ttgaaagtgc	840
cagtgcctaa	gttgtgtggc	agagtcttgt	gtaataacct	tttgtgtacag	atttaaggtc	900
tcggaattgg	tgtaactgtg	gagaagatgt	tgactcctgt	ttttgttcaa	taagtccaac	960
tcttgacatt	tgggttggtt	gcagggaaag	atgggggaatt	ttgttttccg	aaaaaaaaaa	1020
aaaaaaaaaa						1030

pctau2012001598-seql

<210> 408
 <211> 1326
 <212> DNA
 <213> Zea mays

<400> 408
 atcccagcag caggtagcag cagcgcatga ctttgtatct gcaactcggt tcgactagcc 60
 tgcacgccgg gagccctcgt ctgcccttc tccacctccg aaggtagggt acctgatact 120
 tccctgcccc gcgcacgcga tccgcggttt tagtttgctg gcctcggttc cgatcgattc 180
 ctatcctccc tccggggcgg cgtccaccga cttcatcadc cactcgccgg ccgaaccgat 240
 ttgtttcggt tgccgttcga tccgcagatc ctgagagttc aaccgcgag tgagcatgtc 300
 gtcctactcc ccgccgccgc cgccgccgcg ggaccagtcc atggacaccg aggcacccaa 360
 cgcgcccatc accagggagc ggaggtcaa cccgatctg caggagcagc tccccaaagg 420
 atatctcgcg agagctctcg aggcggtgga cccgagccac ccgcagggga ccaaggggcg 480
 cgacccccgc ggcatgagcg tgcttcagca gcacgccgcc ttcttcgacc gcaatggcga 540
 cggcgatcgc taccctggg agacgtttca aggactgcga gcgataggat gtggactcac 600
 tgtatcattc gcgttctcca tactgatcaa cctcttcctc agttatcca ctcagccggg 660
 atggttacct tctcctttgc tgtccatccg tatagacaac atccacaagg gtaagcacgg 720
 gagtgattct gaaacctacg aactgaagg gaggtttgat ccatcaaagt tcgatgctat 780
 attcagtaag tacggtcgaa cccatcctaa tgctataaca agagacgagt taagctcaat 840
 gcttcaagga aaccgcaata cgtacgattt ccttggtctg ttggccgctg ccggtgaatg 900
 gctcttactc tacagcttgg cgaaagaca ggatggcctc ttgcagcgcg aaactgtccg 960
 tggctctatt gatgggagcc tatttgagcg actggaagac gacaacaaca agaagaaatc 1020
 gtcattgaat ctgagcagcc ttgtacagct cagggaagtg ctgtcagtag aaaactacca 1080
 gatataccat tggctcgtgt caaataacaa atgcttcggc tttgttcacg cgtcattaac 1140
 tatgagtgct gggatttggt tgtatgtgtg tcgtgctacc agtttcttct cctgtcgtct 1200
 cacacaggta ctgaattacg catgtgtttt ctagtgttcg tgcggaagct gtattataag 1260
 ctgaaaaatg tgcgtttgaa atttatgggc aaaactgtct tcttggctct aaaaaaaaaa 1320
 aaaaaa 1326

<210> 409
 <211> 15172
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> vector sequence

<400> 409
 gcttcctcgt gctttacggt atcgccgctc ccgattcgca gcgcacgcc ttctatcgcc 60
 ttcttgacga gttcttctga aacgcgtgat cgttcaaaca tttggcaata aagtttctta 120
 agattgaatc ctgttgccgg tcttgcatg attatcatat aatttctggt gaattacgtt 180

pctau2012001598-seq1

aagcatgtaa	taattaacat	gtaatgcatg	acgttatatta	tgagatgggt	ttttatgatt	240
agagtcccg	aattatacat	ttaatacgcg	atagaaaaca	aaatatagcg	cgcaaactag	300
gataaattat	cgcgcgcggt	gtcatctatg	ttactagatc	gacgtccgta	cggttaaaac	360
caccccagta	cattaaaaac	gtccgcaatg	tgttattaag	ttgtctaagc	gtcaatttgt	420
ttacaccaca	atatatcctg	ccaccagcca	gccaacagct	ccccgaccgg	cagctcggca	480
caaaatcacc	actcgataca	ggcagcccat	cagtccacta	gacgctcacc	gggctgggtg	540
ccctcgccgc	tgggctggcg	gccgtctatg	gccctgcaaa	cgcgccagaa	acgccgtcga	600
agccgtgtgc	gagacaccgc	agccgccggc	gttgtggata	cctcgcgga	aacttggccc	660
tcactgacag	atgagggg	gacgttgaca	cttgaggggc	cgactcacc	ggcgcgcggt	720
tgacagatga	ggggcaggct	cgatttcggc	cggcgacgtg	gagctggcca	gcctcgcaaa	780
tcggcgaaaa	cgctgattt	tacgcgagtt	ttccacagat	gatgtggaca	agcctgggga	840
taagtgcct	gcggtattga	cacttgagg	gcgcgactac	tgacagatga	ggggcgcgat	900
ccttgacact	tgaggggag	agtgtgaca	gatgaggggc	gcacctattg	acatttgagg	960
ggctgtccac	aggcagaaaa	tccagcattt	gcaaggggtt	ccgcccggtt	ttcggccacc	1020
gctaacctgt	cttttaacct	gcttttaaac	caatatattt	aaaccttggt	tttaaccagg	1080
gctgcgccct	gtgcgcgtga	ccgcgcacgc	cgaagggggg	tgccccccct	tctcgaaccc	1140
tccgggcccc	ctctcgcggt	ggcagcatca	cccataattg	tggttttcaa	atcggtcccg	1200
tcgatactat	gttatacgcc	aactttgaaa	acaactttga	aaaagctggt	ttctgggtatt	1260
taaggtttta	gaatgcaagg	aacagtgaat	tggagttcgt	cttggtataa	ttagcttctt	1320
ggggtattta	aatactgtag	aaaagaggaa	ggaaataata	aatggctaaa	atgagaatat	1380
caccggaatt	gaaaaaactg	atcgaaaaat	accgctgcgt	aaaagatacg	gaaggaatgt	1440
ctcctgctaa	ggtatataag	ctggtgggag	aaaatgaaaa	cctatatatta	aaaatgacgg	1500
acagccggtg	taaagggacc	acctatgatg	tggaacggga	aaaggacatg	atgctatggc	1560
tggaaggaaa	gctgcctggt	ccaaaggtcc	tgcaccttga	acggcatgat	ggctggagca	1620
atctgctcat	gagtgaggcc	gatggcgctc	tttgctcgga	agagtatgaa	gatgaacaaa	1680
gccctgaaaa	gattatcgag	ctgtatgcgg	agtgcacag	gctctttcac	tccatcgaca	1740
tatcggattg	tccctatacg	aatagcttag	acagccgctt	agccgaattg	gattacttac	1800
tgaataacga	tctggccgat	gtggattg	aaaactggga	agaagacacc	ccatttaaag	1860
atccgcgcga	gctgtatgat	tttttaaga	cggaaaagcc	cgaagaggaa	cttgtctttt	1920
cccacggcga	cctgggagac	agcaacatct	ttgtgaaaga	tggcaaagta	agtggcttta	1980
ttgatcttgg	gagaagcggc	agggcggaca	agtggatga	cattgccttc	tgcgctccgt	2040
cgatcaggga	ggatattggg	gaagaacagt	atgtcgagct	atTTTTTgac	ttactgggga	2100
tcaagcctga	ttgggagaaa	ataaaatatt	atTTTTTact	ggatgaattg	ttttagtacc	2160
tagatgtggc	gcaacgatgc	tggcgacaag	caggagcgca	ccgacttctt	ccgcatcaag	2220

pctau2012001598-seq1

tgttttggct	ctcaggccga	ggccccacggc	aagtatttgg	gcaagggggtc	gctggtatttc	2280
gtgcagggca	agattcggaa	taccaagtac	gagaaggacg	gccagacggt	ctacgggacc	2340
gacttcattg	ccgataaggt	ggattatctg	gacaccaagg	caccaggcgg	atcaaatacag	2400
gaataagggc	acattgcccc	ggcgtgagtc	ggggcaatcc	cgcaaggagg	gtgaatgaat	2460
cggacgtttg	accggaaggc	atacaggcaa	gaactgatcg	acgcgggggtt	ttccgccgag	2520
gatgccgaaa	ccatcgcaag	ccgcaccgtc	atgcgtgcgc	cccgcgaaac	cttcacgtcc	2580
gtcggctcga	tggcccagca	agctacggcc	aagatcgagc	gcgacagcgt	gcaactggct	2640
ccccctgccc	tgcccgcgcc	atcggccgcc	gtggagcggtt	cgcgctcgtct	cgaacaggag	2700
gcggcaggtt	tggcgaagtc	gatgaccatc	gacacgcgag	gaactatgac	gaccaagaag	2760
cgaaaaaccg	ccggcgagga	cctggcaaaa	caggtcagcg	aggccaagca	agccgcgttg	2820
ctgaaacaca	cgaagcagca	gatcaaggaa	atgcagcttt	ccttgttcga	tattgcgccg	2880
tggccggaca	cgatgcgagc	gatgccaaac	gacacggccc	gctctgccct	gttcaccacg	2940
cgcaacaaga	aaatcccgcg	cgaggcgctg	caaaacaagg	tcattttcca	cgtcaacaag	3000
gacgtgaaga	tcacctacac	cggcgctcag	ctgcggggccg	acgatgacga	actggtgtgg	3060
cagcaggtgt	tggagtacgc	gaagcgcacc	cctatcggcg	agccgatcac	cttcacgttc	3120
tacgagcttt	gccaggacct	gggctggtcg	atcaatggcc	ggtattacac	gaaggccgag	3180
gaatgcctgt	cgcgccctaca	ggcgacggcg	atgggcttca	cgtccgaccg	cgttgggcac	3240
ctggaatcgg	tgctcgtgct	gcaccgcttc	cgcgctcctgg	accgtggcaa	gaaaacgtcc	3300
cgttgccagg	tcctgatcga	cgaggaaatc	gtcgtgctgt	ttgctggcga	ccactacacg	3360
aaattcatat	gggagaagta	ccgcaagctg	tcgccgacgg	cccgcaggat	gttcgactat	3420
ttcagctcgc	accgggagcc	gtacccgctc	aagctggaaa	ccttcgcct	catgtgcgga	3480
tcggattcca	cccgcgtgaa	gaagtggcgc	gagcaggtcg	gcgaagcctg	cgaagagttg	3540
cgaggcagcg	gcctgggtga	acacgcctgg	gtcaatgatg	acctggtgca	ttgcaaacgc	3600
tagggccttg	tggggtcagt	tccggctggg	ggttcagcag	ccagcgcttt	actgagatcc	3660
tcttccgctt	cctcgctcac	tgactcgctg	cgctcggtcg	ttcggtcgcg	gcgagcggta	3720
tcagctcact	caaaggcggt	aatacggtta	tccacagaat	caggggataa	cgcaggaaag	3780
aacatgtgag	caaaaggcca	gcaaaaggcc	aggaaccgta	aaaaggccgc	gttgctggcg	3840
ttttccata	ggctccgccc	ccctgacgag	catcacaaaa	atcgacgctc	aagtcagagg	3900
tggcgaaacc	cgacaggact	ataaagatac	caggcgtttc	cccctggaag	ctccctcgtg	3960
cgctctcctg	ttccgaccct	gccgcttacc	ggatacctgt	ccgcctttct	cccttcggga	4020
agcgtggcgc	tttctcatag	ctcacgctgt	aggtatctca	gttcggtgta	ggtcgttcgc	4080
tccaagctgg	gctgtgtgca	cgaaccccc	gttcagccc	accgctgcgc	cttatccggt	4140
aactatcgtc	ttgagtccaa	cccggtaaga	cacgacttat	cgccactggc	agcagccact	4200
ggtaacagga	ttagcagagc	gaggtatgta	ggcggtgcta	cagagttctt	gaagtgggtg	4260

pctau2012001598-seq1

cctaactacg	gctacactag	aagaacagta	tttggatatct	gcgctctgct	gaagccagtt	4320
accttcggaa	aaagagttgg	tagctcttga	tccggcaaac	aaaccaccgc	tggtagcggt	4380
ggtttttttg	tttgcaagca	gcagattacg	cgcagaaaaa	aaggatctca	agaagatcct	4440
ttgatctttt	ctacggggtc	tgacgctcag	tggaacgaaa	actcacgtta	agggattttg	4500
gtcatgagat	tatcaaaaag	gatcttcacc	tagatccttt	tggatctcct	gtggttggca	4560
tgcacataca	aatggacgaa	cggataaacc	ttttcacgcc	cttttaata	tccgattatt	4620
ctaataaacg	ctcttttctc	ttaggtttac	ccgccaatat	atcctgtcaa	acactgatag	4680
tttaaactga	aggcgggaaa	cgacaatctg	ctagtggatc	tcccagtcac	gacgttgtaa	4740
aacgggcgcc	ctagaatcta	attattctat	tcagactaaa	ttagtataag	tattttttta	4800
atcaataaat	aataattaat	aatttattag	taggagtgat	tgaatttata	atatattttt	4860
tttaatcatt	taaagaatct	tatatcttta	aattgacaag	agtttttaa	ggggagagt	4920
ttatcatatc	acaagtagga	ttaatgtgtt	atagtttcac	atgcattacg	ataagttgtg	4980
aaagataaca	ttattatata	taacaatgac	aatcactagc	gatcgagtag	tgagagtcgt	5040
cttattacac	tttcttcctt	cgatctgtca	catggcggcg	gcccgaattc	tcacacaagg	5100
tagttgcaag	acactgaagt	ggtggtagt	gtagtagaag	aagcagaatc	ggtagaaaagg	5160
caagacaatg	gagaagatga	agatggtgga	gattctcttc	ccacaacgca	gcaatcaagg	5220
ttttcaaggt	taaggcactc	gtgctttcca	tcatcgaaca	tgaagtcgat	gttatcctcg	5280
aaagcaagct	cgttgaagag	ttctgggtac	tcaattgggt	tctcgttagc	aaggttttga	5340
tcggtaagga	atggggagaa	tccagtatcc	atcatgcaga	agttccaagc	aagttcgttg	5400
ttatctccgc	acctatccat	ttccatgatg	gtggaagaat	caatgcagca	gttaacaacg	5460
gcagcttcct	cagaatatcc	cacaatttca	gcctcttggt	gctcagcctt	ctcttcctct	5520
ttttcttctt	cctcttgagg	tggttcctca	acgtattggt	gcttaacctc	ttccctaggt	5580
tcctctttag	cttctctagt	ctcaacctct	tgcttagcct	caacaagaat	accctcttga	5640
tggttagcct	ggttaactgg	gaatgggaaa	acgcccttct	tcttaagcct	gtcgatgtag	5700
ttggagatat	cgaagttggt	aacagcgtta	gcacctctgt	actcaatagc	agccatatca	5760
taagcagctg	cagcctcttc	ttgagtgttg	taagttccga	ggtagaggta	cttgtttccg	5820
aaaactcttc	caatcctagc	ttcccatctt	ccgttatgat	gatgcctagc	aactccccta	5880
tacttagaaa	ctcccctaga	gaatccagat	gactgccttc	taagggaagc	aagatactct	5940
tctttggtca	ccctctgcat	ctcttcaagt	tctttggtgt	aagtctcagc	tgggaagtta	6000
agaatggtat	ctgggcccc	atacttaaga	gcagcaagat	cataggtatg	agcagcagcc	6060
tcttcagaat	cataagctcc	aaggtaaacc	tgcttgccct	tcttgttttg	gatggagttc	6120
caagaggact	tatcccaaag	gtgagcttcg	aatcttccag	tccatctatg	cctagtaaca	6180
cctctgtaga	tagatgacct	tctggtagaa	gctggagaag	ttgggttatg	agacttatcg	6240
ccagatggag	atgacttctt	agccctctta	gctctctttg	gtcttgagc	ttcagattga	6300

pctau2012001598-seq1

attgggctag	aggtagtagt	agaagaggac	actgaagaag	atggagaact	agagcaggta	6360
gaggtagtga	gcctcttctt	catgaattct	gttcttcttt	actctttgtg	tgactgaggt	6420
ttggcttagt	gctttgggtca	tctatatata	atgataacaa	caatgagaac	aagctttgga	6480
gtgatcggag	ggtctaggat	acatgagatt	caagtggact	aggatctaca	ccgttggatt	6540
ttgagtgtgg	atatgtgtga	ggttaatttt	acttggtaac	ggccacaaag	gcctaaggag	6600
agggtgttag	acccttatcg	gcttgaaccg	ctggaataat	gccacgtgga	agataattcc	6660
atgaatctta	tcgttatcta	tgagtgaaat	tgtgtgatgg	tggagtgggtg	cttgctcatt	6720
ttacttgcct	ggtggacttg	gccctttcct	tatggggaat	ttatatttta	cttactatag	6780
agctttcata	cctttttttt	accttggatt	tagttaatat	ataatggtat	gattcatgaa	6840
taaaaatggg	aaatttttga	atttgtactg	ctaaatgcat	aagattaggt	gaaactgtgg	6900
aatatatatt	tttttcattt	aaaagcaaaa	tttgcctttt	actagaatta	taaatataga	6960
aaaatatata	acattcaaat	aaaaatgaaa	ataagaactt	tcaaaaaaca	gaactatggt	7020
taatgtgtaa	agattagtcg	cacatcaagt	catctgttac	aatatgttac	aacaagtcac	7080
aagcccaaca	aagttagcac	gtctaaataa	actaaagagt	ccacgaaaat	attacaaatc	7140
ataagcccaa	caaagttatt	gatcaaaaaa	aaaaaacgcc	caacaaagct	aaacaaagtc	7200
caaaaaaaaa	ttctcaagtc	tccatcttcc	tttatgaaca	ttgaaaacta	tacacaaaac	7260
aagtcagata	aatctctttc	tgggcctgtc	ttcccaacct	cctacatcac	ttccctatcg	7320
gattgaatgt	tttacttgta	ccttttccgt	tgcaatgata	ttgatagtat	gtttgtgaaa	7380
actaataggg	ttaacaatcg	aagtcatgga	atatggattt	ggtccaagat	tttccgagag	7440
ctttctagta	gaaagcccat	caccagaaat	ttactagtaa	aataaatcac	caattaggtt	7500
tcttattatg	tgccaaattc	aatataatta	tagaggatat	ttcaaataaa	aacgtatgaa	7560
tgttattagt	aaatggtcag	gtaagacatt	aaaaaaatcc	tacgtcagat	attcaacttt	7620
aaaaattcga	tcagtgtgga	attgtacaaa	aatttgggat	ctactatata	tatataatgc	7680
tttacaacac	ttggattttt	ttttggaggc	tggaattttt	aatctacata	tttgtttttg	7740
ccatgcacca	actcattggt	tagtgtaata	ctttgatatt	gtcaaatata	tgtgttcgtg	7800
tatatattga	taagaatttc	tttgaccata	tacacacaca	catatatata	tatatatata	7860
tattatatat	catgcacttt	taattgaaaa	aataatatat	atatatatag	tgcatttttt	7920
ctaacaacca	tatatgttgc	gattgatctg	caaaaatact	gctagagtaa	tgaaaaatat	7980
aatctattgc	tgaaattatc	tcagatgtta	agattttctt	aaagtaaatt	ctttcaaatt	8040
ttagctaaaa	gtcttghtaat	aactaaagaa	taatacacaa	tctcgaccac	ggaaaaaaaa	8100
cacataataa	atttggggcc	cctagaatct	aattattcta	ttcagactaa	attagtataa	8160
gtattttttt	aatcaataaa	taataattaa	taattttatta	gtaggagtga	ttgaatttat	8220
aatatatatt	ttttaatcat	ttaaagaatc	ttatatcttt	aaattgacaa	gagtttttaa	8280
tgggggagag	gttatcatat	cacaagtagg	attaatgtgt	tatagtttca	catgcattac	8340

pctau2012001598-seq1

gataagttgt	gaaagataac	attattatat	ataacaatga	caatcactag	cgatcgagta	8400
gtgagagtcg	tcttattaca	ctttcttcct	tcgatctgtc	acatggcggc	ggcccgcggc	8460
cgcttcatta	ctcgagccag	gaggatggat	cgatgctggt	ctgagaccct	gctaccggtt	8520
gctgactgaa	ctgctcggca	cggtccttca	tttcacgggc	cttgctcgcc	aactttgtct	8580
tggccgactc	caactgatcc	gctccgggtg	gatgtttccc	cgtcaggtaa	cggtagatcc	8640
aggacagcac	agacagagcg	gcaacaccaa	atcccccgct	tgccagaaaa	cccgtccca	8700
acaggaagat	ggtgatgact	gcagatcaga	aaaactcaga	ttaatcgaca	aattcgatcg	8760
cacaaactag	aaactaacac	cagatctaga	tagaaatcac	aaatcgaaga	gtaattattc	8820
gacaaaactc	aaattatttg	aacaaatcgg	atgatattcta	tgaaacccta	atcgagaatt	8880
aagatgatat	ctaacgatca	aaccagaaa	atcgtcttcg	atctaagatt	aacagaatct	8940
aaaccaaaga	acatatagca	aattgggagc	gaacgaaaac	aaaatcgaag	attttgagag	9000
aataaggaac	acagaaattt	acctgcaggg	accagtacag	gcgagaagat	caccaggaga	9060
ggtgtggcga	ttgtcagcgc	aatgaccgtt	ccagccaggg	tcaaccgga	taacaccaac	9120
aggctacctc	cggcagtaac	cgcggtcgct	gcctttacaa	cacgctgagc	acgcggttgc	9180
agttgcaagt	ggggggcacg	tgtttggtgc	tgctgcccg	agtgtctctgc	catgggtttt	9240
tttaacggag	caagcgccg	ctgttcttct	ttactctttg	tgtgactgag	gtttggtcta	9300
gtgctttggt	catctatata	taatgataac	aacaatgaga	acaagctttg	gagtgatcgg	9360
agggtctagg	atacatgaga	ttcaagtgga	ctaggatcta	caccgttgga	ttttgagtgt	9420
ggatatgtgt	gaggttaatt	ttacttggtg	acggccacaa	aggcctaagg	agaggtgttg	9480
agacccttat	cggcttgaac	cgctggaata	atgccacgtg	gaagataatt	ccatgaatct	9540
tatcgttatc	tatgagtga	attgtgtgat	ggtggagtgg	tgcttgctca	ttttacttgc	9600
ctggtggact	tggccctttc	cttatgggga	atttatattt	tacttactat	agagctttca	9660
tacctttttt	ttaccttgga	tttagttaat	atataatggt	atgattcatg	aataaaaaatg	9720
ggaaattttt	gaatttgtag	tgctaaatgc	ataagattag	gtgaaactgt	ggaatatata	9780
tttttttcat	ttaaaagcaa	aatttgcctt	ttactagaat	tataaatata	gaaaaatata	9840
taacattcaa	ataaaaaatga	aaataagaac	tttcaaaaaa	cagaactatg	tttaatgtgt	9900
aaagattagt	cgcacatcaa	gtcatctgtt	acaatatggt	acaacaagtc	ataagcccaa	9960
caaagttagc	acgtctaaat	aaactaaaga	gtccacgaaa	atattacaaa	tcataagccc	10020
aacaaagtta	ttgatcaaaa	aaaaaaaaacg	cccaacaaag	ctaaacaaag	tccaaaaaaa	10080
acttctcaag	tctccatctt	cctttatgaa	cattgaaaac	tatacacaaa	acaagtcaga	10140
taaatctctt	tctgggcctg	tcttcccaac	ctcctacatc	acttccctat	cggattgaat	10200
gttttacttg	taccttttcc	gttgcaatga	tattgatagt	atgtttgtga	aaactaatag	10260
ggttaacaat	cgaagtcatg	gaatatggat	ttggtccaag	attttccgag	agctttctag	10320
tagaaagccc	atcaccagaa	atttactagt	aaaataaatc	accaattagg	tttcttatta	10380

pctau2012001598-seq1

tgtgccaaat	tcaatataat	tatagaggat	atttcaaatg	aaaacgtatg	aatgttatta	10440
gtaaattggtc	aggtaaagaca	ttaaaaaaat	cctacgtcag	atattcaact	ttaaaaattc	10500
gatcagtgtg	gaattgtaca	aaaatttggg	atctactata	tatatataat	gctttacaac	10560
acttggattt	ttttttggag	gctggaattt	ttaatctaca	tatttgtttt	ggccatgcac	10620
caactcattg	tttagtgtaa	tactttgatt	ttgtcaaata	tatgtgttcg	tgtatatattg	10680
tataagaatt	tctttgacca	tatacacaca	cacatatata	tatatatata	tatattatat	10740
atcatgcact	tttaattgaa	aaaataatat	atatatatat	agtgcatttt	ttctaacaac	10800
catatatgtt	gcgattgatc	tgcaaaaata	ctgctagagt	aatgaaaaat	ataatctatt	10860
gctgaaatta	tctcagatgt	taagattttc	ttaaagtaaa	ttctttcaaa	ttttagctaa	10920
aagtcttgta	ataactaaag	aataatacac	aatctcgacc	acggaaaaaa	aacacataat	10980
aaatttgggc	gcgccgcgta	ttggctagag	cagcttgcca	acatgggtgga	gcacgacact	11040
ctcgtctact	ccaagaatat	caaagatata	gtctcagaag	accaaagggc	tattgagact	11100
tttcaacaaa	gggtaatatc	gggaaacctc	ctcggattcc	attgcccagc	tatctgtcac	11160
ttcatcaaaa	ggacagtaga	aaaggaaggt	ggcacctaca	aatgccatca	ttgcgataaa	11220
ggaaaggcta	tcgttcaaga	tgcctctgcc	gacagtggtc	ccaaagatgg	acccccaccc	11280
acgaggagca	tcgtggaaaa	agaagacgtt	ccaaccacgt	cttcaaagca	agtggattga	11340
tgtgataaca	tgggtggagca	cgacactctc	gtctactcca	agaatatcaa	agatacagtc	11400
tcagaagacc	aaagggctat	tgagactttt	caacaaaggg	taatatcggg	aaacctcctc	11460
ggattccatt	gccagctat	ctgtcacttc	atcaaaagga	cagtagaaaa	ggaaggtggc	11520
acctacaaat	gccatcattg	cgataaagga	aaggctatcg	ttcaagatgc	ctctgccgac	11580
agtgttccca	aagatggacc	cccaccacag	aggagcatcg	tggaaaaaga	agacgttcca	11640
accacgtctt	caaagcaagt	ggattgatgt	gatattctcca	ctgacgtaag	ggatgacgca	11700
caatcccact	atccttcgca	agaccttcct	ctatataagg	aagttcattt	catttgagaga	11760
ggacacgctg	aaatcaccag	tctctctcta	caaatctatc	tctgcatcg	catggcgatt	11820
ttggattctg	ctggcgttac	tacggtgacg	gagaacgggtg	gcggagagtt	cgtcgatctt	11880
gataggcttc	gtcgacggaa	atcgagatcg	gattcttcta	acggacttct	tctctctggt	11940
tccgataata	attctccttc	ggatgatgtt	ggagctcccg	ccgacgttag	ggatcggatt	12000
gattccgttg	ttaacgatga	cgctcaggga	acagccaatt	tggccggaga	taataacggt	12060
ggtggcgata	ataacggtgg	tggaagaggc	ggcggagaag	gaagaggaaa	cgccgatgct	12120
acgtttacgt	atcgaccgtc	ggttccagct	catcggaggg	cgagagagag	tccacttagc	12180
tccgacgcaa	tcttcaaaca	gagccatgcc	ggattattca	acctctgtgt	agtagttctt	12240
attgctgtaa	acagtagact	catcatcgaa	aatcttatga	agtatggttg	gttgatcaga	12300
acggatttct	ggttttagttc	aagatcgctg	cgagattggc	cgcttttcat	gtgttgata	12360
tccctttcga	tctttccttt	ggctgccttt	acggttgaga	aattgggtact	tcagaaatac	12420

pctau2012001598-seq1

atatcagaac	ctgttgtcat	ctttcttcat	attattatca	ccatgacaga	ggttttgtat	12480
ccagtttacg	tcaccctaag	gtgtgattct	gcttttttat	caggtgtcac	tttgatgctc	12540
ctcacttgca	ttgtgtggct	aaagttgggt	tcttatgctc	atactagcta	tgacataaga	12600
tccctagcca	atgcagctga	taaggccaat	cctgaagtct	cctactacgt	tagcttgaag	12660
agcttggcat	atttcatggg	cgctcccaca	ttgtgttatc	agccaagtta	tccacgttct	12720
gcatgtatac	ggaaggggtg	ggtggctcgt	caatttgcaa	aactgggtcat	attcacccga	12780
ttcatgggat	ttataataga	acaatatata	aatcctattg	tcaggaactc	aaagcatcct	12840
ttgaaaggcg	atcttctata	tgctattgaa	agagtgttga	agctttcagt	tccaaattta	12900
tatgtgtggc	tctgcatggt	ctactgcttc	ttccaccttt	ggttaaacad	attggcagag	12960
cttctctgct	tcggggatcg	tgaattctac	aaagattggg	ggaatgcaaa	aagtgtggga	13020
gattactgga	gaatgtggaa	tatgcctggt	cataaatgga	tggttcgaca	tatatacttc	13080
ccgtgcttgc	gcagcaagat	accaaagaca	ctcgccatta	tcattgcttt	cctagtctct	13140
gcagtctttc	atgagctatg	catcgcagtt	ccttgtcgtc	tcttcaagct	atgggctttt	13200
cttgggatta	tgtttcaggt	gcctttgggt	ttcatcacia	actatctaca	ggaaagggtt	13260
ggctcaacgg	tggggaacat	gatcttctgg	ttcatcttct	gcattttcgg	acaaccgatg	13320
tgtgtgcttc	tttattacca	cgacctgatg	aaccgaaaag	gatcgatgtc	atgagcgatc	13380
gcgatcgttc	aaacatttgg	caataaagtt	tcttaagatt	gaatcctggt	gccgggtctt	13440
cgatgattat	catataattt	ctgttgaatt	acgttaagca	tgtaataatt	aacatgtaat	13500
gcatgacgtt	atttatgaga	tgggttttta	tgattagagt	cccgcaatta	tacatttaat	13560
acgcgataga	aaacaaaata	tagcgcgcaa	actaggataa	attatcgcgc	gcgggtgtcat	13620
ctatgttact	agatccctgc	agggcgattt	ggctagagca	gcttgccaac	atggtggagc	13680
acgacactct	cgtctactcc	aagaatatca	aagatacagt	ctcagaagac	caaagggcta	13740
ttgagacttt	tcaacaaagg	gtaatatcgg	gaaacctcct	cggattccat	tgcccagcta	13800
tctgtcactt	catcaaaagg	acagtagaaa	aggaaggtgg	cacctacaaa	tgccatcatt	13860
gcgataaagg	aaaggctatc	gttcaagatg	cctctgccga	cagtgggtccc	aaagatggac	13920
ccccaccac	gaggagcatc	gtggaaaaag	aagacgttcc	aaccacgtct	tcaaagcaag	13980
tggattgatg	tgataacatg	gtggagcacg	acactctcgt	ctactccaag	aatatcaaag	14040
atacagtctc	agaagaccaa	agggctattg	agacttttca	acaaagggtta	atatcgggaa	14100
acctcctcgg	attccattgc	ccagctatct	gtcacttcat	caaaaggaca	gtagaaaagg	14160
aaggtggcac	ctacaaatgc	catcattgcg	ataaaggaaa	ggctatcggt	caagatgcct	14220
ctgccgacag	tgggtccaaa	gatggacccc	caccacagag	gagcatcgtg	gaaaaagaag	14280
acgttccaac	cacgtcttca	aagcaagtgg	attgatgtga	tatctccact	gacgtaaggg	14340
atgacgcaca	atcccactat	ccttcgcaag	accttctctt	atataaggaa	gttcattttca	14400
tttggagagg	acacgctgaa	atcaccagtc	tctctctaca	aatctatctc	tctcgagatg	14460

pctau2012001598-seq1

attgaacaag atggattgca cgcaggttct ccggccgctt gggaggagag gctattcggc	14520
tatgactggg cacaacagac aatcggctgc tctgatgccg ccgtgttccg gctgtcagcg	14580
caggggaggc cggttctttt tgtcaagacc gacctgtccg gtgccctgaa tgaacttcaa	14640
gacgaggcag cgcggctatc gtggctggcc acgacgggcg ttccttgccg agctgtgctc	14700
gacgttgctca ctgaagcggg aagggactgg ctgctattgg gcgaagtgcc ggggcaggat	14760
ctcctgtcat ctcaccttgc tcctgccgag aaagtatcca tcatggctga tgcaatgcgg	14820
cggctgcata cgcttgatcc ggctacctgc ccattcgacc accaagcgaa acatcgcac	14880
gagcgagcac gtactcggat ggaagccggg cttgtcgatc aggatgatct ggacgaagag	14940
catcaggggc tcgcgccagc cgaactgttc gccaggctca aggcgcgcac gcccgacggc	15000
gaggatctcg tcgtgactca tggcgatgcc tgcttgccga atatcatggt ggaaaatggc	15060
cgcttttctg gattcatcga ctgtggccgg ctgggtgtgg cggaccgcta tcaggacata	15120
gcgttggtta cccgtgatat tgctgaagag cttggcggcg aatgggctga cc	15172

<210> 410
 <211> 17146
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> vector sequence

<400> 410	
gcttcctcgt gctttacggt atcgccgctc ccgattcgca gcgcacgcc ttctatcgcc	60
ttcttgacga gttcttctga aacgcgtgat cgttcaaaca tttggcaata aagtttctta	120
agattgaatc ctgttgccgg tcttgcatg attatcatat aatttctgtt gaattacgtt	180
aagcatgtaa taattaacat gtaatgcatg acgttattta tgagatgggt ttttatgatt	240
agagtccgc aattatacat ttaatacgcg atagaaaaca aaatatagcg cgcaaactag	300
gataaattat cgcgcgcggg gtcattctatg ttactagatc gacgtccgta cggttaaaac	360
caccccagta cattaaaaac gtccgcaatg tgttattaag ttgtctaagc gtcaatttgt	420
ttacaccaca atatacctg ccaccagcca gccaacagct ccccgaccgg cagctcggca	480
caaaatcacc actcgataca ggcagcccat cagtccacta gacgctcacc gggctgggtg	540
ccctcgccgc tgggctggcg gccgtctatg gccctgcaaa cgcgccagaa acgccgtcga	600
agccgtgtgc gagacaccgc agccgccggc gttgtggata cctcgcggaa aacttgccc	660
tcactgacag atgaggggag gacgttgaca cttgaggggc cgactcacc ggcgcggcgt	720
tgacagatga ggggcaggct cgatttcggc cggcgacgtg gagctggcca gcctcgcaaa	780
tcggcgaaaa cgcctgattt tacgcgagtt tcccacagat gatgtggaca agcctgggga	840
taagtgccct gcggtattga cacttgaggg gcgcgactac tgacagatga ggggcgcgat	900
ccttgacact tgagggggcag agtgctgaca gatgaggggc gcacctattg acatttgagg	960
ggctgtccac aggcagaaaa tccagcattt gcaagggttt ccgccggtt ttcggccacc	1020

pctau2012001598-seq1

gctaacctgt	cttttaacct	gctttttaaac	caatatattat	aaaccttggt	tttaaccagg	1080
gctgcgccct	gtgcgcgtga	ccgcgcacgc	cgaagggggg	tgccccccct	tctcgaaccc	1140
tccggggccg	ctctcgcgtt	ggcagcatca	cccataattg	tggtttcaaa	atcgggtccg	1200
tcgatactat	gttatacgcc	aactttgaaa	acaactttga	aaaagctggt	ttctggtatt	1260
taaggtttta	gaatgcaagg	aacagtgaat	tggagttcgt	cttggtataa	ttagcttctt	1320
ggggtattta	aatactgtag	aaaagaggaa	ggaataata	aatggctaaa	atgagaatat	1380
caccggaatt	gaaaaaactg	atcgaaaaat	accgctgcgt	aaaagatacg	gaaggaatgt	1440
ctcctgctaa	ggtatataag	ctggtgggag	aaaatgaaaa	cctatatatta	aaaatgacgg	1500
acagccggta	taaagggacc	acctatgatg	tggaacggga	aaaggacatg	atgctatggc	1560
tggaaggaaa	gctgcctggt	ccaaagggtc	tgcaccttga	acggcatgat	ggctggagca	1620
atctgctcat	gagtgaggcc	gatggcgtcc	tttgctcgga	agagtatgaa	gatgaacaaa	1680
gccctgaaaa	gattatcgag	ctgtatgcgg	agtgcacgag	gctctttcac	tccatcgaca	1740
tatcggattg	tccctatacg	aatagcttag	acagccgctt	agccgaattg	gattacttac	1800
tgaataacga	tctggccgat	gtggattgcg	aaaactggga	agaagacacc	ccatttaaag	1860
atccgcgcga	gctgtatgat	tttttaaaga	cggaaaagcc	cgaagaggaa	cttgtctttt	1920
cccacggcga	cctgggagac	agcaacatct	ttgtgaaaga	tggcaaagta	agtggcttta	1980
ttgatcttgg	gagaagcggc	agggcggaca	agtggtatga	cattgccttc	tgcgtccggt	2040
cgatcaggga	ggatattggg	gaagaacagt	atgtcgagct	atTTTTtgac	ttactgggga	2100
tcaagcctga	ttgggagaaa	ataaaatatt	atattttact	ggatgaattg	ttttagtacc	2160
tagatgtggc	gcaacgatgc	tggcgacaag	caggagcgca	ccgacttctt	ccgcatcaag	2220
tgTTTTggct	ctcaggccga	ggccacggc	aagtatttgg	gcaaggggtc	gctggtattc	2280
gtgcagggca	agattcggaa	taccaagtac	gagaaggacg	gccagacggt	ctacgggacc	2340
gacttcattg	ccgataaggt	ggattatctg	gacaccaagg	caccaggcgg	atcaaatcag	2400
gaataagggc	acattgcccc	ggcgtgagtc	ggggcaatcc	cgcaaggagg	gtgaatgaat	2460
cggacgtttg	accggaaggc	atacaggcaa	gaactgatcg	acgcgggggt	ttccgccgag	2520
gatgccgaaa	ccatcgcaag	ccgcaccgtc	atgcgtgcgc	cccgcgaaac	cttccagtcc	2580
gtcggctcga	tggcccagca	agctacggcc	aagatcgagc	gcgacagcgt	gcaactggct	2640
ccccctgccc	tgcccgcgcc	atcggccgcc	gtggagcggt	cgcgtcgtct	cgaacaggag	2700
gcggcagggt	tggcgaagtc	gatgaccatc	gacacgcgag	gaactatgac	gaccaagaag	2760
cgaaaaaccg	ccggcgagga	cctggcaaaa	caggtcagcg	aggccaagca	agccgcgttg	2820
ctgaaacaca	cgaagcagca	gatcaaggaa	atgcagcttt	ccttgttcga	tattgcgccg	2880
tggccggaca	cgatgcgagc	gatgccaaac	gacacggccc	gctctgccct	gttcaccacg	2940
cgcaacaaga	aaatcccgcg	cgaggcgctg	caaaacaagg	tcattttcca	cgtcaacaag	3000
gacgtgaaga	tcacctacac	cggcgtcgag	ctgcgggccg	acgatgacga	actggtgtgg	3060

pctau2012001598-seq1

cagcaggtgt	tggagtacgc	gaagcgcacc	cctatcggcg	agccgatcac	cttcacgttc	3120
tacgagcttt	gccaggacct	gggctggtcg	atcaatggcc	ggtattacac	gaaggccgag	3180
gaatgcctgt	cgcgccctaca	ggcgacggcg	atgggcttca	cgtccgaccg	cgttggggcac	3240
ctggaatcgg	tgtcgtctgct	gcaccgcttc	cgcgctcctgg	accgtggcaa	gaaaacgtcc	3300
cgttgccagg	tcctgatcga	cgaggaaatc	gtcgtgctgt	ttgctggcga	ccactacacg	3360
aaattcatat	gggagaagta	ccgcaagctg	tcgccgacgg	cccgaaggat	gttcgactat	3420
ttcagctcgc	accggggagcc	gtacccgctc	aagctggaaa	ccttccgcct	catgtgcgga	3480
tcggattcca	cccgcgtgaa	gaagtggcgc	gagcaggtcg	gcgaagcctg	cgaagagttg	3540
cgaggcagcg	gcctgggtga	acacgcctgg	gtcaatgatg	acctggtgca	ttgcaaacgc	3600
tagggccttg	tggggctcagt	tccggctggg	ggttcagcag	ccagcgcttt	actgagatcc	3660
tcttccgctt	cctcgtctac	tgactcgtcg	cgctcggtcg	ttcggctgcg	gcgagcggta	3720
tcagctcact	caaaggcggg	aatacggtta	tccacagaat	caggggataa	cgcaggaaaag	3780
aacatgtgag	caaaaggcca	gcaaaaggcc	aggaaccgta	aaaaggccgc	gttgctggcg	3840
ttttccata	ggctccgccc	ccctgacgag	catcacaaaa	atcgacgctc	aagtcagagg	3900
tggcgaaacc	cgacaggact	ataaagatac	caggcgcttc	cccctggaag	ctccctcgtg	3960
cgctctcctg	ttccgaccct	gccgcttacc	ggatacctgt	ccgcctttct	cccttcggga	4020
agcgtggcgc	tttctcatag	ctcacgctgt	aggtatctca	gttcggtgta	ggtcgttcgc	4080
tccaagctgg	gctgtgtgca	cgaaccccc	gttcagcccc	accgctgcgc	cttatccggt	4140
aactatcgtc	ttgagtccaa	cccggtaaga	cacgacttat	cgccactggc	agcagccact	4200
ggtaacagga	ttagcagagc	gaggtatgta	ggcggtgcta	cagagtctct	gaagtgggtg	4260
cctaactacg	gctacactag	aagaacagta	tttggtatct	gcgctctgct	gaagccagtt	4320
accttcggaa	aaagagttgg	tagctcttga	tccggcaaac	aaaccaccgc	tggtagcggg	4380
ggtttttttg	tttgcaagca	gcagattacg	cgcagaaaaa	aaggatctca	agaagatcct	4440
ttgatctttt	ctacggggtc	tgacgctcag	tggaacgaaa	actcacgtta	agggattttg	4500
gtcatgagat	tatcaaaaag	gatcttcacc	tagatccttt	tggatctcct	gtggttgga	4560
tgcacataca	aatggacgaa	cggataaacc	ttttcacgcc	cttttaata	tccgattatt	4620
ctaataaacg	ctcttttctc	ttaggttttac	ccgccaatat	atcctgtcaa	acactgatag	4680
tttaaaactga	aggcgggaaa	cgacaatctg	ctagtggatc	tcccagtcac	gacgttgtaa	4740
aacgggcgtc	tgcgatcgct	gaagttccta	tacttttcag	agaataggaa	cttcggaata	4800
ggaacttccc	atgggatcta	gtaacataga	tgacaccgcg	cgcgataatt	tatcctagtt	4860
tgcgcgctat	attttgtttt	ctatcgcgta	ttaaattgat	aattgcggga	ctctaatacat	4920
aaaaacccat	ctcataaata	acgtcatgca	ttacatgtta	attattacat	gcttaacgta	4980
attcaacaga	aattatatga	taatcatcgc	aagaccggca	acaggattca	atcttaagaa	5040
actttattgc	caaatgtttg	aacgatcacg	ctagcggata	acaatttcac	acagggatat	5100

pctau2012001598-seq1

cactagtaaa	aggtaccgag	ctcctgcagt	atcgatgcgg	ccgcaaagtc	gacgaattct	5160
cattagcaga	actcaagatg	ctgatcctct	ggaacgttga	acttgagctt	gtgttcctcg	5220
aaaagcttgc	acaactcttt	gatgtaacgc	tggtgaagtc	tatcaacttc	ctctctagaa	5280
ggctgaggag	tcatttgaac	ctcgataggc	tttccaacga	tagtagtgat	aggctgtctg	5340
aaaggcatga	gtccgaaaga	gtattggaaa	actccccttc	catggaaaag	tggaaggctg	5400
attcccataa	tcttttggag	tctgttcttg	atccatctaa	gccaagttcc	aggagtgttc	5460
tcaacctggg	tgaagagggt	gttctctccg	aatgagaaga	taggaacaag	agcagcacca	5520
tgcataagag	caagtctgat	gaatccctta	cggttcttca	agagaagtct	gtaagcacca	5580
ggcttagcat	caagagcctc	ttgagcacct	ccaacgatga	tagcaagaag	gtttccacca	5640
ccctttctgc	taaggatgtg	atcagcagaa	actttctcgc	tagacacgag	tccaccagac	5700
atgatgtaat	ctctgaagaa	tggagccctg	aaccaaacgg	taagcatcat	aaggtaggat	5760
ctgattccag	ggaacaaaga	ggtgaatcca	gtagactcag	tacagagggt	aaggaaagca	5820
ccagcagcaa	gaacaccatg	aggatggaat	ccagcaatgt	agttacggct	aggatcaagc	5880
tcagcagtct	taacgagaga	cacaggggaag	taatccttca	tgtacttcca	gatggccaat	5940
cttctgaaga	attggatagg	tctaccacct	tgtctagggt	tatcccaatc	caagtaccac	6000
caggtagcgt	aaagaacaga	gaaaagccag	aacctggtga	acaagagtcc	aacgaagata	6060
acgatgcaga	gttgagcaag	agcaaggaat	gagaaaaccc	actgaagaac	agcgaaagtc	6120
tgcaatcttc	tctccaagga	aacaagaagt	ggagcgaact	cgaccatgaa	ttcagtcctc	6180
cgtgttctct	ccaaatgaaa	tgaacttcct	tatatagagg	aagggctctg	cgaaggatag	6240
tgggattgtg	cgtcatccct	tacgtcagtg	gagatatcac	atcaatccac	ttgctttgaa	6300
gacgtgggtg	gaacgtcttc	ttttccacg	atgctcctcg	tgggtggggg	tccatctttg	6360
ggaccactgt	cggcagaggc	atcttcaacg	atggcctttc	ctttatcgca	atgatggcat	6420
ttgtaggagc	caccttcctt	ttccactatc	ttcacaataa	agtgacagat	agctgggcaa	6480
tggaatccga	ggaggtttcc	ggatattacc	ctttgttgaa	aagtctcaat	tgccctttgg	6540
tcttctgaga	ctgtatcttt	gatatttttg	gagtagacaa	gtgtgtcgtg	ctccaccatg	6600
ttgacgaaga	ttttcttctt	gtcattgagt	cgtaagagac	tctgtatgaa	ctgttcgcca	6660
gtctttacgg	cgagttctgt	taggtcctct	atttgaatct	ttgactccat	gggatccaag	6720
ggccctagaa	tctaattatt	ctattcagac	taaattagta	taagtatttt	tttaatcaat	6780
aaataataat	taataattta	ttagtaggag	tgattgaatt	tataatatat	tttttttaat	6840
catttaaaga	atcttatatc	tttaaattga	caagagtttt	aaatggggag	agtgttatca	6900
tatcacaagt	aggattaatg	tgttatagtt	tcacatgcat	tacgataagt	tgtgaaagat	6960
aacattatta	tatataacaa	tgacaatcac	tagcgatcga	gtagtgagag	tcgtcttatt	7020
acactttctt	ccttcgatct	gtcacatggc	ggcggcccga	attctcacac	aaggtagttg	7080
caagacactg	aagtgggtgg	agtggtagta	gaagaagcag	aatcggtaga	aaggcaagac	7140

pctau2012001598-seq1

aatggagaag	atgaagatgg	tggagattct	cttcccacaa	cgagcaatc	aaggttttca	7200
aggttaaggc	actcgtgctt	tccatcatcg	aacatgaagt	cgatgttatc	ctcgaaagca	7260
agctcgttga	agagttctgg	gtactcaatt	gggttctcgt	tagcaagggt	ttgatcggt	7320
aggaatgggg	agaatccagt	atccatcatg	cagaagttcc	aagcaagttc	gttgttatct	7380
ccgcacctat	ccatttccat	gatggtggaa	gaatcaatgc	agcagttaac	aacggcagct	7440
tcctcagaat	atcccacaat	ttcagcctct	tgttgctcag	ccttctcttc	ctctttttct	7500
tcttcctctt	gaggtgggtc	ctcaacgtat	tgttgcttaa	cctcttccct	aggttcctct	7560
ttagcttctc	tagtctcaac	ctcttgctta	gcctcaacaa	gaataccctc	ttgatgggt	7620
gcctgggtta	ctgggaatgg	gaaaacgccc	ttcttcttaa	gcctgtcgat	gtagttggag	7680
atatcgaagt	tggtaacagc	gtagcacct	ctgtactcaa	tagcagccat	atcataagca	7740
gctgcagcct	cttcttgagt	gttgtaagtt	ccgaggtaga	ggtagttgtt	tccgaaaact	7800
cttccaatcc	tagcttccca	tcttccgtta	tgatgatgcc	tagcaactcc	cctatactta	7860
gaaactcccc	tagagaatcc	agatgactgc	cttctaaggg	aagcaagata	ctcttctttg	7920
gtcacctctc	gcattctctc	aagttctttg	gtgtaagtct	cagctgggaa	gttaagaatg	7980
gtatctgggc	cccaatactt	aagagcagca	agatcatagg	tatgagcagc	agcctcttca	8040
gaatcataag	ctccaaggta	aacctgcttg	cccttcttgt	tttgatgga	gttccaagag	8100
gacttatccc	aaagggtgagc	ttcgaatctt	ccagtccatc	tatgcctagt	aacacctctg	8160
tagatagatg	accttctggt	agaagctgga	gaagttgggt	tatgagactt	atcgccagat	8220
ggagatgact	tcttagccct	cttagctctc	tttggctctg	gagcttcaga	ttgaattggg	8280
ctagaggtag	tagtagaaga	ggacactgaa	gaagatggag	aactagagca	ggtagaggta	8340
gtgagcctct	tcttcatgaa	ttctgttctt	ctttactctt	tgtgtgactg	aggtttggtc	8400
tagtgctttg	gtcatctata	tataatgata	acaacaatga	gaacaagctt	tggagtgatc	8460
ggaggggtcta	ggatacatga	gattcaagtg	gactaggatc	tacaccgttg	gattttgagt	8520
gtggatatgt	gtgagggtta	ttttacttgg	taacggccac	aaaggcctaa	ggagaggtgt	8580
tgagaccctt	atcggttga	accgctggaa	taatgccacg	tggaagataa	ttccatgaat	8640
cttatcggtta	tctatgagtg	aaattgtgtg	atgggtggagt	ggtagcttgct	cattttactt	8700
gcctgggtgga	cttggccctt	tccttatggg	gaatttatat	tttacttact	atagagcttt	8760
catacctttt	ttttaccttg	gatttagtta	atatataatg	gtatgattca	tgaataaaaa	8820
tgggaaattt	ttgaatttgt	actgctaaat	gcataagatt	aggtgaaact	gtggaatata	8880
tatttttttc	atttaaaagc	aaaatttgcc	ttttactaga	attataaata	tagaaaaata	8940
tataacattc	aaataaaaaat	gaaaataaga	actttcaaaa	aacagaacta	tgtttaaatgt	9000
gtaaagatta	gtcgcacatc	aagtcattctg	ttacaatatg	ttacaacaag	tcataagccc	9060
aacaaagtta	gcacgtctaa	ataaactaaa	gagtccacga	aaatattaca	aatcataagc	9120
ccaacaaagt	tattgatcaa	aaaaaaaaaa	cgcccaacaa	agctaaacaa	agtcacaaaa	9180

pctau2012001598-seq1

aaacttctca	agtctccatc	ttccttttatg	aacattgaaa	actatacaca	aaacaagtca	9240
gataaatctc	tttctggggc	tgtcttccca	acctcctaca	tcacttccct	atcggattga	9300
atgttttact	tgtacctttt	ccgttgcaat	gatattgata	gtatgtttgt	gaaaactaat	9360
agggttaaca	atcgaagtca	tggaatatgg	at ttgtgtcca	agat tttccg	agagctttct	9420
agtagaaagc	ccatcaccag	aaatttacta	gtaaaataaa	tcaccaatta	ggtttcttat	9480
tatgtgccaa	attcaatata	attatagagg	atattttcaa	tgaaaacgta	tgaatgttat	9540
tagtaaatgg	tcaggtaaga	cattaaaaaa	atcctacgtc	agatattcaa	ctttaaaaat	9600
tcgatcagtg	tggaattgta	caaaaatttg	ggatctacta	tatatatata	atgctttaca	9660
acacttggat	tttttttttg	aggctggaat	ttttaatcta	catatttggt	ttggccatgc	9720
accaactcat	tgttttagtgt	aatactttga	ttttgtcaaa	tatatgtgtt	cgtgtatatt	9780
tgtataagaa	tttctttgac	catatacaca	cacacatata	tatatatata	tatatattat	9840
atatcatgca	cttttaattg	aaaaaataat	atatatatat	atagtgcatt	ttttctaaca	9900
accatatatg	ttgcgattga	tctgcaaaaa	tactgctaga	gtaatgaaaa	atataatcta	9960
ttgctgaaat	tatctcagat	gttaagattt	tcttaaagta	aattctttca	aatttttagct	10020
aaaagtcttg	taataactaa	agaataatac	acaatctcga	ccacggaaaa	aaaacacata	10080
ataaatttgg	ggcccctaga	atctaattat	tctattcaga	ctaaattagt	ataagtattt	10140
ttttaatcaa	taaataataa	ttaataattt	attagtagga	gtgattgaat	ttataatata	10200
ttttttttaa	tcatttaaa	aatcttata	ctttaaattg	acaagagttt	taaatgggga	10260
gagtgttatc	atatcacaag	taggattaat	gtgttatagt	ttcacatgca	ttacgataag	10320
ttgtgaaaga	taacattatt	atatataaca	atgacaatca	ctagcgatcg	agtagtgaga	10380
gtcgtcttat	tacactttct	tccttcgatc	tgtcacatgg	cggcggcccg	cggccgcttc	10440
attactcgag	ccaggaggat	ggatcgatgc	tggtctgaga	ccctgctacc	ggttgctgac	10500
tgaactgctc	ggcacggtcc	ttcattttcac	gggccttgct	cgccaacttt	gtcttggccg	10560
actccaactg	atccgctccg	ggtggatgtt	tccccgtcag	gtaacggtag	atccaggaca	10620
gcacagacag	agcggcaaca	ccaaatcccc	cgcttgccag	aaaacccgct	cccaacagga	10680
agatgggtgat	gactgcagat	cagaaaaact	cagattaatc	gacaaattcg	atcgcacaaa	10740
ctagaaacta	acaccagatc	tagatagaaa	tcacaaatcg	aagagtaatt	attcgacaaa	10800
actcaaatta	tttgaacaaa	tcggatgata	tctatgaaac	cctaatcgag	aattaagatg	10860
atatctaacg	atcaaaccga	gaaaatcgtc	ttcgatctaa	gattaacaga	atctaaacca	10920
aagaacatat	acgaaattgg	gatcgaacga	aaacaaaatc	gaagattttg	agagaataag	10980
gaacacagaa	atttacctgc	agggaccagt	acaggcgaga	agatcaccag	gagagggtgtg	11040
gcgattgtca	gcgcaatgac	cgttccagcc	aggggtcaacc	cggataacac	caacaggcta	11100
cctccggcag	taaccgcggg	cgctgccttt	acaacacgct	gagcacgcgg	ttgcagttgc	11160
aagtgggggg	cacgtgtttg	ttgctgctgc	ccgtagtgct	ctgccatggt	tttttttaac	11220

pctau2012001598-seq1

ggagcaagcg	gccgctgttc	ttctttactc	tttgtgtgac	tgaggtttgg	tctagtgcctt	11280
tggatcatcta	tatataatga	taacaacaat	gagaacaagc	tttggagtga	tcggaggggtc	11340
taggatacat	gagattcaag	tggactagga	tctacaccgt	tggattttga	gtgtggatat	11400
gtgtgagggtt	aattttactt	ggtaacggcc	acaaaggcct	aaggagaggt	gttgagaccc	11460
ttatcggcctt	gaaccgctgg	aataatgcca	cgtggaagat	aattccatga	atcttatcgt	11520
tatctatgag	tgaaattgtg	tgatgggtgga	gtgggtgcttg	ctcattttac	ttgcctgggtg	11580
gacttggccc	tttccttatg	gggaatttat	attttactta	ctatagagct	ttcatacctt	11640
ttttttacct	tggatttagt	taatataata	tggtatgatt	catgaataaa	aatgggaaat	11700
ttttgaattt	gtactgctaa	atgcataaga	ttaggtgaaa	ctgtggaata	tatatTTTTT	11760
tcatttaaaa	gcaaaatttg	ccttttacta	gaattataaa	tatagaaaaa	tatataacat	11820
tcaaataaaa	atgaaaataa	gaactttcaa	aaaacagaac	tatgtttaat	gtgtaaagat	11880
tagtcgcaca	tcaagtcac	tgttacaata	tgttacaaca	agtcataagc	ccaacaaagt	11940
tagcacgtct	aaataaacta	aagagtccac	gaaaatatta	caaatcataa	gccaacaaaa	12000
gttattgatc	aaaaaaaaaa	aacgccaac	aaagctaaac	aaagtccaaa	aaaaacttct	12060
caagtctcca	tcttccttta	tgaacattga	aaactataca	caaaacaagt	cagataaatc	12120
tctttctggg	cctgtcttcc	caacctccta	catcacttcc	ctatcggatt	gaatgtttta	12180
cttgtagctt	ttccgttgca	atgatattga	tagtatgttt	gtgaaaacta	ataggggttaa	12240
caatcgaagt	catggaatat	ggatttggtc	caagattttc	cgagagcttt	ctagtagaaa	12300
gcccatacacc	agaaatttac	tagtaaaata	aatcaccaat	taggtttctt	attatgtgcc	12360
aaattcaata	taattataga	ggatatttca	aatgaaaacg	tatgaatgtt	attagtaaat	12420
ggtcaggtaa	gacattaaaa	aaatcctacg	tcagatattc	aactttaaaa	attcgatcag	12480
tgtggaattg	tacaaaaatt	tgggatctac	tatatatata	taatgcttta	caacacttgg	12540
atTTTTTTTT	ggaggctgga	atTTTTaatc	tacatatTTg	TTTTggccat	gcaccaactc	12600
attgtttagt	gtaatacttt	gattttgtca	aatatatgtg	ttcgtgtata	tttgataaag	12660
aatttctttg	accatataca	cacacacata	tatatatata	tatatatatt	atatatcatg	12720
cacttttaat	tgaaaaaata	atatatatat	atatagtgca	TTTTttctaa	caaccatata	12780
tgttgcgatt	gatctgcaaa	aatactgcta	gagtaatgaa	aaatataatc	tattgctgaa	12840
attatctcag	atgttaagat	tttcttaaag	taaattcttt	caaatttttag	ctaaaagtct	12900
tgttaataact	aaagaataat	acacaatctc	gaccacggaa	aaaaaacaca	taataaatTT	12960
gggcgcgccg	cgtattggct	agagcagctt	gccaacatgg	tggagcacga	cactctcgtc	13020
tactccaaga	atatcaaaga	tacagtctca	gaagaccaaa	gggctattga	gacttttcaa	13080
caaagggtaa	tatcgggaaa	cctcctcgga	ttccattgcc	cagctatctg	tcacttcac	13140
aaaaggacag	tagaaaagga	aggtggcacc	tacaaatgcc	atcattgcga	taaaggaaag	13200
gctatcggtc	aagatgcctc	tgccgacagt	ggtcccaaag	atggaccccc	accacagagg	13260

pctau2012001598-seq1

agcatcgtgg	aaaaagaaga	cgttccaacc	acgtcttcaa	agcaagtgga	ttgatgtgat	13320
aacatggtgg	agcacgacac	tctcgtctac	tccaagaata	tcaaagatac	agtctcagaa	13380
gaccaaaggg	ctattgagac	ttttcaacaa	agggtaatat	cgggaaacct	cctcggattc	13440
cattgcccag	ctatctgtca	cttcatcaaa	aggacagtag	aaaaggaagg	tggcacctac	13500
aaatgccatc	attgcgataa	aggaaaggct	atcgttcaag	atgcctctgc	cgacagtggg	13560
cccaaagatg	gacccccacc	cacgaggagc	atcggtgaaa	aagaagacgt	tccaaccacg	13620
tcttcaaagc	aagtggattg	atgtgatatc	tccactgacg	taagggatga	cgcacaatcc	13680
cactatcctt	cgcaagacct	tcctctatat	aaggaagttc	atttcatttg	gagaggacac	13740
gctgaaatca	ccagtctctc	tctacaaatc	tatctctgcg	atcgcatggc	gattttggat	13800
tctgctggcg	ttactacggt	gacggagaac	ggtggcggag	agttcgtcga	tcttgatagg	13860
cttcgtcgac	ggaaatcgag	atcggattct	tctaacggac	ttcttctctc	tggttccgat	13920
aataattctc	cttcggatga	tgttgagct	cccgccgacg	ttagggatcg	gattgattcc	13980
gttgtaacg	atgacgtca	gggaacagcc	aatttggccg	gagataataa	cggtggtggc	14040
gataataacg	gtggtggaag	aggcggcggg	gaaggaagag	gaaacgccga	tgctacgttt	14100
acgtatcgac	cgtcggttcc	agtcatcgg	agggcgagag	agagtccact	tagctccgac	14160
gcaatcttca	aacagagcca	tgccggatta	ttcaacctct	gtgtagtagt	tcttattgct	14220
gtaaacagta	gactcatcat	cgaaaatctt	atgaagtatg	gttggttgat	cagaacggat	14280
ttctggttta	gttcaagatc	gctgcgagat	tggccgcttt	tcatgtgttg	tatatccctt	14340
tcgatctttc	ctttggctgc	ctttacggtt	gagaaattgg	tacttcagaa	atacatatca	14400
gaacctgttg	tcatctttct	tcatattatt	atcaccatga	cagaggtttt	gtatccagtt	14460
tacgtcacc	taagggtgtg	ttctgctttt	ttatcaggtg	tcactttgat	gctcctcact	14520
tgcatgtgt	ggctaaagtt	ggtttcttat	gctcatacta	gctatgacat	aagatcccta	14580
gccaatgcag	ctgataaggc	caatcctgaa	gtctcctact	acgttagctt	gaagagcttg	14640
gcatatttca	tggtcgctcc	cacattgtgt	tatcagccaa	gttatccacg	ttctgcatgt	14700
atacggaagg	gttgggtggc	tcgtcaattt	gcaaaactgg	tcatattcac	cggattcatg	14760
ggatttataa	tagaacaata	tataaatcct	attgtcagga	actcaaagca	tcctttgaaa	14820
ggcgatcttc	tatatgctat	tgaaagagtg	ttgaagcttt	cagttccaaa	tttatatgtg	14880
tggctctgca	tgttctactg	cttcttcac	ctttggttaa	acatattggc	agagcttctc	14940
tgcttcgggg	atcgtaatt	ctacaaagat	tgggtggaatg	caaaaagtgt	gggagattac	15000
tggagaatgt	ggaatatgcc	tgttcataaa	tggatggttc	gacatatata	cttcccgtgc	15060
ttgcgcagca	agataccaaa	gacactcgcc	attatcattg	ctttcctagt	ctctgcagtc	15120
tttcatgagc	tatgcatcgc	agttccttgt	cgtctcttca	agctatgggc	ttttcttggg	15180
attatgtttc	aggtgccttt	ggtcttcatc	acaaactatc	tacaggaaag	gtttggctca	15240
acggtgggga	acatgatctt	ctggttcatc	ttctgcattt	tcggacaacc	gatgtgtgtg	15300

pctau2012001598-seq1

cttctttatt	accacgacct	gatgaaccga	aaaggatcga	tgtcatgagc	gatcgcgatc	15360
gttcaaacat	ttggcaataa	agtttcttaa	gattgaatcc	tgttgccggt	cttgcgatga	15420
ttatcatata	atttctgttg	aattacgtta	agcatgtaat	aattaacatg	taatgcatga	15480
cgttatattat	gagatggggt	tttatgatta	gagtcccgca	attatacatt	taatacgcga	15540
tagaaaacaa	aatatagcgc	gcaaactagg	ataaattatc	gcgcgcggtg	tcatctatgt	15600
tactagatcc	ctgcaggggc	tattggctag	agcagcttgc	caacatggtg	gagcacgaca	15660
ctctcgtcta	ctccaagaat	atcaaagata	cagtctcaga	agaccaaagg	gctattgaga	15720
cttttcaaca	aagggttaata	tcgggaaacc	tcctcggatt	ccattgcccc	gctatctgtc	15780
acttcatcaa	aaggacagta	gaaaaggaag	gtggcaccta	caaatgccat	cattgcgata	15840
aaggaaaggc	tatcgttcaa	gatgcctctg	ccgacagtgg	tcccaaagat	ggacccccac	15900
ccacgaggag	catcgtggaa	aaagaagacg	ttccaaccac	gtcttcaaag	caagtggatt	15960
gatgtgataa	catggtggag	cacgacactc	tcgtctactc	caagaatatc	aaagatacag	16020
tctcagaaga	ccaaagggtc	attgagactt	ttcaacaaag	ggtaatatcg	ggaaacctcc	16080
tcggattcca	ttgccagct	atctgtcact	tcatcaaaag	gacagtagaa	aaggaagggtg	16140
gcacctacaa	atgccatcat	tgcgataaag	gaaaggctat	cgttcaagat	gcctctgccg	16200
acagtgggtc	caaagatgga	ccccaccca	cgaggagcat	cgtggaaaaa	gaagacgttc	16260
caaccacgtc	ttcaaagcaa	gtggattgat	gtgatattctc	cactgacgta	agggatgacg	16320
cacaatccca	ctatccttcg	caagaccttc	ctctatataa	ggaagttcat	ttcatttgga	16380
gaggacacgc	tgaaatcacc	agtctctctc	tacaaatcta	tctctctcga	gatgattgaa	16440
caagatggat	tgcacgcagg	ttctccggcc	gcttgggtgg	agaggctatt	cggctatgac	16500
tgggcacaac	agacaatcgg	ctgctctgat	gccgccgtgt	tccggctgtc	agcgcagggg	16560
aggccggttc	tttttgtcaa	gaccgacctg	tccggtgccc	tgaatgaact	tcaagacgag	16620
gcagcgcggc	tatcgtggct	ggccacgacg	ggcgttcctt	gcgcagctgt	gctcgacgtt	16680
gtcactgaag	cgggaaggga	ctggctgcta	ttgggcgaag	tgccggggca	ggatctcctg	16740
tcatctcacc	ttgctcctgc	cgagaaagta	tccatcatgg	ctgatgcaat	gcggcggctg	16800
catacgcttg	atccggctac	ctgcccattc	gaccaccaag	cgaaacatcg	catcgagcga	16860
gcacgtactc	ggatggaagc	cggctcttgc	gatcaggatg	atctggacga	agagcatcag	16920
gggctcgcgc	cagccgaact	gttcgccagg	ctcaaggcgc	gcatgcccga	cggcgaggat	16980
ctcgtcgtga	ctcatggcga	tgcctgcttg	ccgaatatca	tgggtggaaaa	tggccgcttt	17040
tctggattca	tcgactgtgg	ccggctgggt	gtggcggacc	gctatcagga	catagcgttg	17100
gctacccgtg	atattgctga	agagcttggc	ggcgaatggg	ctgacc		17146

<210> 411
 <211> 11142
 <212> DNA
 <213> Artificial Sequence

pctau2012001598-seql

<220>

<223> TDNA sequence

<400> 411

tcctgtgggtt	ggcatgcaca	tacaaatgga	cgaacggata	aaccttttca	cgccctttta	60
aatatccgat	tattctaata	aacgctcttt	tctcttaggt	ttacccgcca	atatatcctg	120
tcaaacactg	atagttttaa	ctgaaggcgg	gaaacgacaa	tctgctagtg	gatctcccag	180
tcacgacgtt	gtaaaacggg	cgccctagaa	tctaattatt	ctattcagac	taaattagta	240
taagtatttt	tttaataaat	aaataataat	taataattta	ttagtaggag	tgattgaatt	300
tataatatat	tttttttaat	catttaaaga	atcttatatc	tttaaattga	caagagtttt	360
aaatggggag	agtgttatca	tatcacaagt	aggattaatg	tgttatagtt	tcacatgcat	420
tacgataagt	tgtgaaagat	aacattatta	tatataacaa	tgacaatcac	tagcgatcga	480
gtagtgagag	tcgtcttatt	acactttctt	ccttcgatct	gtcacatggc	ggcggcccga	540
attctcacac	aaggtagttg	caagacactg	aagtggtagt	agtggtagta	gaagaagcag	600
aatcggtaga	aaggcaagac	aatggagaag	atgaagatgg	tggagattct	cttcccacaa	660
cgcagcaatc	aaggttttca	aggttaaggc	actcgtgctt	tccatcatcg	aacatgaagt	720
cgatgttatc	ctcgaaagca	agctcgttga	agagttctgg	gtactcaatt	gggttctcgt	780
tagcaagggt	ttgatcggtg	aggaatgggg	agaatccagt	atccatcatg	cagaagttcc	840
aagcaagttc	gttggttatc	ccgcacctat	ccattttccat	gatgggtggaa	gaatcaatgc	900
agcagttaac	aacggcagct	tcctcagaat	atcccacaat	ttcagcctct	tgttgctcag	960
ccttctcttc	ctctttttct	tcttcctctt	gaggtgggtc	ctcaacgtat	tgttgcttaa	1020
cctcttccct	aggttcctct	ttagcttctc	tagtctcaac	ctcttgctta	gcctcaacaa	1080
gaataccctc	ttgatgggta	gcctgggttaa	ctgggaatgg	gaaaacgccc	ttcttcttaa	1140
gcctgtcgat	gtagttggag	atatacgaag	tggtaacagc	gtagcacct	ctgtactcaa	1200
tagcagccat	atcataagca	gctgcagcct	cttcttgagt	gttgtaagtt	ccgaggtaga	1260
ggtacttggt	tccgaaaact	cttccaatcc	tagcttccca	tcttccgtta	tgatgatgcc	1320
tagcaactcc	cctatactta	gaaactcccc	tagagaatcc	agatgactgc	cttctaaggg	1380
aagcaagata	ctcttctttg	gtcaccctct	gcatctcttc	aagttctttg	gtgtaagtct	1440
cagctgggaa	gttaagaatg	gtatctgggc	cccaatactt	aagagcagca	agatcatagg	1500
tatgagcagc	agcctcttca	gaatcataag	ctccaaggta	aacctgcttg	cccttcttgt	1560
tttggtatga	gttccaagag	gacttatccc	aaaggtgagc	ttcgaatctt	ccagtccatc	1620
tatgcctagt	aacacctctg	tagatagatg	accttctggg	agaagctgga	gaagttgggt	1680
tatgagactt	atcgccagat	ggagatgact	tcttagccct	cttagctctc	tttggctctg	1740
gagcttcaga	ttgaattggg	ctagaggtag	tagtagaaga	ggacactgaa	gaagatggag	1800
aactagagca	ggtagaggta	gtgagcctct	tcttcatgaa	ttctgttctt	ctttactctt	1860
tgtgtgactg	aggtttggtc	tagtgctttg	gtcatctata	tataatgata	acaacaatga	1920

pctau2012001598-seq1

gaacaagctt	tggagtgatc	ggaggggtcta	ggatacatga	gattcaagtg	gactaggatc	1980
tacaccgttg	gattttgagt	gtggatatgt	gtgaggttaa	ttttacttgg	taacggccac	2040
aaaggcctaa	ggagaggtgt	tgagaccctt	atcggcttga	accgctggaa	taatgccacg	2100
tggaagataa	ttccatgaat	cttatcgтта	tctatgagtg	aaattgtgtg	atggtggagt	2160
ggtgcttgct	cattttactt	gcctggtgga	cttggccctt	tccttatggg	gaatttatat	2220
tttacttact	atagagcttt	catacctttt	ttttaccttg	gatttagtta	atatataatg	2280
gtatgattca	tgaataaaaa	tgggaaattt	ttgaatttgt	actgctaaat	gcataagatt	2340
aggtgaaact	gtggaatata	tatttttttc	atttaaaagc	aaaatttgcc	ttttactaga	2400
attataaata	tagaaaaata	tataacattc	aaataaaaaat	gaaaataaga	actttcaaaa	2460
aacagaacta	tgtttaatgt	gtaaagatta	gtcgcacatc	aagtcattctg	ttacaatatg	2520
ttacaacaag	tcataagccc	aacaaagtta	gcacgtctaa	ataaactaaa	gagtccacga	2580
aatattaca	aatcataagc	ccaacaaagt	tattgatcaa	aaaaaaaaaa	cgcccaacaa	2640
agctaaacaa	agtccaaaaa	aaactttctc	agtctccatc	ttcctttatg	aacattgaaa	2700
actatacaca	aaacaagtca	gataaatctc	tttctggggc	tgtcttccca	acctcctaca	2760
tcacttccct	atcggattga	atgttttact	tgtacctttt	ccgttgcaat	gatattgata	2820
gtatgtttgt	gaaaactaat	agggttaaca	atcgaagtca	tggaatatgg	atttggtcca	2880
agattttccg	agagctttct	agtagaaagc	ccatcaccag	aaatttacta	gtaaaataaa	2940
tcaccaatta	ggtttcttat	tatgtgccaa	attcaatata	attatagagg	atatttcaaa	3000
tgaaaacgta	tgaatgttat	tagtaaattg	tcaggtaaga	cattaaaaaa	atcctacgtc	3060
agatattcaa	ctttaaaaaat	tcgatcagtg	tggaattgta	caaaaatttg	ggatctacta	3120
tatatatata	atgctttaca	acacttggaat	tttttttttg	aggctggaat	ttttaatcta	3180
catatttggt	ttggccatgc	accaactcat	tgtttagtgt	aatactttga	ttttgtcaaa	3240
tatatgtgtt	cgtgtatatt	tgtataagaa	tttctttgac	catatacaca	cacacatata	3300
tatatatata	tatatattat	atatcatgca	cttttaattg	aaaaaataat	atatatatat	3360
atagtgcatt	ttttctaaca	accatatatg	ttgcgattga	tctgcaaaaa	tactgctaga	3420
gtaatgaaaa	atataatcta	ttgctgaaat	tatctcagat	gttaagattt	tcttaaagta	3480
aattcctttca	aatttttagct	aaaagtcttg	taataactaa	agaataatac	acaatctcga	3540
ccacggaaaa	aaaacacata	ataaatttgg	ggcccctaga	atctaattat	tctattcaga	3600
ctaaattagt	ataagtattt	ttttaatcaa	taaataataa	ttaataattt	attagtagga	3660
gtgattgaat	ttataatata	ttttttttta	tcatttaaag	aatcttatat	ctttaaattg	3720
acaagagttt	taaatgggga	gagtgttatc	atatcacaag	taggattaat	gtgttatagt	3780
ttcacatgca	ttacgataag	ttgtgaaaga	taacattatt	atatataaca	atgacaatca	3840
ctagcgatcg	agtagtgaga	gtcgtcttat	tacactttct	tccttcgac	tgtcacatgg	3900
cggcggcccc	cggccgcttc	attactcgag	ccaggaggat	ggatcgatgc	tggctcgaga	3960

pctau2012001598-seq1

ccctgctacc	ggttgctgac	tgaactgctc	ggcacggtcc	ttcatttcac	gggccttgct	4020
cgccaacttt	gtcttggccg	actccaactg	atccgctccg	ggtggatggt	tccccgtcag	4080
gtaacggtag	atccaggaca	gcacagacag	agcggcaaca	ccaaatcccc	cgcttgccag	4140
aaaacccgct	cccaacagga	agatggtgat	gactgcagat	cagaaaaact	cagattaatc	4200
gacaaattcg	atcgcacaaa	ctagaaacta	acaccagatc	tagatagaaa	tcacaaatcg	4260
aagagtaatt	attcgacaaa	actcaaatta	tttgaacaaa	tcggatgata	tctatgaaac	4320
cctaatcgag	aattaagatg	atatctaacg	atcaaaccga	gaaaatcgtc	ttcgatctaa	4380
gattaacaga	atctaaacca	aagaacatat	acgaaattgg	gatcgaacga	aaacaaaatc	4440
gaagattttg	agagaataag	gaacacagaa	atttacctgc	agggaccagt	acaggcgaga	4500
agatcaccag	gagagggtg	gcgattgtca	gcgcaatgac	cgttccagcc	aggggtcaacc	4560
cggataacac	caacaggcta	cctccggcag	taaccgcggt	cgctgccttt	acaacacgct	4620
gagcacgcgg	ttgcagttgc	aagtgggggg	cacgtgtttg	ttgctgctgc	ccgtagtgct	4680
ctgccatggt	tttttttaac	ggagcaagcg	gccgctgttc	ttctttactc	tttgtgtgac	4740
tgaggtttgg	tctagtgcct	tggtcattca	tatataatga	taacaacaat	gagaacaagc	4800
tttgagtgta	tcggagggtc	taggatacat	gagattcaag	tggactagga	tctacaccgt	4860
tggattttga	gtgtggatat	gtgtgaggtt	aattttactt	ggtaacggcc	acaaggcct	4920
aaggagaggt	gttgagacct	ttatcggctt	gaaccgctgg	aataatgcca	cgtggaagat	4980
aattccatga	atcttatcgt	tatctatgag	tgaattgtg	tgatggtgga	gtggtgcttg	5040
ctcattttac	ttgcctgggt	gacttggccc	tttccttatg	gggaatttat	attttactta	5100
ctatagagct	ttcatacctt	ttttttacct	tggatttagt	taatataata	tggtatgatt	5160
catgaataaa	aatgggaaat	ttttgaattt	gtactgctaa	atgcataaga	ttaggtgaaa	5220
ctgtggaata	tatatTTTTT	tcatttaaaa	gcaaaatttg	ccttttacta	gaattataaa	5280
tatagaaaaa	tatataacat	tcaaataaaa	atgaaaataa	gaactttcaa	aaaacagaac	5340
tatgtttaat	gtgtaaagat	tagtcgcaca	tcaagtcatc	tgttacaata	tgttacaaca	5400
agtcataagc	ccaacaaagt	tagcacgtct	aaataaacta	aagagtccac	gaaaatatta	5460
caaatcataa	gccaacaaa	gttattgatc	aaaaaaaaaa	aacgccaac	aaagctaaac	5520
aaagtccaaa	aaaaacttct	caagtctcca	tcttccttta	tgaacattga	aaactataca	5580
caaaacaagt	cagataaatc	tctttctggg	cctgtcttcc	caacctccta	catcacttcc	5640
ctatcggatt	gaatgtttta	cttgtacctt	ttccgttgca	atgatattga	tagtatgttt	5700
gtgaaaacta	ataggggtta	caatcgaagt	catggaatat	ggatttggtc	caagattttc	5760
cgagagcttt	ctagtagaaa	gcccatcacc	agaaatttac	tagtaaaata	aatcaccaat	5820
taggtttctt	attatgtgcc	aaattcaata	taattataga	ggatatttca	aatgaaaacg	5880
tatgaatggt	attagtaaat	ggtcaggtaa	gacattaaaa	aaatcctacg	tcagatattc	5940
aactttaaaa	attcgatcag	tgtggaattg	tacaaaaatt	tgggatctac	tatatatata	6000

pctau2012001598-seq1

taatgcttta	caacacttgg	atTTTTTTTT	ggaggctgga	atTTTTaatc	tacatatattg	6060
ttttggccat	gcaccaactc	attgttttagt	gtaatacttt	gattttgtca	aatatatgtg	6120
ttcgtgtata	tttgataag	aatttctttg	accatataca	cacacacata	tatatatata	6180
tatatatatt	atatatcatg	cacttttaat	tgaaaaaata	atatatatat	atatagtgc	6240
ttttttctaa	caaccatata	tgttgcgatt	gatctgcaaa	aatactgcta	gagtaatgaa	6300
aaatataatc	tattgctgaa	attatctcag	atgttaagat	tttcttaaag	taaattcttt	6360
caaatttttag	ctaaaagtct	tgtaataact	aaagaataat	acacaatctc	gaccacggaa	6420
aaaaaacaca	taataaatTT	gggcgcgccg	cgtattggct	agagcagctt	gccaacatgg	6480
tggagcacga	cactctcgtc	tactccaaga	atatcaaaga	tacagtctca	gaagacccaa	6540
gggctattga	gacttttcaa	caaagggtaa	tatcgggaaa	cctcctcgga	ttccattgcc	6600
cagctatctg	tcacttcac	aaaaggacag	tagaaaagga	aggtggcacc	tacaaatgcc	6660
atcattgcga	taaaggaaag	gctatcgttc	aagatgcctc	tgccgacagt	ggccccaaag	6720
atggaccccc	accacgagg	agcatcgtgg	aaaaagaaga	cgttccaacc	acgtcttcaa	6780
agcaagtgga	ttgatgtgat	aacatgggtg	agcacgacac	tctcgtctac	tccaagaata	6840
tcaaagatac	agtctcagaa	gaccaaaggg	ctattgagac	ttttcaacaa	agggtaatat	6900
cgggaaacct	cctcggattc	cattgcccag	ctatctgtca	cttcatcaaa	aggacagtag	6960
aaaaggaagg	tggcacctac	aatgccatc	attgcgataa	aggaaaggct	atcgtttcaag	7020
atgcctctgc	cgacagtgg	cccaaagatg	gacccccacc	cacgaggagc	atcgtggaaa	7080
aagaagacgt	tccaaccacg	tcttcaaagc	aagtggattg	atgtgatatc	tccactgacg	7140
taagggatga	cgcacaatcc	cactatcctt	cgcaagacct	tcctctatat	aaggaagttc	7200
atttcatttg	gagaggacac	gctgaaatca	ccagtctctc	tctacaaatc	tatctctgcg	7260
atcgcattgg	gattttggat	tctgctggcg	ttactacgg	gacggagAAC	ggtggcggag	7320
agttcgtcga	tcttgatagg	cttcgtcgac	ggaaatcgag	atcggattct	tctaacggac	7380
ttcttctctc	tggttccgat	aataattctc	cttcggatga	tgttggagct	cccgccgacg	7440
ttagggatcg	gattgattcc	gttgtaacg	atgacgtc	gggaacagcc	aatttggccg	7500
gagataataa	cggtggtggc	gataataacg	gtgggtggaag	aggcggcgga	gaaggaagag	7560
gaaacgccga	tgctacgttt	acgtatcgac	cgtcggttcc	agctcatcgg	agggcgagag	7620
agagtccact	tagctccgac	gcaatcttca	aacagagcca	tgccggatta	ttcaacctct	7680
gtgtagtagt	tcttattgct	gtaaacagta	gactcatcat	cgaaaatctt	atgaagtatg	7740
gttggttgat	cagaacggat	ttctggttta	gttcaagatc	gctgcgagat	tggccgcttt	7800
tcatgtgttg	tatatccctt	tcgatctttc	ctttggctgc	ctttacggtt	gagaaattgg	7860
tacttcagaa	atacatatca	gaacctgttg	tcatctttct	tcatattatt	atcaccatga	7920
cagaggTTTT	gtatccagtt	tacgtcacc	taaggtgtga	ttctgctttt	ttatcagggtg	7980
tcactttgat	gctcctcact	tgcatgtgt	ggctaaagtt	ggtttcttat	gctcatacta	8040

pctau2012001598-seq1

gctatgacat	aagatcccta	gccaatgcag	ctgataaggc	caatcctgaa	gtctcctact	8100
acgttagctt	gaagagcttg	gcatatttca	tggtcgctcc	cacattgtgt	tatcagccaa	8160
gttatccacg	ttctgcatgt	atacgaagg	gttgggtggc	tcgtcaattt	gcaaaactgg	8220
tcatattcac	cggattcatg	ggatttataa	tagaacaata	tataaatcct	attgtcagga	8280
actcaaagca	tcctttgaaa	ggcgatcttc	tatatgctat	tgaaagagtg	ttgaagcttt	8340
cagttccaaa	tttatatgtg	tggtcttgca	tggttctactg	cttcttccac	ctttggttaa	8400
acatattggc	agagcttctc	tgcttcgggg	atcgtgaatt	ctacaaagat	tggtggaatg	8460
caaaaagtgt	gggagattac	tggagaatgt	ggaatatgcc	tgttcataaa	tggatggttc	8520
gacatatata	cttcccgtgc	ttgcgcagca	agataccaaa	gacactcgcc	attatcattg	8580
ctttcctagt	ctctgcagtc	tttcatgagc	tatgcatcgc	agttccttgt	cgtctcttca	8640
agctatgggc	ttttcttggg	attatgtttc	aggtgccttt	ggctcttcac	acaaactatc	8700
tacaggaaaag	gtttggctca	acggtgggga	acatgatctt	ctgggttcac	ttctgcattt	8760
tcggacaacc	gatgtgtgtg	cttctttatt	accacgacct	gatgaaccga	aaaggatcga	8820
tgtcatgagc	gatcgcgatc	gttcaaacat	ttggcaataa	agtttcttaa	gattgaatcc	8880
tggtgccggt	cttgcgatga	ttatcatata	atttctgttg	aattacgtta	agcatgtaat	8940
aattaacatg	taatgcatga	cgttatttat	gagatggggt	tttatgatta	gagtcccgca	9000
attatacatt	taatacgcg	tagaaaacaa	aatatagcgc	gcaaactagg	ataaattatc	9060
gcgcgcggtg	tcatctatgt	tactagatcc	ctgcagggcg	tattggctag	agcagcttgc	9120
caacatggtg	gagcacgaca	ctctcgtcta	ctccaagaat	atcaaagata	cagtctcaga	9180
agaccaaagg	gctattgaga	cttttcaaca	aagggttaata	tcgggaaacc	tcctcggatt	9240
ccattgccc	gctatctgtc	acttcatcaa	aaggacagta	gaaaaggaag	gtggcaccta	9300
caaatgccat	cattgcgata	aaggaaaggc	tatcgttcaa	gatgcctctg	ccgacagtgg	9360
tcccaaagat	ggacccccac	ccacgaggag	catcgtggaa	aaagaagacg	ttccaaccac	9420
gtcttcaaag	caagtggatt	gatgtgataa	catggtggag	cacgacactc	tcgtctactc	9480
caagaatatc	aaagatacag	tctcagaaga	caaaggggt	attgagactt	ttcaacaaag	9540
ggtaatatcg	ggaaacctcc	tcggattcca	ttgccagct	atctgtcact	tcatcaaaag	9600
gacagtagaa	aaggaagggtg	gcacctacaa	atgccatcat	tgcgataaag	gaaaggctat	9660
cgttcaagat	gcctctgccg	acagtgggtc	caaagatgga	ccccaccca	cgaggagcat	9720
cgtggaaaaa	gaagacgttc	caaccacgtc	ttcaaagcaa	gtggattgat	gtgatatctc	9780
cactgacgta	agggatgacg	cacaatccca	ctatccttcg	caagaccttc	ctctatataa	9840
ggaagtcat	ttcatttgg	gaggacacgc	tgaaatcacc	agtctctctc	tacaaatcta	9900
tctctctcga	gatgattgaa	caagatggat	tgacgcagg	ttctccggcc	gcttgggtgg	9960
agaggctatt	cggctatgac	tgggcacaac	agacaatcgg	ctgctctgat	gccgccgtgt	10020
tccggctgtc	agcgcagggg	aggccgggtc	tttttgtcaa	gaccgacctg	tccggtgccc	10080

pctau2012001598-seq1

tgaatgaact	tcaagacgag	gcagcgcggc	tatcgtggct	ggccacgacg	ggcgttcctt	10140
gcgcagctgt	gctcgacgtt	gtcactgaag	cgggaagggg	ctggctgcta	ttgggcgaag	10200
tgccgggggca	ggatctcctg	tcatctcacc	ttgctcctgc	cgagaaagta	tccatcatgg	10260
ctgatgcaat	gcggcggctg	catacgcttg	atccggctac	ctgcccattc	gaccaccaag	10320
cgaaacatcg	catcgagcga	gcacgtactc	ggatggaagc	cggctcttgc	gatcaggatg	10380
atctggacga	agagcatcag	gggctcgcgc	cagccgaact	gttcgccagg	ctcaaggcgc	10440
gcatgcccga	cggcgaggat	ctcgtcgtga	ctcatggcga	tgcctgcttg	ccgaatatca	10500
tggtggaaaa	tggccgcttt	tctggattca	tcgactgtgg	ccggctgggt	gtggcggacc	10560
gctatcagga	catagcgctt	gctaccctg	atattgctga	agagcttggc	ggcgaatggg	10620
ctgaccgctt	cctcgtgctt	tacggtatcg	ccgctcccga	ttcgcagcgc	atcgctttct	10680
atcgctttct	tgacgagttc	ttctgaaacg	cgtgatcggt	caaacatttg	gcaataaagt	10740
ttcttaagat	tgaatcctgt	tgccggctct	gcgatgatta	tcatataatt	tctgttgaat	10800
tacgttaagc	atgtaataat	taacatgtaa	tgcatgacgt	tatttatgag	atgggttttt	10860
atgattagag	tcccgcaatt	atacatttaa	tacgcgatag	aaaacaaaat	atagcgcgca	10920
aactaggata	aattatcgcg	cgcggtgtca	tctatgttac	tagatcgacg	tccgtacggt	10980
taaaaccacc	ccagtacatt	aaaaacgtcc	gcaatgtggt	attaagttgt	ctaagcgtca	11040
atttgtttac	accacaatat	atcctgccac	cagccagcca	acagctcccc	gaccggcagc	11100
tcggcacaaa	atcaccactc	gatacaggca	gcccatacgt	cc		11142

<210> 412
 <211> 13116
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> TDNA sequence

<400> 412	
tcctgtggtt	ggcatgcaca tacaaatgga cgaacggata aaccttttca cgccctttta 60
aatatccgat	tattctaata aacgctcttt tctcttaggt ttaccgcgca atatatcctg 120
tcaaacactg	atagttttaa ctgaaggcgg gaaacgacaa tctgctagtg gatctcccag 180
tcacgacgtt	gtaaaacggg cgtctgcgat cgctgaagtt cctatacttt tcagagaata 240
ggaacttcgg	aataggaact tcccatggga tctagtaaca tagatgacac cgcgcgcgat 300
aatttatcct	agtttgcgcg ctatatcttg ttttctatcg cgtattaaat gtataattgc 360
gggactctaa	tcataaaaac ccatctcata aataacgtca tgcattacat gtttaattatt 420
acatgcttaa	cgtaattcaa cagaaattat atgataatca tcgcaagacc ggcaacagga 480
ttcaatctta	agaaacttta ttgccaaatg tttgaacgat cacgctagcg gataacaatt 540
tcacacaggg	atatcactag taaaaggtag cgagctcctg cagtatcgat gcggccgcaa 600
agtcgacgaa	ttctcattag cagaactcaa gatgctgatc ctctggaacg ttgaacttga 660

pctau2012001598-seq1

gcttgtgttc	ctcgaaaagc	ttgcacaact	ctttgatgta	acgctgggtga	agtctatcaa	720
cttcctctct	agaaggctga	ggagtcattt	gaacctcgat	aggctttcca	acgatagtag	780
tgataggctg	tctgaaaggc	atgagtcgga	aagagtattg	gaaaactccc	cttccatgga	840
aaagtggaag	gctgattccc	ataatctttt	ggagtctgtt	ctggatccat	ctaagccaag	900
ttccaggagt	gttctcaacc	tggttgaaga	ggttgttctc	tccgaatgag	aagataggaa	960
caagagcagc	accatgcata	agagcaagtc	tgatgaatcc	cttacggttc	ttcaagagaa	1020
gtctgtaagc	accaggtcta	gcatcaagag	cctcttgagc	acctccaacg	atgatagcaa	1080
gaaggtttcc	accacccttt	ctgctaagga	tgtgatcagc	agaaactttc	tcgctagaca	1140
cgagtccacc	agacatgatg	taatctctga	agaatggagc	cctgaaccaa	acggtaaagca	1200
tcataaggta	ggatctgatt	ccagggaaca	aagaggtgaa	tccagtagac	tcagtacaga	1260
ggtaaggaa	agcaccagca	gcaagaacac	catgaggatg	gaatccagca	atgtagttag	1320
ggctaggatc	aagctcagca	gtcttaacga	gagacacagg	gaagtaatcc	ttcatgtact	1380
tccagatggc	caatcttctg	aagaattgga	taggtctacc	accttgtcta	ggcttatccc	1440
aatccaagta	ccaccaggta	gcgtaaagaa	cagagaaaag	ccagaacctg	gtgaacaaga	1500
gtccaacgaa	gataacgatg	cagagttgag	caagagcaag	gaatgagaaa	acctactgaa	1560
gaacagcgaa	agtctgcaat	cttctctccc	aaggaacaag	aagtggagcg	aactcgacca	1620
tgaattcagt	cccccggtgt	ctctccaaat	gaaatgaact	tccttatata	gaggaaggggt	1680
cttgcggaag	atagtgggat	tgtgcgtcat	cccttacgct	agtggagata	tcacatcaat	1740
ccacttgctt	tgaagacgtg	gttgaacgt	cttctttttc	cacgatgctc	ctcgtgggtg	1800
ggggtccatc	tttgggacca	ctgtcggcag	aggcatcttc	aacgatggcc	tttcctttat	1860
cgcaatgatg	gcattttag	gagccacctt	ccttttccac	tatcttcaca	ataaagttag	1920
agatagctgg	gcaatggaat	ccgaggaggt	ttccggatat	taccctttgt	tgaaaagtct	1980
caattgccct	ttggctctct	gagactgtat	ctttgatatt	tttggagtag	acaagtgtgt	2040
cgtgctccac	catgttgacg	aagattttct	tcttgtcatt	gagtcgtaag	agactctgta	2100
tgaactgttc	gccagtcttt	acggcgagtt	ctgttaggtc	ctctatttga	atctttgact	2160
ccatgggatc	caagggccct	agaatcta	tattctattc	agactaaatt	agtataagta	2220
ttttttta	caataaata	taattaata	tttattagta	ggagtgattg	aatttataat	2280
atattttttt	taatcattta	aagaatctta	tatcttttaa	ttgacaagag	ttttaaatgg	2340
ggagagtgtt	atcatatcac	aagtaggatt	aatgtgttat	agtttcacat	gcattacgat	2400
aagttgtgaa	agataacatt	attatatata	acaatgacaa	tcactagcga	tcgagtagtg	2460
agagtcgtct	tattacactt	tcttccttcg	atctgtcaca	tggcggcggc	ccgaattctc	2520
acacaaggta	gttgcaagac	actgaagtgg	tggtagtggg	agtagaagaa	gcagaatcgg	2580
tagaaaggca	agacaatgga	gaagatgaag	atggtggaga	ttctcttccc	acaacgcagc	2640
aatcaagggt	ttcaagggtta	aggcactcgt	gctttccatc	atcgaacatg	aagtcgatgt	2700

pctau2012001598-seq1

tatcctcgaa	agcaagctcg	ttgaagagtt	ctgggtactc	aattgggttc	tcgttagcaa	2760
ggttttgatc	ggtaaggaat	ggggagaatc	cagtatccat	catgcagaag	ttccaagcaa	2820
gttcgttggt	atctccgcac	ctatccattt	ccatgatggt	ggaagaatca	atgcagcagt	2880
taacaacggc	agcttctca	gaatatccca	caatttcagc	ctcttggtgc	tcagccttct	2940
cttcctcttt	ttcttcttcc	tcttgaggtg	gttcctcaac	gtattggtgc	ttaacctctt	3000
ccctagggtc	ctcttttagct	tctctagtct	caacctcttg	cttagcctca	acaagaatac	3060
cctcttgatg	gtagcctgg	ttaactggga	atgggaaaac	gcccttcttc	ttaagcctgt	3120
cgatgtagtt	ggagatatcg	aagttggtaa	cagcgtagc	acctctgtac	tcaatagcag	3180
ccatatcata	agcagctgca	gcctcttctt	gagtgttgta	agttccgagg	tagaggtact	3240
tgtttccgaa	aactcttcca	atcctagctt	cccatcttcc	gttatgatga	tgcttagcaa	3300
ctcccctata	cttagaaact	cccctagaga	atccagatga	ctgccttcta	agggaagcaa	3360
gatactcttc	tttggtcacc	ctctgcatct	cttcaagttc	tttgggtgta	gtctcagctg	3420
ggaagttaag	aatggtatct	gggcccgaat	acttaagagc	agcaagatca	taggtatgag	3480
cagcagcctc	ttcagaatca	taagctccaa	ggtaaacctg	cttgcccttc	ttgttttgga	3540
tggagttcca	agaggactta	tcccaaaggt	gagcttcgaa	tcttccagtc	catctatgcc	3600
tagtaacacc	tctgtagata	gatgaccttc	tggtagaagc	tggagaagtt	gggttatgag	3660
acttatcgcc	agatggagat	gacttcttag	ccctcttagc	tctctttggt	cttgagctt	3720
cagattgaat	tgggctagag	gtagtagtag	aagaggacac	tgaagaagat	ggagaactag	3780
agcaggtaga	ggtagtgagc	ctcttcttca	tgaattctgt	tcttctttac	tctttgtgtg	3840
actgaggttt	ggtctagtgc	tttggtcatc	tatatataat	gataacaaca	atgagaacaa	3900
gctttggagt	gatcggaggg	tctaggatac	atgagattca	agtggactag	gatctacacc	3960
gttggatttt	gagtgtggat	atgtgtgagg	ttaattttac	ttggtaacgg	ccacaaaggc	4020
ctaaggagag	gtgttgagac	ccttatcggc	ttgaaccgct	ggaataatgc	cacgtggaag	4080
ataattccat	gaatcttatc	gttatctatg	agtgaaattg	tgtgatggtg	gagtgggtgt	4140
tgctcatttt	acttgcctgg	tggacttggc	cctttcctta	tggggaattt	atattttact	4200
tactatagag	ctttcatacc	ttttttttac	cttggattta	gttaatatat	aatggtatga	4260
ttcatgaata	aaaatgggaa	atttttgaat	ttgtactgct	aaatgcataa	gattaggtga	4320
aactgtggaa	tatatatttt	tttcatttaa	aagcaaaatt	tgcccttttac	tagaattata	4380
aatatagaaa	aatatataac	attcaaataa	aaatgaaaat	aagaactttc	aaaaaacaga	4440
actatgttta	atgtgtaaag	attagtcgca	catcaagtca	tctgttacia	tatgttacia	4500
caagtcataa	gccaacaaa	gtagcacgt	ctaaataaac	taaagagtcc	acgaaaatat	4560
tacaaatcat	aagcccaaca	aagttattga	tcaaaaaaaa	aaaacgcca	acaaagctaa	4620
acaaagtcca	aaaaaaactt	ctcaagtctc	catcttcctt	tatgaacatt	gaaaactata	4680
cacaaaacaa	gtcagataaa	tctctttctg	ggcctgtctt	cccaacctcc	tacatcactt	4740

pctau2012001598-seq1

ccctatcgga	ttgaatgttt	tacttgtacc	ttttccgttg	caatgatatt	gatagtatgt	4800
ttgtgaaaac	taatagggtt	aacaatcgaa	gtcatggaat	atggatttgg	tccaagattt	4860
tccgagagct	ttctagtaga	aagcccatca	ccagaaattt	actagtaaaa	taaatcacca	4920
attaggtttc	ttattatgtg	ccaaattcaa	tataattata	gaggatattt	caaatgaaaa	4980
cgatatgaatg	ttattagtaa	atggtcaggt	aagacattaa	aaaaatccta	cgtcagatat	5040
tcaactttta	aaattcgatc	agtgtggaat	tgtacaaaaa	tttgggatct	actatatata	5100
tataatgctt	tacaacactt	ggattttttt	ttggaggctg	gaatttttta	tctacatatt	5160
tgttttggcc	atgcaccaac	tcattgttta	gtgtaatact	ttgattttgt	caaatatatg	5220
tgttcgtgta	tatttgtata	agaatttctt	tgaccatata	cacacacaca	tatatatata	5280
tatatatata	ttatatatca	tgcactttta	attgaaaaaa	taatatatat	atatatagtg	5340
cattttttct	aacaaccata	tatgttgcca	ttgatctgca	aaaatactgc	tagagtaatg	5400
aaaaatataa	tctattgctg	aaattatctc	agatgttaag	atcttcttaa	agtaaattct	5460
ttcaaatttt	agctaaaagt	cttgtaataa	ctaaagaata	atacacaatc	tcgaccacgg	5520
aaaaaaaaaca	cataataaat	ttggggcccc	tagaatctaa	ttattctatt	cagactaaat	5580
tagtataagt	atctttttta	tcaataaata	ataattaata	atctatttagt	aggagtgatt	5640
gaatttataa	tatatTTTTT	ttaatcattt	aaagaatctt	atatctttta	attgacaaga	5700
gttttaaatg	gggagagtgt	tatcatatca	caagtaggat	taatgtgtta	tagtttcaca	5760
tgcattacga	taagttgtga	aagataacat	tattatatat	aacaatgaca	atcactagcg	5820
atcgagtagt	gagagtcgtc	ttattacact	ttcttccttc	gatctgtcac	atggcggcgg	5880
cccgcggccg	cttcattact	cgagccagga	ggatggatcg	atgctggtct	gagaccctgc	5940
taccggttgc	tgactgaact	gctcggcacg	gtccttcatt	tcacgggcct	tgctcgccaa	6000
ctttgtcttg	gccgactcca	actgatccgc	tccgggtgga	tgtttccccg	tcaggtaacg	6060
gtagatccag	gacagcacag	acagagcggc	aacaccaa	ccccgccttg	ccagaaaacc	6120
cgctcccaac	aggaagatgg	tgatgactgc	agatcagaaa	aactcagatt	aatcgacaaa	6180
ttcgatcgca	caaactagaa	actaacacca	gatctagata	gaaatcaca	atcgaagagt	6240
aattattcga	caaaactcaa	attatttgaa	caaatcggat	gatattctatg	aaacccta	6300
cgagaattaa	gatgatattt	aacgatcaaa	cccagaaaat	cgtcttcgat	ctaagattaa	6360
cagaatctaa	accaaagaac	atatacga	ttgggatcga	acgaaaacaa	aatcgaagat	6420
tttgagagaa	taaggaacac	agaaatttac	ctgcagggac	cagtacaggc	gagaagatca	6480
ccaggagagg	tgtggcgatt	gtcagcgcaa	tgaccgttcc	agccagggtc	aaccgggata	6540
acaccaacag	gctacctccg	gcagtaaccg	cggtcgtgc	ctttacaaca	cgctgagcac	6600
gcggttgag	ttgcaagtgg	ggggcacgtg	tttgttgctg	ctgcccgtag	tgctctgcc	6660
tggttttttt	taacggagca	agcggccgct	gttcttcttt	actctttgtg	tgactgaggt	6720
ttggtctagt	gctttggtca	tctatatata	atgataacaa	caatgagaac	aagctttgga	6780

pctau2012001598-seq1

gtgatcggag	ggtctaggat	acatgagatt	caagtggact	aggatctaca	ccgttggatt	6840
ttgagtgtgg	atatgtgtga	ggttaatfff	acttggtaac	ggccacaaag	gcctaaggag	6900
aggtgttgag	acccttatcg	gcttgaaccg	ctggaataat	gccacgtgga	agataattcc	6960
atgaatctta	tcgttatcta	tgagtgaaat	tgtgtgatgg	tggagtgggtg	cttgctcatt	7020
ttacttgcct	ggtggacttg	gccctttcct	tatggggaat	ttatatfffa	cttactatag	7080
agctttcata	cctttttfff	accttggatt	tagttaatat	ataatgggtat	gattcatgaa	7140
taaaaatggg	aaatfffftga	atftgtactg	ctaaatgcat	aagattaggt	gaaactgtgg	7200
aatatatat	tttttcattt	aaaagcaaaa	tttgcctfff	actagaatta	taaatataga	7260
aaaatatata	acattcaa	aaaaatgaaa	ataagaactt	tcaaaaaaca	gaactatgft	7320
taatgtgtaa	agattagtcg	cacatcaagt	catctgttac	aatatgttac	aacaagtc	7380
aagcccaaca	aagttagcac	gtctaaataa	actaaagagt	ccacgaaaat	attacaaatc	7440
ataagcccaa	caaagttatt	gatcaaaaaa	aaaaaacgcc	caacaaagct	aaacaaagtc	7500
caaaaaaaac	ttctcaagtc	tccatcttcc	tttatgaaca	ttgaaaacta	tacacaaaac	7560
aagtcagata	aatctctttc	tgggcctgtc	ttcccaacct	cctacatcac	ttccctatcg	7620
gattgaatgt	tttacttght	ccttttccgt	tgcaatgata	ttgatagtat	gtttgtgaaa	7680
actaataggg	ttaacaatcg	aagtcatgga	atatggattt	ggtccaagat	tttccgagag	7740
ctttctagta	gaaagcccat	caccagaaat	ttactagtaa	aataaatcac	caattaggft	7800
tcttattatg	tgccaaattc	aatataatta	tagaggatat	ttcaaatgaa	aacgtatgaa	7860
tgttattagt	aaatggtcag	gtaagacatt	aaaaaatcc	tacgtcagat	attcaacttt	7920
aaaaattcga	tcagtgtgga	attgtacaaa	aatttgggat	ctactatata	tatataatgc	7980
tttacaacac	ttggatfff	ttttggaggc	tggaatfff	aatctacata	tttgttttgg	8040
ccatgcacca	actcattgft	tagtgtaata	ctttgatfff	gtcaaatata	tgtgttcgtg	8100
tatatttght	taagaatfff	tttgaccata	tacacacaca	catatatata	tatatatata	8160
tattatatat	catgcacttt	taattgaaaa	aataatatat	atatatatag	tgcattfff	8220
ctaacaacca	tatatgttgc	gattgatctg	caaaaatact	gctagagtaa	tgaaaaatat	8280
aatctattgc	tgaattatc	tcagatgtta	agatfffctt	aaagtaaatt	ctttcaaatt	8340
ttagctaaaa	gtcttghta	aactaaagaa	taatacacaa	tctcgaccac	ggaaaaaaa	8400
cacataataa	atttgggcgc	gccgcgtatt	ggctagagca	gcttgccaac	atggtggagc	8460
acgacactct	cgtctactcc	aagaatatca	aagatacagt	ctcagaagac	caaagggcta	8520
ttgagacttt	tcaacaaagg	gtaatatcgg	gaaacctcct	cggattccat	tgcccagcta	8580
tctgtcactt	catcaaaagg	acagtagaaa	aggaagggtg	cacctacaaa	tgccatcatt	8640
gcgataaagg	aaaggctatc	gttcaagatg	cctctgccga	cagtgggtccc	aaagatggac	8700
ccccaccac	gaggagcatc	gtggaaaaag	aagacgttcc	aaccacgtct	tcaaagcaag	8760
tggattgatg	tgataacatg	gtggagcacg	acactctcgt	ctactccaag	aatatcaaag	8820

pctau2012001598-seq1

atacagtctc	agaagaccaa	agggctattg	agacttttca	acaaagggta	atatcgggaa	8880
acctcctcgg	attccattgc	ccagctatct	gtcacttcat	caaaaggaca	gtagaaaagg	8940
aaggtggcac	ctacaaatgc	catcattgcg	ataaaggaaa	ggctatcggt	caagatgcct	9000
ctgccgacag	tgggtcccaa	gatggacccc	caccacgag	gagcatcggt	gaaaaagaag	9060
acgttccaac	cacgtcttca	aagcaagtgg	attgatgtga	tatctccact	gacgtaaggg	9120
atgacgcaca	atcccactat	ccttcgcaag	accttctct	atataaggaa	gttcattttca	9180
tttggagagg	acacgctgaa	atcaccagtc	tctctctaca	aatctatctc	tgcgatcgca	9240
tggcgatttt	ggattctgct	ggcgttacta	cggtgacgga	gaacgggtggc	ggagagttcg	9300
tcgatcttga	taggcttcgt	cgacggaaat	cgagatcgga	ttcttctaac	ggacttcttc	9360
tctctgggtc	cgataataat	tctccttcgg	atgatgttgg	agctcccgcc	gacgttaggg	9420
atcggattga	ttccgttggt	aacgatgacg	ctcagggaa	agccaatttg	gccggagata	9480
ataacggtgg	tggcgataat	aacggtggtg	gaagaggcgg	cggagaagga	agaggaaacg	9540
ccgatgctac	gtttacgtat	cgaccgtcgg	ttccagctca	tcggagggcg	agagagagtc	9600
cacttagctc	cgacgcaatc	ttcaaacaga	gccatgccgg	attattcaac	ctctgtgtag	9660
tagttcttat	tgctgtaa	agtagactca	tcacgaaaa	tcttatgaag	tatggttggt	9720
tgatcagaac	ggatttctgg	tttagttcaa	gatcgctgcg	agattggccg	cttttcatgt	9780
gttgatatatc	cttttcgac	tttcctttgg	ctgcctttac	ggttgagaaa	ttggtacttc	9840
agaaatacat	atcagaacct	gttgtcatct	ttcttcatat	tattatcacc	atgacagagg	9900
ttttgtatcc	agtttacgtc	accctaaggt	gtgattctgc	ttttttatca	ggtgtcactt	9960
tgatgctcct	cacttgcat	gtgtggctaa	agttggtttc	ttatgctcat	actagctatg	10020
acataagatc	cctagccaat	gcagctgata	aggccaatcc	tgaagtctcc	tactacgtta	10080
gcttgaagag	cttgcatat	ttcatggctg	ctccacatt	gtgttatcag	ccaagttatc	10140
cacgttctgc	atgtatacgg	aagggttggg	tggctcgta	atttgcaaaa	ctggctatat	10200
tcaccggatt	catgggattt	ataatagaac	aatatataaa	tcctattgtc	aggaactcaa	10260
agcatccttt	gaaaggcgat	cttctatatg	ctattgaaag	agtgttgaag	ctttcagttc	10320
caaatttata	tgtgtggctc	tgcattgtct	actgcttctt	ccacctttgg	ttaaacatat	10380
tggcagagct	tctctgcttc	ggggatcggt	aattctacaa	agattgggtg	aatgcaaaaa	10440
gtgtgggaga	ttactggaga	atgtggaata	tgctgttca	taaatggatg	gttcgacata	10500
tatacttccc	gtgcttgcg	agcaagatac	caaagacact	cgccattatc	attgctttcc	10560
tagtctctgc	agtctttcat	gagctatgca	tcgcagttcc	ttgtcgtctc	ttcaagctat	10620
gggcttttct	tgggattatg	tttcaggtgc	ctttggctct	catcacaac	tatctacagg	10680
aaaggtttgg	ctcaacggtg	gggaacatga	tcttctgggt	catcttctgc	attttcggac	10740
aaccgatgtg	tgtgcttctt	tattaccacg	acctgatgaa	ccgaaaagga	tcgatgtcat	10800
gagcgatcgc	gatcgttcaa	acatttggca	ataaagtttc	ttaagattga	atcctgttgc	10860

pctau2012001598-seq1

cggtcttgcg	atgattatca	tataatttct	gttgaattac	gttaagcatg	taataattaa	10920
catgtaatgc	atgacgttat	ttatgagatg	ggtttttatg	attagagtcc	cgcaattata	10980
catttaatac	gcgatagaaa	acaaaatata	gcgcgcaaac	taggataaat	tatcgcgcg	11040
ggtgtcatct	atgttactag	atccctgcag	ggcgtattgg	ctagagcagc	ttgccaacat	11100
ggtggagcac	gacactctcg	tctactccaa	gaatatcaaa	gatacagtct	cagaagacca	11160
aagggctatt	gagacttttc	aacaaagggt	aatatcgga	aacctcctcg	gattccattg	11220
cccagctatc	tgtcacttca	tcaaaaggac	agtagaaaag	gaagggtggca	cctacaaatg	11280
ccatcattgc	gataaaggaa	aggctatcgt	tcaagatgcc	tctgccgaca	gtgggtcccaa	11340
agatggaccc	ccaccacga	ggagcatcgt	ggaaaaagaa	gacgttccaa	ccacgtcttc	11400
aaagcaagtg	gattgatgtg	ataacatggt	ggagcacgac	actctcgtct	actccaagaa	11460
tatcaaagat	acagtctcag	aagaccaaag	ggctattgag	acttttcaac	aaagggtaat	11520
atcgggaaac	ctcctcggat	tccattgccc	agctatctgt	cacttcatca	aaaggacagt	11580
agaaaaggaa	ggtggcacct	acaaatgcca	tcattgcgat	aaaggaaagg	ctatcgttca	11640
agatgcctct	gccgacagtg	gtcccaaaga	tggaccccca	cccacgagga	gcatcgtgga	11700
aaaagaagac	gttccaacca	cgtcttcaaa	gcaagtggat	tgatgtgata	tctccactga	11760
cgtaagggat	gacgcacaat	cccactatcc	ttcgcaagac	cttcctctat	ataaggaagt	11820
tcatttcatt	tggagaggac	acgctgaaat	caccagtctc	tctctacaaa	tctatctctc	11880
tcgagatgat	tgaacaagat	ggattgcacg	caggttctcc	ggccgcttgg	gtggagaggc	11940
tattcggcta	tgactgggca	caacagacaa	tcggctgctc	tgatgccgcc	gtgttccggc	12000
tgtcagcgca	ggggaggccg	gttctttttg	tcaagaccga	cctgtccggt	gccctgaatg	12060
aacttcaaga	cgaggcagcg	cggctatcgt	ggctggccac	gacgggctgt	ccttgccgag	12120
ctgtgctcga	cgttgtcact	gaagcgggaa	gggactggct	gctattgggc	gaagtgccgg	12180
ggcaggatct	cctgtcatct	caccttgctc	ctgccgagaa	agtatccatc	atggctgatg	12240
caatgcggcg	gctgcatacg	cttgatccgg	ctacctgccc	attcgaccac	caagcgaaac	12300
atcgcatcga	gcgagcacgt	actcggatgg	aagccggtct	tgtcgatcag	gatgatctgg	12360
acgaagagca	tcaggggctc	gcgccagccg	aactgttcgc	caggctcaag	gcgcgcatgc	12420
ccgacggcga	ggatctcgtc	gtgactcatg	gcgatgcctg	cttgccgaat	atcatggtgg	12480
aaaatggccg	cttttctgga	ttcatcgact	gtggccggct	gggtgtggcg	gaccgctatc	12540
aggacatagc	gttggctacc	cgtgatattg	ctgaagagct	tggcggcgaa	tgggctgacc	12600
gcttcctcgt	gctttacggg	atcgccgctc	ccgattcgca	gcgcacgccc	ttctatcgcc	12660
ttcttgacga	gttcttctga	aacgcgtgat	cgttcaaaca	tttggaata	aagtttctta	12720
agattgaatc	ctgttgccgg	tcttgcatg	attatcatat	aatttctgtt	gaattacgtt	12780
aagcatgtaa	taattaacat	gtaatgcatg	acgttattta	tgagatgggt	ttttatgatt	12840
agagtcccg	aattatacat	ttaatacgcg	atagaaaaca	aaatatagcg	cgcaaactag	12900

pctau2012001598-seq1

gataaattat cgcgcgcggt gtcattctatg ttactagatc gacgtccgta cggttaaaac	12960
caccccagta cattaaaaac gtccgcaatg tgttattaag ttgtctaagc gtcaatttgt	13020
ttacaccaca atatatcctg ccaccagcca gccaacagct ccccgaccgg cagctcggca	13080
caaaatcacc actcgataca ggcagcccat cagtcc	13116

<210> 413
 <211> 16749
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> vector sequence

<400> 413	
gcttcctcgt gctttacggt atcgccgctc ccgattcgca gcgcatcgcc ttctatcgcc	60
ttcttgacga gttcttctga aacgcgtgat cgttcaaaca tttggcaata aagtttctta	120
agattgaatc ctgttgccgg tcttgcatg attatcatat aatttctgtt gaattacgtt	180
aagcatgtaa taattaacat gtaatgcatg acgttattta tgagatgggt ttttatgatt	240
agagtcccg c aattatacat ttaatacgcg atagaaaaca aaatatagcg cgcaaactag	300
gataaattat cgcgcgcggt gtcattctatg ttactagatc gacgtccgta cggttaaaac	360
caccccagta cattaaaaac gtccgcaatg tgttattaag ttgtctaagc gtcaatttgt	420
ttacaccaca atatatcctg ccaccagcca gccaacagct ccccgaccgg cagctcggca	480
caaaatcacc actcgataca ggcagcccat cagtccacta gacgtccacc gggctggttg	540
ccctcgccgc tgggctggcg gccgtctatg gccctgcaaa cgcgccagaa acgccgtcga	600
agccgtgtgc gagacaccgc agccgccggc gttgtggata cctcgcgga aacttggccc	660
tcactgacag atgagggggc gacgttgaca cttgaggggc cgactcacc ggcgcgcggt	720
tgacagatga ggggcaggct cgatttcggc cggcgacgtg gagctggcca gcctcgcaaa	780
tcggcgaaaa cgctgattt tacgcgagtt tcccacagat gatgtggaca agcctgggga	840
taagtgccct gcggtattga cacttgaggg gcgcgactac tgacagatga ggggcgcgat	900
ccttgacact tgaggggcag agtgctgaca gatgaggggc gcacctattg acatttgagg	960
ggctgtccac aggcagaaaa tccagcattt gcaaggggtt ccgcccgttt ttcggccacc	1020
gctaacctgt cttttaacct gcttttaaac caatatattat aaaccttggt ttttaaccagg	1080
gctgcgccct gtgcgcgtga ccgcgcacgc cgaagggggg tgccccccct tctcgaaccc	1140
tccggggccg ctctcgcggtt ggcagcatca ccataattg tggtttcaaa atcggctccg	1200
tcgatactat gttatacgcc aactttgaaa acaactttga aaaagctgtt ttctggtatt	1260
taaggtttta gaatgcaagg aacagtgaat tggagttcgt cttgttataa ttagcttctt	1320
ggggtattta aatactgtag aaaagaggaa ggaaataata aatggctaaa atgagaatat	1380
caccggaatt gaaaaaactg atcgaaaaat accgctgcgt aaaagatacg gaaggaatgt	1440
ctcctgctaa ggtatataag ctggtgggag aaaatgaaaa cctatatatta aaaatgacgg	1500

pctau2012001598-seq1

acagccggta	taaagggacc	acctatgatg	tggaacggga	aaaggacatg	atgctatggc	1560
tggaaggaaa	gctgcctgtt	ccaaaggtcc	tgcaccttga	acggcatgat	ggctggagca	1620
atctgctcat	gagtgaggcc	gatggcgtcc	tttgctcgga	agagtatgaa	gatgaacaaa	1680
gccctgaaaa	gattatcgag	ctgtatgcgg	agtgcacag	gctctttcac	tccatcgaca	1740
tatcggattg	tccctatacg	aatagcttag	acagccgctt	agccgaattg	gattacttac	1800
tgaataacga	tctggccgat	gtggattgcg	aaaactggga	agaagacacc	ccatttaaag	1860
atccgcgcga	gctgtatgat	tttttaaaga	cggaaaagcc	cgaagaggaa	cttgtctttt	1920
cccacggcga	cctgggagac	agcaacatct	ttgtgaaaga	tggcaaagta	agtggcttta	1980
ttgatcttgg	gagaagcggc	agggcggaca	agtggatga	cattgccttc	tgcgtccggt	2040
cgatcagggg	ggatattggg	gaagaacagt	atgtcgagct	atTTTTTgac	ttactgggga	2100
tcaagcctga	ttgggagaaa	ataaaatatt	atattttact	ggatgaattg	ttttagtacc	2160
tagatgtggc	gcaacgatgc	tggcgacaag	caggagcgca	ccgacttctt	ccgcatcaag	2220
tgTTTTggct	ctcaggccga	ggcccacggc	aagtatttgg	gcaaggggtc	gctggtattc	2280
gtgcagggca	agattcggaa	taccaagtac	gagaaggacg	gccagacggt	ctacgggacc	2340
gacttcattg	ccgataaggt	ggattatctg	gacaccaagg	caccaggcgg	atcaaatcag	2400
gaataagggc	acattgcccc	ggcgtgagtc	ggggcaatcc	cgcaaggagg	gtgaatgaat	2460
cggacgtttg	accggaaggc	atacaggcaa	gaactgatcg	acgcgggggt	ttccgccgag	2520
gatgccgaaa	ccatcgcaag	ccgcaccgtc	atgcgtgcgc	cccgcgaaac	cttccagtcc	2580
gtcggctcga	tggcccagca	agctacggcc	aagatcgagc	gcgacagcgt	gcaactggct	2640
ccccctgccc	tgcccgcgcc	atcggccgcc	gtggagcggt	cgcgtcgtct	cgaacaggag	2700
gcggcaggtt	tggcgaagtc	gatgaccatc	gacacgcgag	gaactatgac	gaccaagaag	2760
cgaaaaaccg	ccggcgagga	cctggcaaaa	caggtcagcg	aggccaagca	agccgcgttg	2820
ctgaaacaca	cgaagcagca	gatcaaggaa	atgcagcttt	ccttgttcga	tattgcgccg	2880
tggccggaca	cgatgcgagc	gatgccaaac	gacacggccc	gctctgccct	gttcaccacg	2940
cgcaacaaga	aaatcccgcg	cgaggcgctg	caaaacaagg	tcattttcca	cgtaacaag	3000
gacgtgaaga	tcacctacac	cggcgctcag	ctgcgggccg	acgatgacga	actggtgtgg	3060
cagcaggtgt	tggagtacgc	gaagcgcacc	cctatcggcg	agccgatcac	cttcacgttc	3120
tacgagcttt	gccaggacct	gggctggtcg	atcaatggcc	ggtattacac	gaaggccgag	3180
gaatgcctgt	cgcgcttaca	ggcgacggcg	atgggcttca	cgtccgaccg	cgttgggcac	3240
ctggaatcgg	tgctgctgct	gcaccgcttc	cgcgtcctgg	accgtggcaa	gaaaacgtcc	3300
cgttgccagg	tcctgatcga	cgaggaaatc	gtcgtgctgt	ttgctggcga	ccactacacg	3360
aaattcatat	gggagaagta	ccgcaagctg	tcgccgacgg	cccgacggat	gttcgactat	3420
ttcagctcgc	accgggagcc	gtacccgctc	aagctggaaa	ccttccgcct	catgtgcgga	3480
tcggattcca	cccgcgtgaa	gaagtggcgc	gagcaggtcg	gcgaagcctg	cgaagagttg	3540

pctau2012001598-seq1

cgaggcagcg	gcctggtgga	acacgcctgg	gtcaatgatg	acctggtgca	ttgcaaacgc	3600
tagggccttg	tggggtcagt	tccggctggg	ggttcagcag	ccagcgcttt	actgagatcc	3660
tcttccgctt	cctcgctcac	tgactcgctg	cgctcggtcg	ttcggctgcg	gcgagcggta	3720
tcagctcact	caaaggcggg	aatacggtta	tccacagaat	caggggataa	cgcaggaaag	3780
aacatgtgag	caaaaggcca	gcaaaaggcc	aggaaccgta	aaaaggccgc	gttgctggcg	3840
tttttccata	ggctccgccc	ccctgacgag	catcacaaaa	atcgacgctc	aagtcagagg	3900
tggcgaaacc	cgacaggact	ataaagatac	caggcgtttc	cccctggaag	ctccctcgtg	3960
cgctctcctg	ttccgaccct	gccgcttacc	ggatacctgt	ccgcctttct	cccttcggga	4020
agcgtggcgc	tttctcatag	ctcacgctgt	aggtatctca	gttcgggtgta	ggtcgttcgc	4080
tccaagctgg	gctgtgtgca	cgaaccccc	gttcagcccg	accgctgcmc	cttatccggt	4140
aactatcgtc	ttgagtccaa	cccggtaaga	cacgacttat	cgccactggc	agcagccact	4200
ggtaacagga	ttagcagagc	gaggtatgta	ggcggtgcta	cagagtctct	gaagtgggtg	4260
cctaactacg	gctacactag	aagaacagta	tttggatatct	gcgctctgct	gaagccagtt	4320
accttcggaa	aaagagttgg	tagctcttga	tccggcaaac	aaaccaccgc	tggtagcggg	4380
ggtttttttg	tttgcaagca	gcagattacg	cgcagaaaaa	aaggatctca	agaagatcct	4440
ttgatctttt	ctacggggtc	tgacgctcag	tggaacgaaa	actcacgtta	agggattttg	4500
gtcatgagat	tatcaaaaag	gatcttcacc	tagatccttt	tggatctcct	gtggttggca	4560
tgcacataca	aatggacgaa	cggataaacc	ttttcacgcc	cttttaata	tccgattatt	4620
ctaataaacg	ctcttttctc	ttaggtttac	ccgccaatat	atcctgtcaa	acactgatag	4680
tttaactga	aggcgggaaa	cgacaatctg	ctagtggatc	tcccagtcac	gacgttgtaa	4740
aacgggcgtc	tgcgatcgct	gaagttccta	tacttttcag	agaataggaa	cttcggaata	4800
ggaacttccc	atgggatcta	gtaacataga	tgacaccgcg	cgcgataatt	tatcctagtt	4860
tgcgcgctat	attttgtttt	ctatcgcgta	ttaaattgtat	aattgcggga	ctctaatacat	4920
aaaaacccat	ctcataaata	acgtcatgca	ttacatgtta	attattacat	gcttaacgta	4980
attcaacaga	aattatatga	taatcatcgc	aagaccggca	acaggattca	atcttaagaa	5040
actttattgc	caaagtgttg	aacgatcacg	ctagcggata	acaatttcac	acagggatat	5100
cactagtaaa	aggtaccgag	ctcctgcagt	atcgatgcgg	ccgcaaagtc	gacgaattct	5160
cattagcaga	actcaagatg	ctgatcctct	ggaacgttga	acttgagctt	gtgttcctcg	5220
aaaagcttgc	acaactcttt	gatgtaacgc	tggtgaagtc	tatcaacttc	ctctctagaa	5280
ggctgaggag	tcatttgaac	ctcgataggc	tttccaacga	tagtagtgat	aggctgtctg	5340
aaaggcatga	gtccgaaaga	gtattggaaa	actccccttc	catggaaaag	tggaaggctg	5400
attcccataa	tcttttgtag	tctgttctgg	atccatctaa	gccaagtctc	aggagtgttc	5460
tcaacctggg	tgaagagggt	gttctctccg	aatgagaaga	taggaacaag	agcagcacca	5520
tgcataagag	caagtctgat	gaatccctta	cggttcttca	agagaagtct	gtaagcacca	5580

pctau2012001598-seq1

ggtctagcat	caagagcctc	ttgagcacct	ccaacgatga	tagcaagaag	gtttccacca	5640
ccctttctgc	taaggatgtg	atcagcagaa	actttctcgc	tagacacgag	tccaccagac	5700
atgatgtaat	ctctgaagaa	tggagccctg	aaccaaacgg	taagcatcat	aaggtaggat	5760
ctgattccag	ggaacaaaga	ggtgaatcca	gtagactcag	tacagagggt	aaggaaagca	5820
ccagcagcaa	gaacaccatg	aggatggaat	ccagcaatgt	agttacggct	aggatcaagc	5880
tcagcagtct	taacgagaga	cacagggaag	taatccttca	tgtacttcca	gatggccaat	5940
cttctgaaga	attggatagg	tctaccacct	tgtctaggct	tatcccaatc	caagtaccac	6000
caggtagcgt	aaagaacaga	gaaaagccag	aacctggtga	acaagagtcc	aacgaagata	6060
acgatgcaga	gttgagcaag	agcaaggaat	gagaaaaccc	actgaagaac	agcgaaagtc	6120
tgcaatcttc	tctcccaagg	aacaagaagt	ggagcgaact	cgaccatgaa	ttcagtcctc	6180
cgtgttctct	ccaaatgaaa	tgaacttcct	tatatagagg	aagggtcttg	cgaaggatag	6240
tgggattgtg	cgatcatcct	tacgtcagtg	gagatatcac	atcaatccac	ttgctttgaa	6300
gacgtggttg	gaacgtcttc	ttttccacg	atgctcctcg	tgggtggggg	tccatctttg	6360
ggaccactgt	cggcagaggc	atcttcaacg	atggcctttc	ctttatcgca	atgatggcat	6420
ttgtaggagc	caccttcctt	ttccactatc	ttcacaataa	agtgacagat	agctgggcaa	6480
tggaatccga	ggaggtttcc	ggatattacc	ctttgttgaa	aagtctcaat	tgccctttgg	6540
tcttctgaga	ctgtatcttt	gatatttttg	gagtagacaa	gtgtgtcgtg	ctccaccatg	6600
ttgacgaaga	ttttcttctt	gtcattgagt	cgtaagagac	tctgtatgaa	ctgttcgcca	6660
gtctttacgg	cgagttctgt	taggtcctct	atgtgaatct	ttgactccat	gggatccaag	6720
ggccctagaa	tctaattatt	ctattcagac	taaattagta	taagtatttt	tttaatcaat	6780
aaataataat	taataattta	ttagtaggag	tgattgaatt	tataatatat	tttttttaat	6840
catttaaaga	atcttatatc	tttaaatgga	caagagtttt	aaatggggag	agtgttatca	6900
tatcacaagt	aggattaatg	tgttatagtt	tcacatgcat	tacgataagt	tgtgaaagat	6960
aacattatta	tatataacaa	tgacaatcac	tagcgatcga	gtagtgagag	tcgtcttatt	7020
acactttctt	ccttcgatct	gtcacatggc	ggcggcccga	attctcacac	aaggtagttg	7080
caagacactg	aagtgggtgg	agtggtagta	gaagaagcag	aatcggtaga	aaggcaagac	7140
aatggagaag	atgaagatgg	tggagattct	cttcccacaa	cgcagcaatc	aaggttttca	7200
aggttaaggc	actcgtgctt	tccatcatcg	aacatgaagt	cgatgttatc	ctcgaaagca	7260
agctcgttga	agagttctgg	gtactcaatt	gggttctcgt	tagcaagggt	ttgatcggtg	7320
aggaatgggg	agaatccagt	atccatcatg	cagaagttcc	aagcaagttc	gttggttatct	7380
ccgcacctat	ccatttccat	gatggtggaa	gaatcaatgc	agcagttaac	aacggcagct	7440
tcctcagaat	atcccacaat	ttcagcctct	tggtgctcag	ccttctcttc	ctctttttct	7500
tcttctcttt	gaggtgggtc	ctcaacgtat	tggtgcttaa	cctcttccct	aggttctctt	7560
ttagcttctc	tagtctcaac	ctcttgctta	gcctcaacaa	gaataccctc	ttgatgggtta	7620

pctau2012001598-seq1

gcctgggttaa	ctgggaatgg	gaaaacgccc	ttctttcttaa	gcctgtcgat	gtagttggag	7680
atatcgaagt	tggtaacagc	gtagcacct	ctgtactcaa	tagcagccat	atcataagca	7740
gctgcagcct	cttcttgagt	gttgtaagtt	ccgaggtaga	ggtacttggt	tccgaaaact	7800
cttccaatcc	tagcttccca	tcttccgtta	tgatgatgcc	tagcaactcc	cctatactta	7860
gaaactcccc	tagagaatcc	agatgactgc	cttctaaggg	aagcaagata	ctcttctttg	7920
gtcacccctct	gcatctcttc	aagttctttg	gtgtaagtct	cagctgggaa	gttaagaatg	7980
gtatctgggc	ccaataactt	aagagcagca	agatcatagg	tatgagcagc	agcctcttca	8040
gaatcataag	ctccaaggta	aacctgcttg	cccttcttgt	tttgatgga	gttccaagag	8100
gacttatccc	aaaggtgagc	ttcgaatctt	ccagtccttc	tatgcctagt	aacacctctg	8160
tagatagatg	accttctggg	agaagctgga	gaagttgggt	tatgagactt	atcgccagat	8220
ggagatgact	tcttagccct	cttagctctc	tttgggtctg	gagcttcaga	ttgaattggg	8280
ctagaggtag	tagtagaaga	ggacactgaa	gaagatggag	aactagagca	ggtagaggta	8340
gtgagcctct	tcttcatgaa	ttcactagt	attaaatttt	gttgggtgctt	tgagcatata	8400
acaagcatgg	tatatatagg	cacgtaaaca	agttgagaaa	ttttactttg	agtttgacat	8460
aaccaataaa	agttagtgt	gtttattacc	tcactcagtt	tgcaccgcaa	ctgtcggttag	8520
tgatgtttac	ctttcctttt	tctattattt	attagtatta	tataatatat	atatatgtgt	8580
gatgagactt	gaaattgttt	agcaccgcaa	atgtccttct	tgaggggagg	ttttcttttg	8640
ctgaggttgg	ggtgtcacat	acacccccct	ctatggactc	aacgtccttg	ctgaggttta	8700
ccccacacta	catgagattt	ttctagactc	aatactatga	tatttctcgc	cttatcgga	8760
ttgggttaa	tcagttgaag	ttaggggtcat	atcgataaaa	ttgacacatg	atcgactctg	8820
atattaaaca	gattctctcc	ctcgaacctc	actcactttc	ctttttctat	tctttattag	8880
tattatataa	tagatccgtt	ccaaccattc	acgtacataa	gaagagagat	atTTTTTTTT	8940
aatggactaa	catgacaaat	aaaacaaaca	aaggagtaat	gatcactaca	acaaattaga	9000
ttatgagga	caaataattt	catcatctat	aatcatgtt	tcgtcactaa	aaattttgtg	9060
tgacgaaaaa	gatttcgtca	atcagttgtc	actaaaaata	tacaaagacg	atttaatgat	9120
gtttaccttt	ccttttctat	tctttattag	tattatataa	taaatatatg	tgtgatgaga	9180
cttgaaattg	tttagcaccg	caaatgtcct	tggtgaaggg	aggttttctt	ttgctgaggt	9240
tggggtgtca	catacaccct	ctctatggac	tgaacgtcct	ttttgaggtt	tattttacac	9300
tgcatgagat	ttttctagat	tcaacattat	gatttctaga	ctcaacacta	cgatcgtcac	9360
taaagactat	tttttatata	taaaaaaaat	actttgtcct	taaatgtata	aattagggat	9420
aaatttatta	ttataaaaaa	ggttaataat	tttgtgatta	aatctattat	tttgtcactg	9480
aaagtgtttg	cttttaccga	cgacatatat	gtcactaaat	attatcataa	gtagtgacaa	9540
ttacaattgt	cacaaaataa	aaaaaattat	tcatattcaa	caaaaaaggg	tactacgaca	9600
atacatTTTT	tgtcactgaa	agtaatcaag	ttgtgataaa	ttaatTTatt	taatgacaaa	9660

pctau2012001598-seq1

aatattttgta	tcaaaattca	cccatgatca	tataataaaa	ataactaaaa	ttatactaaa	9720
gcataaatga	caagaaaatc	taactaaaac	atatcaaata	ttactcctaa	acaagacat	9780
ataagtaaaa	atttcttcca	aagtatcaat	aacgtggtga	cacatagctt	gcaatcaatc	9840
ttgcttcaat	tttcaccttt	tatacctgta	aaaagaaaga	gaaaataaaa	caatgattta	9900
aaaatcgaat	tcccgcggcc	cctagaatct	aattatttcta	ttcagactaa	attagtataa	9960
gtattttttt	aatcaataaa	taataattaa	taattttatta	gtaggagtga	ttgaattttat	10020
aatatatattt	ttttaatcat	ttaaagaatc	ttatatcttt	aaattgacaa	gagtttttaa	10080
tgggggagagt	gttatcatat	cacaagtagg	attaatgtgt	tatagtttca	catgcattac	10140
gataagttgt	gaaagataac	attattatat	ataacaatga	caatcactag	cgatcgagta	10200
gtgagagtcg	tcttattaca	ctttcttcct	tcgatctgtc	acatggcggc	ggcccgcggc	10260
cgcttcatta	ctcgagccag	gaggatggat	cgatgctggt	ctgagaccct	gctaccggtt	10320
gctgactgaa	ctgctcggca	cggctccttca	tttcacgggc	cttgctcgcc	aactttgtct	10380
tggccgactc	caactgatcc	gctccgggtg	gatgtttccc	cgtcaggtaa	cggtagatcc	10440
aggacagcac	agacagagcg	gcaacaccaa	atccccgct	tgccagaaaa	cccgtccca	10500
acaggaagat	ggtgatgact	gcagatcaga	aaaactcaga	ttaatcgaca	aattcgatcg	10560
cacaaactag	aaactaacac	cagatctaga	tagaaatcac	aaatcgaaga	gtaattattc	10620
gacaaaactc	aaattatttg	aacaaatcgg	atgatattcta	tgaacccta	atcgagaatt	10680
aagatgatat	ctaacgatca	aaccagaaa	atcgtcttcg	atctaagatt	aacagaatct	10740
aaaccaaaga	acatatacga	aattgggatc	gaacgaaaac	aaaatcgaag	attttgagag	10800
aataaggaac	acagaaattt	acctgcaggg	accagtacag	gcgagaagat	caccaggaga	10860
ggtgtggcga	ttgtcagcgc	aatgaccgtt	ccagccaggg	tcaaccgga	taacaccaac	10920
aggctacctc	cggcagtaac	cgcggtcgct	gcctttacaa	cacgctgagc	acgcggttgc	10980
agttgcaagt	ggggggcacg	tgtttgttgc	tgctgcccg	agtgtctctgc	catggaaatt	11040
ttgttggtgc	tttgagcata	taacaagcat	ggtatatata	ggcacgtaaa	caagttgaga	11100
aattttactt	tgagtttgac	ataaccaata	aaagttagt	ctgtttatta	cctcactcag	11160
tttgaccgc	aactgtcggt	agtgatgttt	acctttcctt	tttctattat	ttattagtat	11220
tatataatat	atatatatgt	gtgatgagac	ttgaaattgt	ttagcaccgc	aaatgtcctt	11280
cttgagggga	ggttttcttt	tgctgagggt	ggggtgtcac	atacaccccc	ctctatggac	11340
tcaacgtcct	tgctgagggt	tacccacac	tacatgagat	ttttctagac	tcaatactat	11400
gatattttctc	gccttatcgg	aattgggtta	actcagttga	agttagggtc	atatcgataa	11460
aattgacaca	tgatcgactc	tgatattaaa	cagattctct	ccctcgaacc	tcactcactt	11520
tcctttttct	attctttatt	agtattatat	aatagatccg	ttccaaccat	tcacgtacat	11580
aagaagagag	atattttttt	ttaatggact	aacatgacaa	ataaaacaaa	caaaggagta	11640
atgatcacta	caacaaatta	gattatgagg	gacaaataat	ttcatcatct	ataaatcatg	11700

pctau2012001598-seq1

tttcgtcact	aaaaattttg	tgtgacgaaa	aagatttcgt	caatcagttg	tcactaaaaa	11760
tatacaaaga	cgatttaaatg	atgtttacct	ttccttttct	attcttttatt	agtattatat	11820
aataaatata	tgtgtgatga	gacttgaaat	tgtttagcac	cgcaaagtgc	cttgttgaag	11880
ggagggtttc	ttttgctgag	gttgggggtgt	cacatacacc	ccctctatgg	actgaacgtc	11940
ctttttgagg	tttattttac	actgcatgag	atttttctag	attcaacatt	atgattttcta	12000
gactcaacac	tacgatcgtc	actaaagact	attttttata	tataaaaaaa	atactttgtc	12060
cttaaagtga	taaattaggg	ataaatattat	tattataaaa	aagggttaata	attttgtgat	12120
taaatctatt	attttgtcac	tgaaagtgtt	tgcttttacc	gacgacatat	atgtcactaa	12180
atattatcat	aagtagtgac	aattacaatt	gtcacaaaat	aaaaaaaaatt	attcatattc	12240
aacaaaaaag	ggtactacga	caatacattt	tttgtcactg	aaagtaatca	agttgtgata	12300
aattaattta	tttaattgaca	aaaatatattg	tatcaaaatt	cacccatgat	catataataa	12360
aaataactaa	aattatacta	aagcataaat	gacaagaaaa	tctaactaaa	acatatcaaa	12420
tattactcct	aaacaaagac	atataagtaa	aaattttcttc	caaagtatca	ataacgtggt	12480
gacacatagc	ttgcaatcaa	tcttgcttca	attttcacct	tttataacctg	taaaaagaaa	12540
gagaaaataa	aacaatgatt	taaaggcgcg	ccgcgtattg	gctagagcag	cttgccaaca	12600
tggtggagca	cgacactctc	gtctactcca	agaatatcaa	agatacagtc	tcagaagacc	12660
aaagggctat	tgagactttt	caacaaaggg	taatatcggg	aaacctcctc	ggattccatt	12720
gcccagctat	ctgtcacttc	atcaaaagga	cagtagaaaa	ggaaggtggc	acctacaaat	12780
gccatcattg	cgataaagga	aaggctatcg	ttcaagatgc	ctctgccgac	agtgggccca	12840
aagatggacc	cccaccacg	aggagcatcg	tggaaaaaga	agacgttcca	accacgtctt	12900
caaagcaagt	ggattgatgt	gataacatgg	tggagcacga	cactctcgtc	tactccaaga	12960
atatcaaaga	tacagtctca	gaagaccaa	gggctattga	gacttttcaa	caaagggtaa	13020
tatcgggaaa	cctcctcgga	ttccattgcc	cagctatctg	tcacttcatc	aaaaggacag	13080
tagaaaagga	aggtggcacc	tacaaatgcc	atcattgcga	taaaggaaaag	gctatcgttc	13140
aagatgcctc	tgccgacagt	ggtcccaaag	atggaccccc	accacgagg	agcatcgtag	13200
aaaaagaaga	cgttccaacc	acgtcttcaa	agcaagtgga	ttgatgtgat	atctccactg	13260
acgtaaggga	tgacgcacaa	tcccactatc	cttcgcaaga	ccttcctcta	tataaggaag	13320
ttcatttcat	ttggagagga	cacgctgaaa	tcaccagtct	ctctctacaa	atctatctct	13380
gcgatcgcat	ggcgattttg	gattctgctg	gcgttactac	ggtgacggag	aacggtggcg	13440
gagagttcgt	cgatcttgat	aggcttcgtc	gacggaaatc	gagatcggat	tcttctaacg	13500
gacttcttct	ctctggttcc	gataataatt	ctccttcgga	tgatgttgga	gctcccgccg	13560
acgttaggga	tcggattgat	tccgttggtta	acgatgacgc	tcagggaaca	gccaatattg	13620
ccggagataa	taacggtggg	ggcgataata	acggtggtgg	aagaggcggc	ggagaaggaa	13680
gaggaaacgc	cgatgctacg	tttacgtatc	gaccgtcggt	tccagctcat	cggagggcga	13740

pctau2012001598-seq1

gagagagtcc	acttagctcc	gacgcaatct	tcaaacagag	ccatgccgga	ttattcaacc	13800
tctgtgtagt	agttcttatt	gctgtaaaca	gtagactcat	catcgaaaat	cttatgaagt	13860
atgggttggt	gatcagaacg	gatttctggt	ttagttcaag	atcgctgcga	gattggccgc	13920
ttttcatgtg	ttgtatatcc	ctttcgatct	ttcctttggc	tgccctttacg	gttgagaaat	13980
tggtaacttca	gaaatacata	tcagaacctg	ttgtcatctt	tcttcatatt	attatcacca	14040
tgacagaggt	tttgtatcca	gtttacgtca	ccctaagggtg	tgattctgct	tttttatcag	14100
gtgtcacttt	gatgctcctc	acttgcattg	tgtggctaaa	gttgggtttct	tatgctcata	14160
ctagctatga	cataagatcc	ctagccaatg	cagctgataa	ggccaatcct	gaagtctcct	14220
actacgttag	cttgaagagc	ttggcatatt	tcatggctgc	tcccacattg	tgttatcagc	14280
caagttatcc	acgttctgca	tgtatacgga	agggttgggt	ggctcgtcaa	tttgcaaaac	14340
tggatcatatt	caccggattc	atgggattta	taatagaaca	atatataaat	cctattgtca	14400
ggaactcaaa	gcatcctttg	aaaggcgatc	ttctatatgc	tattgaaaga	gtgttgaagc	14460
tttcagttcc	aaatttatat	gtgtggctct	gcatgttcta	ctgcttcttc	cacctttggt	14520
taaacatatt	ggcagagctt	ctctgcttcg	gggatcgtga	attctacaaa	gattggtgga	14580
atgcaaaaag	tgtgggagat	tactggagaa	tgtggaatat	gcctgttcat	aaatggatgg	14640
ttcgacatat	atacttcccg	tgcttgcgca	gcaagatacc	aaagacactc	gccattatca	14700
ttgctttcct	agtctctgca	gtctttcatg	agctatgcac	cgcagttcct	tgctgtctct	14760
tcaagctatg	ggcttttctt	gggattatgt	ttcaggtgcc	tttgggtcttc	atcacaaact	14820
atctacagga	aaggtttggc	tcaacggtgg	ggaacatgat	cttctgggtc	atcttctgca	14880
ttttcggaca	accgatgtgt	gtgcttcttt	attaccacga	cctgatgaac	cgaaaaggat	14940
cgatgtcatg	agcgatcgcg	atcgttcaaa	catttggcaa	taaagtttct	taagattgaa	15000
tcctgttgcc	ggtcttgcca	tgattatcat	ataatttctg	ttgaattacg	ttaagcatgt	15060
aataattaac	atgtaatgca	tgacgttatt	tatgagatgg	gtttttatga	ttagagtccc	15120
gcaattatac	atttaatacg	cgatagaaaa	caaaatatag	cgcgcaaact	aggataaatt	15180
atcgcgcgcg	gtgtcatcta	tgttactaga	tccttgcagg	gcgtattggc	tagagcagct	15240
tgccaacatg	gtggagcacg	acactctcgt	ctactccaag	aatatcaaag	atacagtctc	15300
agaagaccaa	agggtctattg	agacttttca	acaaagggtg	atatcgggaa	acctcctcgg	15360
attccattgc	ccagctatct	gtcacttcat	caaaaggaca	gtagaaaagg	aagggtggcac	15420
ctacaaatgc	catcattgcg	ataaaggaaa	ggctatcggt	caagatgcct	ctgccgacag	15480
tgggtcccaa	gatggacccc	caccacgag	gagcatcgtg	gaaaaagaag	acgttccaac	15540
cacgtcttca	aagcaagtgg	attgatgtga	taacatgggtg	gagcacgaca	ctctcgtcta	15600
ctccaagaat	atcaaagata	cagtctcaga	agaccaaagg	gctattgaga	cttttcaaca	15660
aagggttaata	tcgggaaacc	tcctcggatt	ccattgccca	gctatctgtc	acttcatcaa	15720
aaggacagta	gaaaaggaag	gtggcaccta	caaatgcat	cattgcgata	aaggaaaggc	15780

pctau2012001598-seq1

tatcgttcaa gatgcctctg cgcacagtgg tcccaaagat ggacccccac ccacgaggag	15840
catcgtggaa aaagaagacg ttccaaccac gtcttcaaag caagtggatt gatgtgatat	15900
ctccactgac gtaagggatg acgcacaatc ccactatcct tcgcaagacc ttcctctata	15960
taaggaagtt catttcattt ggagaggaca cgctgaaatc accagtctct ctctacaaat	16020
ctatctctct cgagatgatt gaacaagatg gattgcacgc aggttctccg gccgcttggg	16080
tggagaggct attcggctat gactgggcac aacagacaat cggctgctct gatgccgccg	16140
tgttccggct gtcagcgcag gggaggcccg ttctttttgt caagaccgac ctgtccggtg	16200
ccctgaatga acttcaagac gaggcagcgc ggctatcgtg gctggccacg acgggcgttc	16260
cttgccgacg tgtgctcgac gttgtcactg aagcgggaag ggactggctg ctattgggag	16320
aagtgccggg gcaggatctc ctgtcatctc accttgctcc tgccgagaaa gtatccatca	16380
tggctgatgc aatgcggcgg ctgcatacgc ttgatccggc tacctgcca ttcgaccacc	16440
aagcgaaaca tcgcatcgag cgagcacgta ctcggatgga agccggtctt gtcgatcagg	16500
atgatctgga cgaagagcat caggggctcg cgccagccga actgttcgcc aggctcaagg	16560
cgcgcatgcc cgacggcgag gatctcgtcg tgactcatgg cgatgcctgc ttgccgaata	16620
tcatggtgga aaatggccgc ttttctggat tcatcgactg tggccggctg ggtgtggcgg	16680
accgctatca ggacatagcg ttggctaccc gtgatattgc tgaagagctt ggcggcgaat	16740
gggctgacc	16749

<210> 414
 <211> 137
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> linker sequence

<400> 414	
atttaaagtc ggccgcaat tcgtcgattg aggacgtccc tactagacct gctggacctc	60
ctcctgctac ttactacgat tctctcgctg tgcataatgg cagtcatgcc cgggcctgca	120
ggcggccgca tttaa	137

<210> 415
 <211> 9734
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> synthesized DNA fragment

<400> 415	
cctgcagggc gatcgcttga aacatccctg aagtgtctca tttattttta tttattcttt	60
gctgataaaa aaataaaata aaagaagcta agcacacggt caaccattgc tctactgcta	120
aaagggttat gtgtagtgtt ttactgcata aattatgcag caaacaagac aactcaaatt	180
aaaaaatttc ctttgcttgt tttttgttg tctctgactt gactttcttg tggaagttgg	240

pctau2012001598-seq1

ttgtataagg	attgggacac	cattgtcctt	cttaatttaa	ttttattctt	tgctgataaa	300
aaaaaaaaatt	tcatatagtg	ttaaataata	atttgttaaa	taaccaaaaa	gtcaaatatg	360
tttactctcg	tttaaataat	tgagattcgt	ccagcaaggc	taaacgattg	tatagattta	420
tgacaatatt	tactttttta	tagataaatg	ttatattata	ataaatttat	atacatatat	480
tatatgttat	ttattattat	tttaaadcct	tcaatatttt	atcaaaccac	ctcataattt	540
tttttttatc	tgtaagaagc	aataaaatta	aatagaccca	ctttaaggat	gatccaacct	600
ttatacagag	taagagagtt	caaatagtac	cctttcatat	acatatcaac	taaaatatta	660
gaaatatcat	ggatcaaacc	ttataaagac	attaaataag	tggataagta	taatataata	720
atgggtagta	tataatatat	aaatggatac	aaacttctct	ctttataatt	gttatgtctc	780
cttaacatcc	taatataata	cataagtggg	taatataata	tatataaatg	gagacaaact	840
tcttccatta	taattgttat	gtcttcttaa	cacttatgtc	tcgttcacaa	tgctaagggt	900
agaattgttt	agaaagtctt	atagtacaca	tttgtttttg	tactatttga	agcattccat	960
aagccgtcac	gattcagatg	atttataata	ataagaggaa	atttatcata	gaacaataag	1020
gtgcatagat	agagtgttaa	tatatcataa	catcctttgt	ttattcatag	aagaagttag	1080
atggagctca	gttattatac	tgttacatgg	tcggatacaa	tattccatgc	tctccatgag	1140
ctcttacacc	tacatgcatt	ttagttcata	cttcatatgt	cactcgacaa	actcaagctc	1200
cttgaccctg	tccttggcgt	agatgtcctt	gtacctgtca	tagaggctct	gcacagcgtg	1260
catgtaagag	tcatggagct	ccctgatgat	ctcatccttg	gtctgaccgt	acttgataga	1320
tgggactggg	attggcttac	ccacaatggg	gtaaattggg	tgctgtgtg	ggaggagacc	1380
aaagtcgtag	ttgaacacac	cccttccatg	gaagagtggc	atgggtgaaac	caagagcatg	1440
ctggatcttc	ttctgccacc	tgtgaagctt	agagttctca	ttgctctcaa	tctgctcata	1500
aagctcattc	tcaccaaagg	agatggtagg	gacaaggcta	gcaccagtct	gcacagcaat	1560
cttgatgaaa	ccaaacctcc	tcttcaacac	aagggtccatc	actcctggtc	tggcgttcaa	1620
ggactcagag	gcacctccca	ccacaatagc	aatggacctt	cctggaccac	ccttggtgag	1680
gatgttctgg	caggaggctt	tagacactga	gcagagtcca	caggccatga	ggtaatccct	1740
gtagagtggg	atgttgaagt	tgctggtgag	ggtgagcaag	gaaggcctga	ttcctgggaa	1800
gaggtcatcg	aaaccagtag	cgttggtgga	gaaggtgcag	aagctaccca	tgctgatgat	1860
accatgtggg	tggtaaccaa	agatgtagtt	cttgatggg	tcaaggtcag	cctccttgat	1920
aacatgagct	gggaagtatc	cagcaaagag	cttccaccaa	gctgcgttcc	tgagccacct	1980
gataggtctg	cctccgttct	caggggcctt	atcccacacc	aaaatccagg	tcaagtagag	2040
gataattggg	aaccaaagga	gcacagggat	gctgcagagg	aagaagaaga	tgaacatgca	2100
gatgctcatc	atggagcacc	agaggagaac	agcaagggtt	tggagccttc	tcctaagtgg	2160
gaccctgagt	ggagcatacc	taggggatgg	gatggcctca	aggggtgtgct	tcaccttctg	2220
ttggagggtgc	tggtccttgc	tagccatatg	gtaattgtaa	atgtaattgt	aatgttgttt	2280

pctau2012001598-seq1

gttgtttggt	gttgttggta	attgttgtaa	aaagtatatc	ttaaattctt	taatacgggtg	2340
tactaggata	ttgaactggt	tcttgatgat	gaaaacctgg	gccgagattg	cagctattta	2400
tagtcatagg	tcttgtaaac	atgcatggac	atttggccac	ggggtggcat	gcagtttgac	2460
gggtgttgaa	ataaacaaaa	atgaggtggc	ggaagagaat	acgagtttga	ggttgggtta	2520
gaaacaacaa	atgtgagggc	tcatgatggg	ttgagttggt	gaatgttttg	ggctgctcga	2580
ttgacacctt	tgtgagtacg	tgttgttgtg	catggctttt	ggggtccagt	tttttttct	2640
tgacgcggcg	atcctgatca	gctagtggat	aagtgatgtc	cactgtgtgt	gattgcgttt	2700
ttgtttgaat	tttatgaact	tagacattgc	tatgcaaagg	atactctcat	tgtgttttgt	2760
cttcttttgt	tccttggcct	tttcttatga	tccaagagac	tagtcagtgt	tgtggcattc	2820
gagactacca	agattaatta	tgatggggga	aggataagta	actgattagt	acggactggt	2880
accaaattaa	ttaataagcg	gcaaataag	ggcatggatg	cgatcgcgcg	gccgctcctc	2940
gaagagaagg	gttaataaca	cattttttta	catttttaac	acaaatttta	gttattttaa	3000
aatttattaa	aaaatttaaa	ataagaagag	gaactcttta	aataaatcta	acttacaaaa	3060
tttatgattt	ttaataagtt	ttcaccaata	aaaaatgtca	taaaaatatg	ttaaaaagta	3120
tattatcaat	attctcttta	tgataaataa	aaagaaaaaa	aaaataaaaag	ttaagtgaaa	3180
atgagattga	agtgacttta	ggtgtgtata	aatatatcaa	ccccgccaac	aatttattta	3240
atccaaatat	attgaagtat	attattccat	agcctttatt	tatttatata	tttattatat	3300
aaaagcttta	tttgttctag	gttgttcatg	aaatatTTTT	ttggttttat	ctccgttgta	3360
agaaaatcat	gtgctttgtg	tcgccactca	ctattgcagc	tttttcatgc	attggtcaga	3420
ttgacggttg	attgtatttt	tgttttttat	ggttttgtgt	tatgacttaa	gtcttcatct	3480
ctttatctct	tcatcagggt	tgatggttac	ctaatatggg	ccatgggtac	atgcatgggt	3540
aaattaggtg	gccaactttg	ttgtgaacga	tagaattttt	tttatattaa	gtaaactatt	3600
tttatattat	gaaataataa	taaaaaaaaa	attttatcat	tattaacaaa	atcatattag	3660
ttaatttggt	aactctataa	taaaagaaat	actgtaacat	tcacattaca	tggtaacatc	3720
ttccaccct	ttcatttggt	ttttgtttga	tgactttttt	tcttgtttaa	atttatttcc	3780
cttcttttaa	atttggaata	cattatcatc	atatataaac	taaaatacta	aaaacaggat	3840
tacacaaatg	ataaataata	acacaaatat	ttataaatct	agctgcaata	tattttaaact	3900
agctatatcg	atattgtaaa	ataaaactag	ctgcattgat	actgataaaa	aaatatcatg	3960
tgctttctgg	actgatgatg	cagtatactt	ttgacattgc	ctttatttta	tttttcagaa	4020
aagctttctt	agttctgggt	tcttcattat	ttgtttccca	tctccattgt	gaattgaatc	4080
atttgcttcg	gtcacaaaat	acaatttagc	taggtacatg	cattggtcag	attcacgggt	4140
tattatgtca	tgacttaagt	tcatggtagt	acattacctg	ccacgcatgc	attatatagg	4200
ttagatttga	taggcaaatt	tggttgtcaa	caatataaat	ataaataatg	tttttatatt	4260
acgaaataac	agtgatcaaa	acaaacagtt	ttatctttat	taacaagatt	ttgtttttgt	4320

pctau2012001598-seq1

ttgatgacgt	tttttaatgt	ttacgctttc	ccccttcttt	tgaattttaga	acactttatc	4380
atcataaaat	caaataactaa	aaaaattaca	tatttcataa	ataataacac	aaatattttt	4440
aaaaaatctg	aaataataat	gaacaatatt	acatattatc	acgaaaattc	attaataaaa	4500
atattatata	aataaaatgt	aatagtagtt	atatgtagga	aaaaagtact	gcacgcataa	4560
tatatacaaa	aagattaaaa	tgaactatta	taaataataa	cactaaatta	atggtgaatc	4620
atatcaaaat	aatgaaaaag	taaataaaat	ttgtaattaa	cttctatatg	tattacacac	4680
acaaataata	aataatagta	aaaaaaatta	tgataaatat	ttaccatctc	ataagatatt	4740
taaaataatg	ataaaaatat	agattatttt	ttatgcaact	agctagccaa	aaagagaaca	4800
cgggtatata	taaaaagagt	acctttaaat	tctactgtac	ttccttttatt	cctgacgttt	4860
ttatatcaag	tggacatacg	tgaagatfff	aattatcagt	ctaaatattt	cattagcact	4920
taatactfff	ctgtttttatt	cctatcctat	aagtagtccc	gattctccca	acattgctta	4980
ttcacacaac	taactaagaa	agtcttccat	agcccccaa	gaattctfff	acaacaatta	5040
ccaacaacaa	caaacaacaa	acaacattac	aattacattt	acaattacca	aatgaagaag	5100
aggctcacta	cctctacctg	ctctagttct	ccatcttctt	cagtgtcctc	ttctactact	5160
acctctagcc	caattcaatc	tgaagctcca	agaccaaaga	gagctaagag	ggctaagaag	5220
tcatctccat	ctggcgataa	gtctcataac	ccaacttctc	cagcttctac	cagaaggcca	5280
tctatctaca	gagggtgttac	taggcataga	tggactggaa	gattcgaagc	tcacctttgg	5340
gataagtcct	cttggaaactc	catccaaaac	aagaagggca	agcagggttta	ccttggagct	5400
tatgattctg	aagaggctgc	tgctcatacc	tatgatcttg	ctgctcttaa	gtattggggc	5460
ccagatacca	ttcttaactt	cccagctgag	acttacacca	aagaacttga	agagatgcag	5520
agggtgacca	aagaagagta	tcttgcttcc	cttagaaggc	agtcactctg	attctctagg	5580
ggagtttcta	agtatagggg	agttgctagg	catcatcata	acggaagatg	ggaagctagg	5640
attggaagag	ttttcggaag	caagtacctc	tacctcgga	cttacaacac	tcaagaagag	5700
gctgcagctg	cttatgatat	ggctgctatt	gagtacagag	gtgctaacgc	tgttaccaac	5760
ttcgatatct	ccaactacat	cgacaggctt	aagaagaagg	gcgttttccc	attcccagtt	5820
aaccaggcta	accatcaaga	gggtattctt	gttgaggcta	agcaagaggt	tgagactaga	5880
gaagctaaag	aggaacctag	ggaagagggt	aagcaacaat	acgttgagga	accacctcaa	5940
gaggaagaag	aaaaagagga	agagaaggct	gagcaacaag	aggctgaaat	tgtgggatat	6000
tctgaggaag	ctgccgttgt	taactgctgc	attgattctt	ccaccatcat	ggaaatggat	6060
aggtgcggag	ataacaacga	acttgcttgg	aacttctgca	tgatggatac	tggattctcc	6120
ccattcctta	ccgatcaaaa	ccttgctaac	gagaacccaa	ttgagtaccc	agaactcttc	6180
aacgagcttg	ctttcgagga	taacatcgac	ttcatgttcg	atgatggaaa	gcacgagtcg	6240
cttaaccttg	aaaaccttga	ttgctgcgtt	gtgggaagag	aatctccacc	atcttcatct	6300
tctccattgt	cttgcccttc	taccgattct	gcttcttcta	ctaccactac	caccacttca	6360

pctau2012001598-seq1

gtgtcttgca	actaccttgt	gtgagaattc	gacacaagtg	tgagagtact	aaataaatgc	6420
tttggttgta	cgaaatcatt	acactaaata	aaataatcaa	agcttatata	tgccctccgc	6480
taaggccgaa	tgcaaagaaa	ttggttcttt	ctcgttatct	tttgccactt	ttactagtag	6540
gtattaatta	ctacttaatc	atctttgttt	acggctcatt	atatccggtc	tagaggatcc	6600
aaggccgcga	agttaaaagc	aatgtttgtca	cttgctgtac	taacacatga	tgtgatagtt	6660
tatgctagct	agctataaca	taagctgtct	ctgagtgtgt	tgtatatata	taaagatcat	6720
cactggtgaa	tggtgatcgt	gtacgtaccc	tacttagtag	gcaatggaag	cacttagagt	6780
gtgctttgtg	catggccttg	cctctgtttt	gagacttttg	taatgttttc	gagtttaaat	6840
ctttgccttt	gcgcggccgc	atttaaataca	tccatgccct	tcatttgccg	cttattaatt	6900
aatttggtaa	cagtccgtac	taatcagtta	cttatccttc	ccccatcata	attaatcttg	6960
gtagtctcga	atgccacaac	actgactagt	ctcttggtac	ataagaaaaa	gccaaagAAC	7020
aaaagaagac	aaaacacaat	gagagtatcc	tttgcatagc	aatgtctaag	ttcataaaat	7080
tcaaacaaaa	acgcaatcac	acacagtgga	catcacttat	ccactagctg	atcaggatcg	7140
ccgcgtcaag	aaaaaaaaac	tggaccccaa	aagccatgca	caacaacacg	tactcacaaa	7200
ggtgtcaatc	gagcagccca	aaacattcac	caactcaacc	catcatgagc	cctcacatct	7260
gttgtttcta	acccaacctc	aaactcgtat	tctcttccgc	cacctcatct	ttgtttatct	7320
caacacccgt	caaactgcat	gccaccccg	ggccaaatgt	ccatgcatgt	taacaagacc	7380
tatgactata	aatagctgca	atctcggccc	aggttttcat	catcaagaac	cagttcaata	7440
tcctagtaca	ccgtattaaa	gaatttaaga	tatactctcg	agttttacaa	caattaccaa	7500
caacaacaaa	caacaacaaa	cattacaatt	acattttaca	ttaccaaagt	gtcgagttcg	7560
ctccacttct	tgttccttgg	gagagaagat	tgacagactt	cgctgttctt	cagtgggttt	7620
tctcattcct	tgctcttgct	caactctgca	tcgttatctt	cgttggactc	ttgttcacca	7680
ggttctggct	tttctctgtt	ctttacgcta	cctggtggta	cttggattgg	gataagccta	7740
gacaaggtgg	tagacctatc	caattcttca	gaagattggc	catctggaag	tacatgaagg	7800
attacttccc	tgtgtctctc	gttaagactg	ctgagcttga	tcctagccgt	aactacattg	7860
ctggattcca	tcctcatggt	gttcttgctg	ctggtgcttt	ccttaacctc	tgtactgagt	7920
ctactggatt	cacctctttg	ttccctggaa	tcagatccta	ccttatgatg	cttaccgttt	7980
ggttcagggc	tccattcttc	agagattaca	tcattgtctg	tggactcgtg	tctagcgaga	8040
aagtttctgc	tgatcacatc	cttagcagaa	agggtggtgg	aaaccttctt	gctatcatcg	8100
ttggaggtgc	tcaagaggct	cttgatgcta	gacctggtgc	ttacagactt	ctcttgaaga	8160
accgtaaggg	attcatcaga	cttgctctta	tgcatggtgc	tgctcttggt	cctatcttct	8220
cattcggaga	gaacaacctc	ttcaaccagg	ttgagaacac	tcctggaact	tggcttagat	8280
ggatccagaa	cagactccaa	aagattatgg	gaatcagcct	tccacttttc	catggaaggg	8340
gagttttcca	atactctttc	ggactcatgc	ctttcagaca	gcctatcact	actatcgttg	8400

pctau2012001598-seq1

gaaagcctat cgaggttcaa atgactcctc agccttctag agaggaagtt gatagacttc	8460
accagcgta catcaaagag ttgtgcaagc ttttcgagga acacaagctc aagttcaacg	8520
ttccagagga tcagcatctt gagttctgct aatgactcga gaagtatgaa ctaaaatgca	8580
tgtaggtgta agagctcatg gagagcatgg aatattgtat ccgaccatgt aacagtataa	8640
taactgagct ccatctcact tcttctatga ataaacaaag gatgttatga tatattaaca	8700
ctctatctat gcaccttatt gttctatgat aaatttcctc ttattattat aaatcatctg	8760
aatcgtgacg gcttatggaa tgcttcaa atgtacaaaa caaatgtgta ctataagact	8820
ttctaaacaa ttctaaccctt agcattgtga acgagacata agtggttaaga agacataaca	8880
attataatgg aagaagtttg tctccattta tatattatat attaccact tatgtattat	8940
attaggatgt taaggagaca taacaattat aaagagagaa gtttgtatcc atttatatat	9000
tatatactac cattttatat attatactta tccacttatt taatgtcttt ataaggtttg	9060
atccatgata tttctaatat tttagttgat atgtatatga aagggtacta tttgaactct	9120
cttactctgt ataaaggttg gatcatcctt aaagtgggtc tatttaattt tattgcttct	9180
tacagataaa aaaaaatta tgagttgggt tgataaaata ttgaaggatt taaaataata	9240
ataaataaca tataatatat gtatataaat ttattataat ataacattta tctataaaaa	9300
agtaaatatt gtcataaatc tatacaatcg tttagccttg ctggacgaat ctcaattatt	9360
taaacgagag taaacatatt tgactttttg gttatttaac aaattattat ttaacactat	9420
atgaaatfff tttttttatc agcaaagaat aaaattaaat taagaaggac aatggtgtcc	9480
caatccttat acaaccaact tccacaagaa agtcaagtca gagacaacaa aaaaacaagc	9540
aaaggaaatt ttttaatttg agttgtcttg tttgctgcat aatttatgca gtaaaacact	9600
acacataacc cttttagcag tagagcaatg gttgaccgtg tgcttagctt cttttatfff	9660
atftttttat cagcaaagaa taaataaaat aaaatgagac acttcaggga tgtttcaaat	9720
ttaaadcctg cagg	9734

<210> 416
 <211> 10374
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> synthesized DNA fragment

<400> 416	
ttaattaata atgtataagc attaaattaa agatactaga atcaattttg taattataaa	60
cgtaattaaa tcttaaatct tcgaggttag accttcagga aatagagatg cttgcagaag	120
atcccgcagt ctgcgatcgc gacgtcctag aatctaatta ttctattcag actaaattag	180
tataagtatt tttttaatca ataaataata attaataatt tattagtagg agtgattgaa	240
tttataatat atftttttta atcatfttaa gaatcttata tctfttaaatt gacaagagtt	300
ttaaattgggg agagtgttat catatgacaa gtaggattaa tgtgttatag tttcacatgc	360

pctau2012001598-seq1

attacgataa	gttgtgaaag	ataacattat	tatatataac	aatgacaatc	actagcgatc	420
gagtagtgag	agtcgtctta	ttacactttc	ttccttcgat	ctgtcacatg	gcggcggccc	480
gaattctcat	caccaccact	cgaaatcatt	gaaggacacc	tgatcatcag	aggagcccac	540
gtaagcatga	tcctctggag	tcatagaaga	ggtggcctgg	tggtgctgaa	gatcagtagt	600
gagatccatg	aggtcgtttg	ggattgggta	gtggtcgtta	aggttgccga	tgagcatagc	660
aatggaggcc	tcttcttcat	ctcttgggat	gtagtcgagc	atggcctctt	cgtaatgatt	720
cctctcacga	gactcgttct	cacgaccctc	attctgctta	ccggagtttc	cgttcatggc	780
gaagtagagg	ttcttggact	cgtcctcata	gatagtgagg	aagtcgccga	tctcagcacc	840
gttttgcttc	acgaactcac	cagtgttctc	gagcacgtac	atcctggact	tgttgttgga	900
ccagaacttg	tacttgaagg	accaggactg	catagagaac	acgtccctca	tctgaacgac	960
gatgccctct	ttgtcagaga	gctttggaag	gttggcctca	gcattctctt	ttggaagcac	1020
aattcttccg	agggatccca	catcggagtt	cttgagttct	ttctcgcaga	gcaccctgag	1080
tttcttggtg	tcgaaagtgg	actgctggta	ggagtccttc	ttgtcggaga	tctgcacggt	1140
gttcttgagg	ttgagcatca	tgagctgcct	cttagaatcc	acaagctcgg	atggtgaaga	1200
gtttggggag	gagttgttac	gggacctcat	catggcgttc	ttacggttga	tcctggccat	1260
cttagtcatc	cttgggtcaa	gaaatcctcc	cctctcttgc	caaaggcatg	ggtttccaac	1320
aagagatccg	aactcggatc	cagtctgagt	cactggaatc	tgtgggtagc	actctgagag	1380
gtacactggc	tgtggattca	taacagcagc	catctgctca	acggggtaag	agaatggtgg	1440
aaggaagtga	tgaggctgag	cctggtgatg	atcgtaagtt	ggcacagtag	tgaagtggga	1500
cctgttggtg	ttgggggtcca	tggaaagctc	ttgcacagag	ttagcgttgg	aggatgggaa	1560
tgggaggaag	ttgtccatgg	tgaattcgct	agcggttgat	gatgttggtg	agactgagaa	1620
ccttgggggt	ctttatatag	atgttcagct	atgcggggat	gcgacctctg	ggtaggagag	1680
cacgtacggg	gccccgctcg	tttgtggctc	tccgtgcgga	catcccgtgc	ggacagtacc	1740
agaaagtgct	ccgtctctgc	tctatacggc	tctatacgcg	tacctcttga	acggtgcgtg	1800
gagaggagtg	gtgtgtcaat	ttccgccccg	ccctcgtgcg	gttccgcatg	catccaatcc	1860
taggtcggaa	ctatcccag	ctgcggatgc	cgatgcggac	ggacaagtgg	gaactatcac	1920
aactatcgag	acgtcgcccc	ttcgactcag	atcttccaag	gcctcgtctc	cgagtccgct	1980
gcttctcgcc	gcgccgatca	cttctccgcc	gccaacaagg	cttgtagtta	ataggaatca	2040
ttcagggatt	gtgattccgg	gcagtagtaa	ttaataatat	agtattagta	tagataatat	2100
gtttcgtttg	ggatcttttg	aacgttgctc	tgttccttgt	tgttcatttt	aaagcttttg	2160
agggatagtt	gcagaactgt	tcggtgatgc	ttcatcctct	caagaactag	at ttgggtaa	2220
agaaacatcc	atgcatggat	atggaatggt	gttcttccga	ttggagatta	ttttataaaa	2280
tttaaaattc	atgatttaaa	aaaacacata	aaaaccacaa	aattcatgat	ttattgacaa	2340
tacgatacaa	aattagcacc	accggctact	ggctcattac	acatttcccc	ttcccctcat	2400

pctau2012001598-seq1

tctcactttg	tggctttatt	attattatta	ttacatatat	tttaccgtta	ttatttcacg	2460
tcacataagc	ttgactagtt	tagaggaagc	tgtcgacctt	tccgttgagt	cccctctcaa	2520
gatccttagc	cacgagagca	gccatatcag	acttacggcc	aaggaactgg	agagccctga	2580
tgcagtagtt	gcagttctgc	ctaagttgaa	ggtatccagc	gagagcgttt	ccctgaccat	2640
caagatcaag	gcaatcgagg	aggatgtaag	cagccttagc	aaatgagtgg	atgagcacca	2700
cagtccatgg	ctcaacaagg	aatgcaacct	cagtaaggga	gtcgtggaag	tgccctgtaca	2760
ttggagaagc	cttgccctgga	tcgagttctt	ggcctctaag	agatgacctt	gcgagagcag	2820
acacagcaag	agcgttatca	agcctgagct	tagtcacgag	gtcgatgtgg	ttgatatcgg	2880
agtaggaatc	cctgtc gatg	gactcgagca	catcagccaa	aagcatagca	gcaaggtgcc	2940
agtgtccatc	aaggatcacg	taccaagact	ggatcctgga	tgggagaagc	tcatggttgg	3000
caacgcaatc	ctgcatgaat	ggctggtact	tagcagtgca	gtggttgtag	acgtagagag	3060
tcctctgaat	agcagcctca	agcctagctg	gagaagctcc	cctatacaac	aaagtctgga	3120
gctgagtcac	acgcctgtaa	agcaacacct	taacaggagt	agcagaggag	agagcagcag	3180
cagcttgttc	ataagtgc at	ggccacctag	cagctgggtg	agaaatttga	gtgtgggact	3240
cgtggtcggg	caaagaatca	gaagtcctaa	ggaacaggtc	gccccacaca	tcagacttct	3300
cttgattaga	tgggacctgc	cttggtggct	cccagcaatc	aaggtttaata	ggagtctcgg	3360
cacctcttct	tgggtggagaa	gcagaagaga	tctgggagtc	ctcatcggac	acaacaagtg	3420
gcctctggta	catagcagca	gacagagtgt	cgaacatgat	cccagagccag	aacatgagag	3480
acagagtaga	cctgtgttcc	tcagaagcca	ccactggatc	aagttgtgga	gactcagatg	3540
aggggggtggg	agtctcgaaa	attccagcga	aagtagaagc	cacggatcca	ccaggaagcc	3600
tattgaaagc	cttacccttt	ctctgcatcc	tagcgaactt	gtgcctgaag	gtgtagagct	3660
gcctattagc	agtctcgagg	aacactggag	caccatcatt	ctcgagaagc	ttatcgagcc	3720
tagctcccat	tccatgtttg	tcatcatcat	cgagcacgga	ctgagtgagg	ctgaagatga	3780
tgttggcgaa	gatcaccttg	aaggatggga	ttccagtagt	gtgctgaagt	gcatgcctag	3840
cctcattcca	agcgttacgc	ctaatagaac	gctcgtcagc	agcaatatca	gctgggacgt	3900
taagtccagc	tcccctttga	gcatgttgag	tccactgaga	agcgaaagcc	acaatagcaa	3960
ggtgaagtgc	cctagcagca	gccttatctt	cttcagcgga	gagagccctt	cctctaagag	4020
aagttgacac	cctatcgagc	ctgcacaccc	taatgcacat	cctgttagac	caatttggtc	4080
cccactcagc	cctttgcttt	ggtgggaggt	aagaaatctg	gtcggagtat	gggcagttgt	4140
gctcagtaag	ccagcaagaa	agggcgtttt	ccatggagtc	gtggtagatc	ctcatgaggt	4200
tcttggatcat	ggtggacctt	gcataccttc	tagcggtaga	gtcgggaagcc	atgcacaaac	4260
tgttttctgg	gaggagtgga	gactgcctat	cagaccaaag	agagccatcc	tgagcaactt	4320
gctgaacgga	gttagtgtca	gtcacggaat	ctggatcggg	agtctgttcc	tcaagctcac	4380
gaatgagatc	atcagtgttt	ggggagtgtg	gtggaagaag	aacagcctgg	aaagacaatg	4440

pctau2012001598-seq1

gggagagggg	cttctcaagt	tgctgtccca	tagagaaatc	gcctgggata	gcaagatccc	4500
atgggaatgg	agcatcagat	tgcacgttag	cagcatcctc	agcgttagct	ggaatagcag	4560
aatgggagaa	gtcgaagaga	tcgccagggt	gagaaagaag	tccctgagtc	caagatggca	4620
aagcatcgtg	ggagttgatc	actggtggag	catcgtgatt	atcggactcg	ggagttggaa	4680
tagtagcagc	ggaagtagaa	ggctcggagg	tagtagtagc	agtcctagcc	ttcttagtcc	4740
tagctcttgg	agcagcaccc	ttagccttag	acctttggga	gctaagccag	ttgaaggtgc	4800
agtccttggt	ccacctcttg	cagttagagc	atgacaccca	tccgttctca	ttagcctcgt	4860
tacggttctc	tggagcatca	catcttctct	tgcccttacg	gcatgggtca	caagagtggg	4920
tctgccttct	tctagtgtcg	gccatggtac	tagtactcat	tattaaaaca	aacataaaca	4980
tattattata	aaaactaaaa	acattacaaa	cttacaaatt	tctctgaagt	tgtatcctca	5040
gtacttcaaa	gaaaatagct	tacaccaa	tttttcttgt	tttcacaaat	gccgaacttg	5100
gttccttata	taggaaaact	caagggcaaa	aatgacacgg	aaaaatataa	aaggataagt	5160
agtgggggat	aagattcctt	tgtgataagg	ttactttccg	cccttacatt	ttccacctta	5220
catgtgtcct	ctatgtctct	ttcacaatca	ccgaccttat	cttcttcttt	tcattgttgt	5280
cgtcagtgtc	tacgtcttca	agattctttt	cttcgcctgg	ttcttctttt	tcaatttcta	5340
cgtattcttc	ttcgtattct	ggcagtatag	gatcttgtat	ctgtacattc	ttcatttttg	5400
aacatagggt	gcatatgtgc	cgcataattga	tctgtcttct	gctgagctca	cataatactt	5460
ccatagtttt	tcccgtaaac	attggattct	tgatgctaca	tcttggcccg	ggtcgatagt	5520
tgtgatagtt	cccacttgtc	cgtccgcac	ggcatccgca	gctcgggata	gttccgacct	5580
aggattggat	gcatgcggaa	ccgcacgagg	gcggggcgga	aattgacaca	ccactcctct	5640
ccacgcaccg	ttcaagaggt	acgcgtatag	agccgtatag	agcagagacg	gagcactttc	5700
tggtagtgtc	cgcacgggat	gtccgcacgg	agagccacaa	acgagcgggg	ccccgtacgt	5760
gctctcctac	cccaggatcg	catccccgca	tagctgaaca	tctatataaa	gacccccaa	5820
gttctcagtc	tcaccaacat	catcaaccgc	tagcggggccc	accatgaaga	agcgactcac	5880
tacctccact	tgctccagct	ctccatcttc	ctctgtgtcc	tcttccacta	ccacctctag	5940
cccaattcaa	tctgaggctc	caagacctaa	gagggccaag	agagctaaga	aatcctctcc	6000
atccggcgac	aagtctcaca	acccaacttc	tccagcttcc	acccgtaggt	catctatcta	6060
cagagggtgt	actaggcaca	ggtggactgg	aaggtttgag	gctcatctct	gggacaagtc	6120
cagctggaac	tccatccaaa	acaagaagg	caagcagggt	tacctcggcg	cttacgattc	6180
tgaagaggct	gctgctcaca	cttacgatct	tgctgtctct	aagtactggg	ggccagacac	6240
tattctcaac	ttcccagctg	agacttacac	caaagaactc	gaagagatgc	agagggtgac	6300
caaagaagag	tacctcgctt	ctctcaggcg	tcagtcatct	ggattttcta	ggggcggtgc	6360
aaagtacagg	ggagttgcta	ggcatcatca	caatggaaga	tgggaggcta	ggatcgggaag	6420
ggtgttcgga	aacaagtacc	tctacctcgg	cacctacaac	acccaagaag	aggcagctgc	6480

pctau2012001598-seq1

tgcttacgac	atggctgcta	ttgagtacag	gggtgctaac	gctgtgacca	acttcgacat	6540
ctccaactac	atcgacaggc	tcaagaaaaa	gggcgtgttc	ccattcccag	tgaaccaggc	6600
taatcaccaa	gagggaaatcc	tcgtcgaagc	taagcaagag	gttgagactc	gtgaggccaa	6660
agaggaaacct	agggaagagg	ttaagcagca	gtacgttgag	gaaccacctc	aagaagagga	6720
agaaaaagag	gaagagaagg	ccgagcaaca	agaggctgag	atttgtgggat	attccgaaga	6780
ggccgctgtt	gtgaactgct	gcatcgactc	ctccaccatc	atggaaatgg	ataggtgcgg	6840
cgacaacaac	gagcttgctt	ggaattttctg	catgatggac	accggcttct	ccccattcct	6900
tactgatcag	aacctcgcca	acgagaaccc	aattgagtac	cccgagcttt	tcaacgagct	6960
ggctttcgag	gataacatcg	acttcatgtt	cgacgacgga	aagcacgagt	gccttaacct	7020
tgagaacctc	gattgctgtg	tgggtgggaag	agaatcccca	ccctcttcat	cttcaccact	7080
cagctgcctt	tctaccgact	ccgcttcttc	cactactact	accactacct	ccgtgtcctg	7140
caactacctc	gtgtaagggc	ccgccgccat	gtgacagatc	gaaggaagaa	agtgtaataa	7200
gacgactctc	actactcgat	cgctagtgat	tgtcattgtt	atatataata	atgttatctt	7260
tcacaactta	tcgtaatgca	tgtgaaacta	taacacatta	atcctacttg	tcatatgata	7320
acactctccc	catttaaaac	tcttgtcaat	ttaaagatat	aagattcttt	aaatgattaa	7380
aaaaaatata	ttataaattc	aatcactcct	actaataaat	tattaattat	tatttattga	7440
ttaaaaaaat	acttatacta	atttagtctg	aatagaataa	ttagattcta	gcccgggcct	7500
gcaggcatgg	agtcaaagat	tcaaatagag	gacctaacag	aactcgccgt	aaagactggc	7560
gaacagttca	tacagagtct	cttacgactc	aatgacaaga	agaaaatctt	cgtcaacatg	7620
gtggagcacg	acacacttgt	ctactccaaa	aatatcaaag	atacagtctc	agaagaccaa	7680
agggcaattg	agacttttca	acaaagggta	atatccggaa	acctcctcgg	attccattgc	7740
ccagctatct	gtcactttat	tgtgaagata	gtggaaaagg	aaggtggctc	ctacaaatgc	7800
catcattgcg	ataaaggaaa	ggccatcggt	gaagatgcct	ctgccgacag	tgggtcccaa	7860
gatggacccc	caccacgag	gagcatcggt	gaaaaagaag	acgttccaac	cacgtcttca	7920
aagcaagtgg	attgatgtga	tatctccact	gacgtaaggg	atgacgcaca	atcccactat	7980
ccttcgcaag	acccttcctc	tatataagga	agttcatttc	atttggagag	aacacggggg	8040
acttctagaa	ccatggctat	cctcgactcc	gctgggtgtga	ctactgttac	tgaaaatggg	8100
ggtggcgagt	tcgtggatct	tgatagactt	aggcgtagga	agtccaggtc	cgattcttct	8160
aacggacttc	tcctctccgg	ctccgataac	aactctccat	ctgatgatgt	tggggctcca	8220
gctgatgtga	gggataggat	tgattccgtg	gtgaacgatg	atgctcaggg	cactgctaata	8280
ctcgctggtg	ataacaatgg	tggcggagat	aacaacggtg	gtggaagagg	tgggtggtgag	8340
ggaagaggta	atgctgacgc	tactttcacc	tacaggccat	ctgttccagc	tcatagaagg	8400
gctagagaga	gccctctttc	ctccgatgct	atcttcaagc	agtctcacgc	tggactcttc	8460
aacctttgcg	tgggtggtgct	tatcgctgtg	aactccaggc	tcattatcga	gaacctgatg	8520

pctau2012001598-seq1

aagtacggct	ggctcatcag	gaccgatttc	tggttctcta	gccgttctct	tagggactgg	8580
ccactcttta	tgtgctgcat	ctccctcagc	attttccac	tcgctgcttt	cactgttgag	8640
aagcttgtgc	tccagaagta	catctctgag	cccgtggtga	ttttcctcca	catcatcatc	8700
accatgaccg	aggtgctcta	cccagtgtac	gttacactta	ggtgcgactc	cgctttcctc	8760
tctggtgtga	ctcttatgct	cctcacttgc	atcgtgtggc	tcaagcttgt	ttcctacgct	8820
cacacctcct	acgacattcg	ttctcttgct	aacgctgccg	acaaggctaa	tcccgaagtg	8880
tcttactacg	tgtcccttaa	gtccctcgcc	tactttatgg	tggtccaac	tctttgctac	8940
cagccatctt	atccaaggtc	cgctgtatt	aggaagggat	gggttgcaag	gcagttcgct	9000
aagctcgtga	tcttactg	attcatgggc	ttcattatcg	agcagtacat	caaccccatc	9060
gtgaggaact	ctaagcacc	tcttaagggc	gatctcctct	acgctattga	gaggggtgtg	9120
aagctctctg	tgccaaacct	ttacgtgtgg	ctctgcatgt	tctactgctt	ttccacctc	9180
tggtcaata	tcctcgctga	gcttctttgc	ttcggcgaca	gagagtctta	caaggattgg	9240
tggaacgcta	agtccgtggg	cgattattgg	aggatgtgga	acatgccagt	gcacaagtgg	9300
atggtgaggc	acatctactt	tccatgcctc	aggccaaga	tcccaaagac	cctcgctatc	9360
attatcgcat	tcctcgtgtc	tgctgtgttc	cacgagtgtg	gcattgctgt	tccatgccga	9420
cttttcaagc	tctgggcttt	cctcggaatc	atgttccagg	ttccactcgt	gttcatcacc	9480
aactacctcc	aagagagggt	cggctccact	gttggcaaca	tgattttctg	gttcatcttc	9540
tgcatcttcg	gccagccaat	gtgcgtgttg	ttgtactacc	acgacctcat	gaaccgtaag	9600
ggctccatgt	cttaatctag	aaggggtgat	cgattatgag	atcgtaaaa	gacactgcta	9660
ggtgttaagg	atggataata	ataataataa	tgagatgaat	gtgttttaag	ttagtgtaac	9720
agctgtaata	aagagagaga	gagagagaga	gagagagaga	gagagagaga	gagagagaga	9780
gaggctgatg	aaatgttatg	tatgtttctt	ggtttttaaa	ataaatgaaa	gcacatgctc	9840
gtgtggttct	atcgaattat	tcggcggttc	ctgtgggaaa	aagtccagaa	gggccgccgc	9900
agctactact	acaaccaagg	ccgtggagga	gggcaacaga	gccagcactt	cgatagctgc	9960
tgcgatgatc	ttaagcaatt	gaggagcgag	tgacatgca	ggggactgga	gcgtgcaatc	10020
ggccagatga	ggcaggacat	ccagcagcag	ggacagcagc	aggaagttga	gaggtggtcc	10080
catcaatcta	aacaagtcgc	tagggacctt	ccgggacagt	gcggcaccca	gcctagccga	10140
tgccagctcc	aggggcagca	gcagtctgca	tggttttgaa	gtggtgatcg	atgagatcgt	10200
ataaagacac	tgctaggtgt	taaggatggg	ataataagat	gtgttttaag	tcattaaccg	10260
taataaaaag	agagagaggc	tgatggaatg	ttatgtatgt	atgtttcttg	gtttttaaaa	10320
ttaaattggaa	agcacatgct	cgtgtgggtt	ctatctgcag	gcggccgcgg	cgcc	10374

<210> 417
 <211> 434
 <212> DNA
 <213> Artificial Sequence

pctau2012001598-seq1

<220>

<223> hpRNAi

<400> 417

gtgagcaatg aaccaagatt tatcaatacc gttacttttg atagcaaaga gggttctcct	60
actcttggtta tgggtccacgg atatggtgcc tctcaggggt tcttctttcg gaatTTTTat	120
gcccttgcgga ggcatttcaa agttattgct attgatcagc ttggctgggg tggttcaagc	180
aggcctgact tcacatgcag aagtacagaa gagactgaag attggtttat tgattccttt	240
gaggagtggc gcaaagccaa aaaccttagc aactttatTT tgcttgggca ctcctttgga	300
gggtatgtcg ctgcaaaata tgctctcaag catccagagc atgttcagca gttgattctg	360
gtaggaccag ctggatttac atcagagact gaacatatgt ccgagcggct taccagttt	420
agagcaacat ggaa	434

<210> 418

<211> 593

<212> DNA

<213> Artificial Sequence

<220>

<223> hpRNAi

<400> 418

actgctgatg ctgtcaggca gtatctatgg ttgtttgagg agcataatgt tcttgaattc	60
ctcgtacttg ctggagatca tctatatcga atggattatg aaaagttcat tcaagcccac	120
agagaaacag atgctgatat tactgttgcc gcactgccaa tggatgaaaa gcgagccact	180
gcatttggtc tcatgaagat tgacgaagaa ggacgcatta ttgaatttgc agagaaaccg	240
aaaggagagc aattgaaagc aatgaaagtg gatactacca ttttaggtct tgatgatgag	300
agagctaaag agatgccttt tatcgcaagt atgggtatat atgtcattag caaagatgtg	360
atgttaaact tacttcgtga taagttccct ggtgccaatg attttggcag tgaagttatt	420
cctggtgcaa cttcgcttgg gatgagagtg caagcttatt tatatgatgg atactgggaa	480
gatattggta ccatcgaagc tttctacaat gccaatTTgg gcattaccaa aaagccagtc	540
ccagatttta gcttctatga ccgatcagct ccaatctaca cccaacctcg ata	593

<210> 419

<211> 5

<212> PRT

<213> Artificial Sequence

<220>

<223> lipase motif

<220>

<221> misc_feature

<222> (2)..(2)

<223> Xaa can be any naturally occurring amino acid

<220>

<221> misc_feature

<222> (4)..(4)

<223> Xaa can be any naturally occurring amino acid

<400> 419

Gly Xaa Ser Xaa Gly
1 5

<210> 420

<211> 6

<212> PRT

<213> Artificial Sequence

<220>

<223> acyltransferase motif

<220>

<221> X

<222> (2)..(5)

<223> any amino acid

<400> 420

His Xaa Xaa Xaa Xaa Asp
1 5

<210> 421

<211> 7

<212> PRT

<213> Artificial Sequence

<220>

<223> probable lipid binding motif

<220>

<221> X

<222> (2)..(4)

<223> any amino acid

<400> 421

Val Xaa Xaa Xaa His Gly Phe
1 5

<210> 422

<211> 1224

<212> DNA

<213> Arabidopsis thaliana

<400> 422

atggcggaag aaatctcaaa gacgaagggt ggatcttctt ctactgcttc ggtggctgat	60
tcattctgctg ctgcgtcggc tgcaacgaat gcggccaaat caagatggaa aattttgtgg	120
cctaattcgc tccggtggat tcctacgtcc accgattaca tcatcgccgc cgagaaacgt	180
cttctctcca tcctcaagac gccttatgta caagagcaag tcagtattgg ttcaggacca	240
ccaggttcta aaatcagggt gtttaggtct acgagcaatg agtcacgtta catcaacact	300
gttacatttg atgccaagga gggagctcct acactcgtca tggttcatgg ttatggtgct	360
tctcaagggt ttttcttccg taattttgat gctcttgcca gtcgatttag agtgatcgct	420

pctau2012001598-seq1

attgatcaac	ttgggtgggg	tggttcaagt	aggcctgatt	ttacatgtag	aagcacagaa	480
gaaactgagg	catggtttat	cgactccttt	gaggaatggc	gtaaagccca	gaatctcagt	540
aactttattc	tattaggaca	ttctttttgga	ggctatgttg	ctgctaaata	cgcgcttaag	600
catcctgaac	atgttcaaca	cttaattctg	gtgggatctg	ctgggttctc	agcagaagca	660
gatgccaaat	cagaatggct	cactaaattt	agagcaacat	ggaaagggtgc	agtcctaaat	720
catttatggg	agtcaaattt	cactcctcag	aagctggtta	gaggattagg	tccttgggggt	780
ccaggtcttg	taaatcggtg	tacaactgca	agatttggtg	cacattcgga	gggaactggg	840
ctaacagaag	aggaagccaa	attgctaacc	gattatgtgt	accatacttt	ggctgcaaag	900
gctagtggag	agttatgctt	gaaatacatc	ttctcatttg	gagcatttgc	taggaagccc	960
ctcttacaaa	gtgcatcaga	gtggaaagtg	ccaacaacgt	ttatctatgg	aatgaatgat	1020
tggatgaact	atcaagggtgc	ggtggaagcg	aggaaatcca	tgaagggtccc	ttgcgaaatc	1080
attcgggttc	cacagggtgg	tcattttgtg	ttcatagaca	acccaattgg	ttttcattct	1140
gcagtgcttt	atgcttgccg	caagtttata	tctcaagact	cctctcatga	tcaacaactc	1200
ctagatggtc	tacgattggg	ttag				1224

<210> 423

<211> 1700

<212> DNA

<213> *Brachypodium distachyon*

<400> 423

tccgcgcccc	aaacgatccc	aacagaagct	ctaattctcca	aagccgccgc	gcgtgttgag	60
ggtggtgcgg	ggaaaagctt	ggtgtcgtga	gccccgtgt	cgcatgcgcc	gcgctgccgc	120
cgtgacgagg	atggcagcga	ccgaggagat	gaggcaggcg	tccgccgccg	ccgccgccac	180
ggtgaccgag	gcctcggcgt	cggcggcccc	gcccgcgggg	tccagggtggg	cgcggtgtg	240
gccggccgcg	ctgcgctgga	tccccacctc	caccgagcgc	atcatcgccg	ccgagaagcg	300
cctcctctcc	gtactcaaaa	ctgggtatgt	ccaagaacaa	gttaacattg	gctcggctcc	360
acccgggtca	aaagtaagat	ggtttagatc	atcaagtgat	gagccaaggt	tcataatac	420
agtgacattt	gatagcaagg	agaatgctcc	cactcttgtc	atggtccatg	gttatggtgc	480
ttcacagggt	ttcttcttta	gaaattttga	tgcccttgca	agccgtttcc	gagtgattgc	540
cattgatcag	cttggttggg	gtggatcaag	tagacctgac	ttcacctgta	aaagtaccga	600
agaaaactgag	gcttggttca	tagattcttt	agaggaatgg	cgtaaagcaa	agaacctcag	660
taattttata	ttgctcggtc	attctttcgg	aggatatgtt	gcagcaaaat	atgccttgca	720
gcacctgaa	cacgtgcagc	acttaatttt	ggtcggttct	gctgggtttt	catcagaaac	780
agatcatagc	tctgagtggg	taaccaagtt	tcgagcaaca	tggaaaggca	tgctagtgaa	840
ccaactatgg	gagtccaatt	ttactcctca	aagaattgta	agaggattgg	gtccttgggg	900
cccagatttg	gttcgcagat	ataccactgc	taggtttggc	tcatattcaa	cagggtgaatt	960
actaacagaa	catgagtccg	gcttgctgac	agattacatt	taccatacat	tagctgccaa	1020

pctau2012001598-seq1

agctagtgga	gagctgtgct	tgaaatatat	tttttccttg	ggggcatttg	caaggaaacc	1080
tcttctgcag	agtgcacatg	actggaaagt	gccgaccact	ttcatatatg	gccatgacga	1140
ttggatgaaa	taccaggggg	cacagcaagc	acgcaaggat	atgaaagtgc	cttgcgaaat	1200
catcagagtc	ccacagggag	gacattttgt	gttcatagat	aacccttccg	ggttccattc	1260
ggcagtcctc	tatgctgtgc	ggaaattttt	atctggagat	gcagaggagg	gtctctctct	1320
tcctgatggc	ctgatatctg	catgacagca	tgaggcgcg	tgtcatacca	attagcggta	1380
tgaacacaaa	gcaaagctat	acggagctag	gaaatgttac	aaatgtcacg	actcaccaga	1440
aatgttacaa	atgtcaccac	tcaccagttt	cctttttgta	tgtatgaatt	gtgtgaatat	1500
acacgtcatt	catatttgcc	ggcgtatcag	tacttcaata	gtgataaaac	atgatcatat	1560
atatatgtat	gatttctcta	gtcggttctc	atcaagtcaa	gttattgtga	ttggtgaatg	1620
atatactttc	caggtcaact	ttgtgtttgc	atgtacaaac	tatcatggaa	catatcagta	1680
tagtttatga	tttgtcttcc					1700

<210> 424
 <211> 1484
 <212> DNA
 <213> Glycine max

<400> 424						
gtcatggatg	cgcgctcactg	ctcgcgttca	ttataatggc	ggaagagata	accaagaacg	60
acgtcggagt	aacctccaaa	accaccagaa	gcagctccag	gttctggcct	cgttggattc	120
ccacttccac	cgatcacatc	atcgctgccg	agaagcgctt	tctttccgtc	gtcaagactg	180
gttatgttca	agagcatggt	aacattggct	ctggctctcc	tggctccaaa	gtgaggtggt	240
tccgttcata	cagcaacgag	ccgcggttta	ttaacaccgt	tacatttgac	agtaaaccac	300
attctccaac	gcttgtcatg	attcatgggt	atgctgcttc	acaggggttc	ttttttcgca	360
atattgacgc	gcttgcgtct	cgatttagag	tcattgctgt	tgatcaactt	ggatggggtg	420
gatcgagcag	acctgatttc	acatgcaaaa	gcactgaaga	aactgaggca	tggtttattg	480
attcttttga	ggaatggaga	aaagccaaaa	acttgagcaa	ttttatactg	ctcggacatt	540
cttttggtgg	ttatgttgct	gccaaatatg	cgctcaagca	ccctgagcat	gtacaacact	600
tgattctggt	tggatctgct	ggattttcat	ctgaatcaga	tgcaaagtct	gagtggataa	660
caaggtttcg	agcaacatgg	aagggggcag	ttttgaacca	tctttgggaa	tcaaatttca	720
cacctcagaa	acttgtcagg	ggtttaggtc	cttgggggtcc	caacatagtc	cgcaagtata	780
caagtgctag	gtttggtaca	cattcaactg	gggaaatact	gactgaagag	gaatcaacat	840
tgctgacaga	ctatgtttac	cacacattgg	cggccaaagc	tagtggagag	ctgtgcttaa	900
aatatatatt	ttcatttgga	gcatttgcta	ggatgccctt	tcttctcagt	gcctcagagt	960
ggaaggtgcc	caccactttc	atgtatgggt	tccaagactg	gatgaattat	caaggtgccc	1020
aagaagctcg	caagcatatg	aaggttccat	gcgaaatcat	caggattcct	caggggtgggc	1080

pctau2012001598-seq1

actttgcgtt cattgacaac ccaactgcct tccattcagc tgttttttat gcttgtcgaa	1140
ggtttcttac acctgatcca gacaatgaat ctcttcctaa agggctaacc tctgcatagg	1200
ttaggtctta attttgtgct attcctgtct atatgtattt taatatTTTT ttttactaat	1260
taaatttcat aattgaatga aatcatatgt atattgtttc agtaaagtgg aatttactga	1320
aaatatttgt aatagcaact tcaacaaaaa tcgatttgta ggagaaattt cttccctgga	1380
aattgttcta ttttaaactc tgttgctcat aagatattat gacttcattc aactaataat	1440
tcatgtcgtt taggaaaagt agttagttat attaaatttg tcaa	1484

<210> 425
 <211> 1662
 <212> DNA
 <213> Zea mays

<400> 425	
accatacggg cggggccgca ccgaccgaac ctaaccgaga gcacgagcat acccgtccc	60
actccgactg cagagcatca gccgaggaga aaagtcggga gaaacgcgcg tgacgtctgc	120
ccgcgtcgta tgcgccgcgc tgccgtcgcc gcgacgacga cgacgaccag gatggcagcc	180
gaggagatga gacgggcctc cgcctcaacg gccacggcgg agatgccggc gtcgccggcg	240
ccggcgcaag cggggtcgag gtgggcgcgg gtgtggccgc gcgcgctccg gtggatcccc	300
acctccacgg accgcatcat cgccgccgag aagcgactcc tcacgatagt caaaactgga	360
tatgtccagg aacgagtcaa cattggctct gctccacctg ggtcaaaagt aagatggttt	420
aggtcagcaa gtgatgaacc aaggttcatt aatactgtaa catttgatag caaggagaat	480
gccccaccc tggttatggt ccatggctat ggagcttcac aggggttctt ctttcgaaac	540
tttgatgcc ttgcaagccg ttttaggtg attgccattg atcagcttgg ctgggggtggt	600
tcaagcagac ctgacttcac atgtaaaagt accgaagaaa ctgaggcatg gttcatagat	660
tctttcgagg agtggcgcaa ggccaagaac ctacgtaatt ttatatgct tggtcactct	720
tttgaggat atgttgctgc aaaatatgcg ctaaagcacc ctgaacacgt tcaacagttg	780
attttggttg gtctgtctgg cttctcatca gaaacagagc atagctctga gtggttaacc	840
aagtttcgag caacatggaa aggcattgcta atgaatcgtc tttgggagtc caattttact	900
cccaaaggg ttattagggg attgggtcct tggggtccag gtctagtaca gagatatacc	960
agtgccaggt ttggtacaag ttctactggt gaattactaa cagatgaaga atcggcattg	1020
atgacagatt atatgtacca tacgttagct gccaaagcta gtggagagct gtgcttgaaa	1080
tatatatttt cttcggggc atttgcaagg aaacctcttc tgcagtgcgc gtccgattgg	1140
aaagtgccga ctactttcat atatgggcag caagattgga tgaactacca aggcgctcag	1200
caagcacgga aggacatgaa agttccttgt gaaataatca gggtgccgca gggtgacat	1260
tttgtgttca tagacaacc ttcagggttc cactcggctg tcttctatgc gtgccgtaat	1320
cttctatcag taaatggaga ggagggttc acatttcctg atggcctaata atctgcgtga	1380
agtggcatgt tcaacaagct tgctcaacaa cagtttacat aaagcaaaga tatacgattg	1440

pctau2012001598-seq1

tggaatcat	tgccatttc	caccaatttg	cttgatacgc	gattatgctg	tgtatatatt	1500
acataacaaa	tgtattagta	tcatttaaatg	cacgatttgt	gaaagggcct	gagtttgtat	1560
ttagcgaatt	ttaggttggg	ttttttccct	ttttcttctt	tcagtgcgct	tgctagtcaa	1620
tcccatacta	taagccgtga	tcatttaaaa	aaaaaaaaaa	aa		1662

<210> 426
 <211> 1763
 <212> DNA
 <213> Sorghum bicolor

<400> 426	
actgcagatg	cgcggtcgtc ggctccggct cgcgaggcg agaacggcga accagcccgt 60
gtctctgttc	cctttcttcc ctttaaaaac acggcaaaaa aaaagctagc cggttacgct 120
accgaaccga	acggctcggc acgcgggcac gggcgcgggg tcgcaccgga aaagcacgag 180
cagagcagac	ctgacgtctc cagactgcag gagcatcatc agtcgaggag gaggaagtgt 240
ggggggggga	aagggaaacg tgcgcgtcgt atgcgcctcg ctgccgtcgc caggacgacc 300
aggatggcag	ccgaggagat gaggcggggc tcggcctccg cggcggtcgc ggccacgacg 360
gaggcggcgc	cggcgccggc gcaagcgggg tccaggtggg cgcggtgtg gccgagcgcg 420
ctccggtgga	tccccacctc cacggatcgc atcatcgccg cggagaagcg gtcctctcgc 480
atagtcaaaa	ctgggtatgt ccaggaacaa gtcaacattg gctcagctcc acctgggtca 540
aaagtaagat	ggtttaggtc agcgagtgat gaaccaaggt tcattaatac tgtaacattt 600
gatggcaagg	agaacgcccc caccctggtt atggtccatg gctatggagc ttcacagggg 660
ttcttctttc	gaaactttga cgcccttgca agtcgtttta ggggtgattgc cattgatcag 720
cttggctggg	gtggttcaag cagacctgac ttcacatgta aaagtaccga agaaactgag 780
gcatggttca	tagattcttt tgaggagtgg cgcaaggcca agaacctcag taattttata 840
ttgcttggtc	actcctttgg aggatatgtt gcggcaaagt atgccctaaa gcaccctgaa 900
cacattcagc	acttagtttt ggttggtcct gctggcttct cgtcagaaac agaccatagc 960
tctgagtggg	taaccaagtt tcgagcaaca tggaaaggca tgctagtga tcatctttgg 1020
gagccaatt	ttactcccca aagagttatt agaggattgg gcccttgggg tccaggtcta 1080
gtacaaagat	ataccagtgc caggtttggg acacgttcaa ctggtgatata actaacagat 1140
caagaatcaa	cattgttgac agattatatt taccatacct tagctgcaa agctagtgga 1200
gagctgtgct	tgaaatatat attttccttc ggggcatttg caaggaaacc tcttctgcag 1260
tgcgcatccg	attggaaagt gccgactact ttcatatatg gtcaggaaga ttggatgaac 1320
taccaagggg	ctcagcaagc acggaaggac atgaaagttc cttgtgaaat aatcaggggtg 1380
ccacagagtg	gacattttgt gtttatagac aacccttcag ggttccactc ggctgtcttc 1440
tacgcgtgcc	gtaatctttt atcccaaat ggggaggagg gcttcacatt tcctggtggc 1500
ctaatatctg	catgaagtgg catgttcaac aatcttatcg tgcccaacaa tagtttatat 1560

pctau2012001598-seq1

gaagcaaaga tatacgatgg tggaaatctt tgctcatttc caccaatctg gaaatatttg	1620
tgccctcttc caccaatttg tttgtatacg gattatgccg tgtatatatt ctgtgttgac	1680
tgaagaaac ataatgtatt aacattatgt aatgtatgta cgattcttta tttgattttc	1740
aacttgcaat acgcaagaac cac	1763

<210> 427
 <211> 1399
 <212> DNA
 <213> Ricinus communis

<400> 427	
cgcccttttta ccagtcaatt tccatttttta tatataagtg cttttgctta atttaagact	60
aactacagcg acgaattcgc gtttatgaaa ttgcttcgcc tacgactgct acgagtatct	120
agctcctcaa tatcatcaat aatggcggaa ggggctgctg ccacatcagc atcagcatca	180
gcgtcagcgt cagcgtcatg ggcaaaaaca agatctctac ggccatctgc tctccgttgg	240
atcccaactt caaccgatca catcatcgcc gccgaaaagc gtcttctctc cctcgtcaag	300
actccctatg ttgtggaaca agtgaatata gggctctggcc caccgggggc gaaggtagagg	360
tggttttcgtt ctaaaagcga cgaggcacgg tttattaaca cggttacttt tgatagcaaa	420
gaggaggatt ctctacact ggttatgggt catggatatg ctgcttctca aggcttcttc	480
tttcgcaatt ttgatgctct tgcttctcgt ttcaggctca ttgctattga tcagctcggg	540
tgggggtggat caagtagacc tgattttacg tgtaagagca ctgaagaaac tgaggcatgg	600
ttcattgact cttttgaggc ttggcgtaaa gagaaaaacc tcagtaactt catcttactt	660
ggacattctt tcggagggta tattgcagct aaatatgcac tcaagcatcc tgagcatggt	720
caacatctga ttttagtggg atctgctgga ttttcatcag aatctgaaga caaatctgag	780
cagcttactc ggttcagagc aacatggaag ggagcagttt tgaatcattt atgggagtct	840
aattttactc ctgagaaggt tattagaggt ttaggtcctt ggggtccaga tctcgtacgc	900
aagtacacaa ctgctagatt tggttcatat tcaactgggt agatattaaa ggaggaggag	960
tccaaattgc ttacagacta tgtgtaccat acctagccg ccaaagctag tggagagcta	1020
tgcttgaaat atatattttc ttttgagca tttgctcgga tgcccttctt acaaagcgcg	1080
tcacaatgga aagtccaac tactttcata tatgggatgc aagattggat gaattatcaa	1140
ggggcccaa gagctcgcaa agatatgaat gtcccatgtg aaatcattag ggttcctcag	1200
ggcgggcact tcgttttcat agacaacca actgggtttc attcagctgt gttatatgcc	1260
tgccggagat ttctctcacc cgatcctgat aatgaatctc ttcctgaagg tctgatatct	1320
gcgtaggaag tgtggtttgt aattatttct tttttatttg ttgtgtataa tttatctgag	1380
aatttccaat tctttcaat	1399

<210> 428
 <211> 1480
 <212> DNA
 <213> Medicago truncatula

pctau2012001598-seql

```

<400> 428
ggttggctca tagttccttt tacctgttga aaacaaaaca tatggagtaa ctttttagtc      60
agaaattcaa agctacgcac ttgattaaac taattatcga aaaatggcgg aagaaattag      120
acaaaaggac gacgtcgatt catcttcgaa atctaaaagc ttctggctctt cactccgttg      180
gattcccact tctaccgatc atatcatcgc cgctgagaaa cgccttcttt ccattatcaa      240
gactggggtat gctcaagagc atgttaatat aggttctggt cctcctggct ctaaagttag      300
atggttccgt tcaaccagta acgagccacg ctttctcaac actgttacat ttgatagtaa      360
acccgattct cctacacttg ttatggttca tggatacgct gcttctcagg gtttcttctt      420
tcgcaatttt gatgctctcg cctctcggtt cagaatcatt gctgttgatc aacttggttg      480
gggaggatca agcagacctg atttcacatg caaaagtacc gaagaaactg aggcattggtt      540
cattgattct ttcgaggaat ggagaaaagc caaaaatctt accaatttca tactgcttgg      600
acattctttt ggtgggttat ttgcttccaa atacgcgctc aagcaccctc agcacgtaca      660
acacttaatt ttggtgggac ctgccgggtt tacagaagaa acagatccaa agactgagtt      720
tggtactaag tttcgagcaa catggaaggg agcagttctg aaccatctat gggaatctaa      780
ttttacacct cagaaaattg tcagagggtt aggtccttgg ggtcctaaca tgggccgcaa      840
atatacaagt gctaggtttg gtacacattc aaccgggcaa aaactgattg acgaggaatc      900
aagtctgctg actgattatg tttatcatac attggcggcc aaagctagtg gggagctgtg      960
tttaaaatat atttttgcat ttggagcatt tgctaggatg ccccttcttc aaagtgtcga     1020
agagtggaag gtgcccacca cattcatata tggttacgaa gattggatga attatgaagg     1080
tgcccaagaa gtcgcaagc atatgaaggt tccatgtgaa attatcaggg tccctaaggc     1140
cggccatttt gtgttcattg acaaccaag tggcttccat tcagctgtgt tttatgcttg     1200
tcgaagggtt cttaccccaa attcggacaa tgaatctctt cccgaagggc tatcgtctgc     1260
ttaggattta attttgcac aatccagtgt atattaatat ggttattaat ttttttttac     1320
ttcataactg aatgaagccg tgtcttgttt ctcaagtgaag tggaatataa tggaaatata     1380
tgtaattgta ataacaataa tattgatttg ttggggaact ttgaggacaa aaacatattc     1440
tggtaaaatt ttgttgaca tgcgacaaac atatgctgtg                               1480

```

```

<210> 429
<211> 407
<212> PRT
<213> Arabidopsis thaliana

```

```

<400> 429
Met Ala Glu Glu Ile Ser Lys Thr Lys Val Gly Ser Ser Ser Thr Ala
1          5          10         15

Ser Val Ala Asp Ser Ser Ala Ala Ala Ser Ala Ala Thr Asn Ala Ala
          20         25         30

```


pctau2012001598-seql

Lys Ser Arg Trp Lys Ile Leu Trp Pro Asn Ser Leu Arg Trp Ile Pro
35 40 45
Thr Ser Thr Asp Tyr Ile Ile Ala Ala Glu Lys Arg Leu Leu Ser Ile
50 55 60
Leu Lys Thr Pro Tyr Val Gln Glu Gln Val Ser Ile Gly Ser Gly Pro
65 70 75 80
Pro Gly Ser Lys Ile Arg Trp Phe Arg Ser Thr Ser Asn Glu Ser Arg
85 90 95
Tyr Ile Asn Thr Val Thr Phe Asp Ala Lys Glu Gly Ala Pro Thr Leu
100 105 110
Val Met Val His Gly Tyr Gly Ala Ser Gln Gly Phe Phe Phe Arg Asn
115 120 125
Phe Asp Ala Leu Ala Ser Arg Phe Arg Val Ile Ala Ile Asp Gln Leu
130 135 140
Gly Trp Gly Gly Ser Ser Arg Pro Asp Phe Thr Cys Arg Ser Thr Glu
145 150 155 160
Glu Thr Glu Ala Trp Phe Ile Asp Ser Phe Glu Glu Trp Arg Lys Ala
165 170 175
Gln Asn Leu Ser Asn Phe Ile Leu Leu Gly His Ser Phe Gly Gly Tyr
180 185 190
Val Ala Ala Lys Tyr Ala Leu Lys His Pro Glu His Val Gln His Leu
195 200 205
Ile Leu Val Gly Ser Ala Gly Phe Ser Ala Glu Ala Asp Ala Lys Ser
210 215 220
Glu Trp Leu Thr Lys Phe Arg Ala Thr Trp Lys Gly Ala Val Leu Asn
225 230 235 240
His Leu Trp Glu Ser Asn Phe Thr Pro Gln Lys Leu Val Arg Gly Leu
245 250 255
Gly Pro Trp Gly Pro Gly Leu Val Asn Arg Tyr Thr Thr Ala Arg Phe
260 265 270
Gly Ala His Ser Glu Gly Thr Gly Leu Thr Glu Glu Glu Ala Lys Leu
275 280 285
Leu Thr Asp Tyr Val Tyr His Thr Leu Ala Ala Lys Ala Ser Gly Glu
290 295 300

pctau2012001598-seql

Leu Cys Leu Lys Tyr Ile Phe Ser Phe Gly Ala Phe Ala Arg Lys Pro
305 310 315 320

Leu Leu Gln Ser Ala Ser Glu Trp Lys Val Pro Thr Thr Phe Ile Tyr
325 330 335

Gly Met Asn Asp Trp Met Asn Tyr Gln Gly Ala Val Glu Ala Arg Lys
340 345 350

Ser Met Lys Val Pro Cys Glu Ile Ile Arg Val Pro Gln Gly Gly His
355 360 365

Phe Val Phe Ile Asp Asn Pro Ile Gly Phe His Ser Ala Val Leu Tyr
370 375 380

Ala Cys Arg Lys Phe Ile Ser Gln Asp Ser Ser His Asp Gln Gln Leu
385 390 395 400

Leu Asp Gly Leu Arg Leu Val
405

<210> 430

<211> 413

<212> PRT

<213> Brachypodium distachyon

<400> 430

Met Arg Arg Ala Ala Ala Val Thr Arg Met Ala Ala Thr Glu Glu Met
1 5 10 15

Arg Gln Ala Ser Ala Ala Ala Ala Ala Thr Val Thr Glu Ala Ser Ala
20 25 30

Ser Ala Ala Pro Pro Ala Gly Ser Arg Trp Ala Arg Val Trp Pro Ala
35 40 45

Ala Leu Arg Trp Ile Pro Thr Ser Thr Glu Arg Ile Ile Ala Ala Glu
50 55 60

Lys Arg Leu Leu Ser Val Leu Lys Thr Gly Tyr Val Gln Glu Gln Val
65 70 75 80

Asn Ile Gly Ser Ala Pro Pro Gly Ser Lys Val Arg Trp Phe Arg Ser
85 90 95

Ser Ser Asp Glu Pro Arg Phe Ile Asn Thr Val Thr Phe Asp Ser Lys
100 105 110

Glu Asn Ala Pro Thr Leu Val Met Val His Gly Tyr Gly Ala Ser Gln
115 120 125

Gly Phe Phe Phe Arg Asn Phe Asp Ala Leu Ala Ser Arg Phe Arg Val
Page 545

pctau2012001598-seql

130

135

140

Ile Ala Ile Asp Gln Leu Gly Trp Gly Gly Ser Ser Arg Pro Asp Phe
 145 150 155 160

Thr Cys Lys Ser Thr Glu Glu Thr Glu Ala Trp Phe Ile Asp Ser Leu
 165 170 175

Glu Glu Trp Arg Lys Ala Lys Asn Leu Ser Asn Phe Ile Leu Leu Gly
 180 185 190

His Ser Phe Gly Gly Tyr Val Ala Ala Lys Tyr Ala Leu Gln His Pro
 195 200 205

Glu His Val Gln His Leu Ile Leu Val Gly Ser Ala Gly Phe Ser Ser
 210 215 220

Glu Thr Asp His Ser Ser Glu Trp Leu Thr Lys Phe Arg Ala Thr Trp
 225 230 235 240

Lys Gly Met Leu Val Asn Gln Leu Trp Glu Ser Asn Phe Thr Pro Gln
 245 250 255

Arg Ile Val Arg Gly Leu Gly Pro Trp Gly Pro Asp Leu Val Arg Arg
 260 265 270

Tyr Thr Thr Ala Arg Phe Gly Ser Tyr Ser Thr Gly Glu Leu Leu Thr
 275 280 285

Glu His Glu Ser Gly Leu Leu Thr Asp Tyr Ile Tyr His Thr Leu Ala
 290 295 300

Ala Lys Ala Ser Gly Glu Leu Cys Leu Lys Tyr Ile Phe Ser Leu Gly
 305 310 315 320

Ala Phe Ala Arg Lys Pro Leu Leu Gln Ser Ala Ser Asp Trp Lys Val
 325 330 335

Pro Thr Thr Phe Ile Tyr Gly His Asp Asp Trp Met Lys Tyr Gln Gly
 340 345 350

Ala Gln Gln Ala Arg Lys Asp Met Lys Val Pro Cys Glu Ile Ile Arg
 355 360 365

Val Pro Gln Gly Gly His Phe Val Phe Ile Asp Asn Pro Ser Gly Phe
 370 375 380

His Ser Ala Val Phe Tyr Ala Cys Arg Lys Phe Leu Ser Gly Asp Ala
 385 390 395 400

Glu Glu Gly Leu Ser Leu Pro Asp Gly Leu Ile Ser Ala

pctau2012001598-seql

405

410

<210> 431
 <211> 387
 <212> PRT
 <213> Glycine max

<400> 431

Met Ala Glu Glu Ile Thr Lys Asn Asp Val Gly Val Thr Ser Lys Thr
 1 5 10 15

Thr Arg Ser Ser Ser Arg Phe Trp Pro Arg Trp Ile Pro Thr Ser Thr
 20 25 30

Asp His Ile Ile Ala Ala Glu Lys Arg Leu Leu Ser Val Val Lys Thr
 35 40 45

Gly Tyr Val Gln Glu His Val Asn Ile Gly Ser Gly Pro Pro Gly Ser
 50 55 60

Lys Val Arg Trp Phe Arg Ser Ser Ser Asn Glu Pro Arg Phe Ile Asn
 65 70 75 80

Thr Val Thr Phe Asp Ser Lys Pro His Ser Pro Thr Leu Val Met Ile
 85 90 95

His Gly Tyr Ala Ala Ser Gln Gly Phe Phe Phe Arg Asn Phe Asp Ala
 100 105 110

Leu Ala Ser Arg Phe Arg Val Ile Ala Val Asp Gln Leu Gly Trp Gly
 115 120 125

Gly Ser Ser Arg Pro Asp Phe Thr Cys Lys Ser Thr Glu Glu Thr Glu
 130 135 140

Ala Trp Phe Ile Asp Ser Phe Glu Glu Trp Arg Lys Ala Lys Asn Leu
 145 150 155 160

Ser Asn Phe Ile Leu Leu Gly His Ser Phe Gly Gly Tyr Val Ala Ala
 165 170 175

Lys Tyr Ala Leu Lys His Pro Glu His Val Gln His Leu Ile Leu Val
 180 185 190

Gly Ser Ala Gly Phe Ser Ser Glu Ser Asp Ala Lys Ser Glu Trp Ile
 195 200 205

Thr Arg Phe Arg Ala Thr Trp Lys Gly Ala Val Leu Asn His Leu Trp
 210 215 220

Glu Ser Asn Phe Thr Pro Gln Lys Leu Val Arg Gly Leu Gly Pro Trp
 225 230 235 240

pctau2012001598-seql

Gly Pro Asn Ile Val Arg Lys Tyr Thr Ser Ala Arg Phe Gly Thr His
245 250 255

Ser Thr Gly Glu Ile Leu Thr Glu Glu Glu Ser Thr Leu Leu Thr Asp
260 265 270

Tyr Val Tyr His Thr Leu Ala Ala Lys Ala Ser Gly Glu Leu Cys Leu
275 280 285

Lys Tyr Ile Phe Ser Phe Gly Ala Phe Ala Arg Met Pro Leu Leu Leu
290 295 300

Ser Ala Ser Glu Trp Lys Val Pro Thr Thr Phe Met Tyr Gly Phe Gln
305 310 315 320

Asp Trp Met Asn Tyr Gln Gly Ala Gln Glu Ala Arg Lys His Met Lys
325 330 335

Val Pro Cys Glu Ile Ile Arg Ile Pro Gln Gly Gly His Phe Ala Phe
340 345 350

Ile Asp Asn Pro Thr Ala Phe His Ser Ala Val Phe Tyr Ala Cys Arg
355 360 365

Arg Phe Leu Thr Pro Asp Pro Asp Asn Glu Ser Leu Pro Lys Gly Leu
370 375 380

Thr Ser Ala
385

<210> 432
<211> 416
<212> PRT
<213> Zea mays

<400> 432

Met Arg Arg Ala Ala Val Ala Ala Thr Thr Thr Thr Thr Arg Met Ala
1 5 10 15

Ala Glu Glu Met Arg Arg Ala Ser Ala Ser Thr Ala Thr Ala Glu Met
20 25 30

Pro Ala Ser Pro Ala Pro Ala Gln Ala Gly Ser Arg Trp Ala Arg Val
35 40 45

Trp Pro Arg Ala Leu Arg Trp Ile Pro Thr Ser Thr Asp Arg Ile Ile
50 55 60

Ala Ala Glu Lys Arg Leu Leu Thr Ile Val Lys Thr Gly Tyr Val Gln
65 70 75 80

pctau2012001598-seql

Glu Arg Val Asn Ile Gly Ser Ala Pro Pro Gly Ser Lys Val Arg Trp
85 90 95
Phe Arg Ser Ala Ser Asp Glu Pro Arg Phe Ile Asn Thr Val Thr Phe
100 105 110
Asp Ser Lys Glu Asn Ala Pro Thr Leu Val Met Val His Gly Tyr Gly
115 120 125
Ala Ser Gln Gly Phe Phe Phe Arg Asn Phe Asp Ala Leu Ala Ser Arg
130 135 140
Phe Arg Val Ile Ala Ile Asp Gln Leu Gly Trp Gly Gly Ser Ser Arg
145 150 155 160
Pro Asp Phe Thr Cys Lys Ser Thr Glu Glu Thr Glu Ala Trp Phe Ile
165 170 175
Asp Ser Phe Glu Glu Trp Arg Lys Ala Lys Asn Leu Ser Asn Phe Ile
180 185 190
Leu Leu Gly His Ser Phe Gly Gly Tyr Val Ala Ala Lys Tyr Ala Leu
195 200 205
Lys His Pro Glu His Val Gln Gln Leu Ile Leu Val Gly Pro Ala Gly
210 215 220
Phe Ser Ser Glu Thr Glu His Ser Ser Glu Trp Leu Thr Lys Phe Arg
225 230 235 240
Ala Thr Trp Lys Gly Met Leu Met Asn Arg Leu Trp Glu Ser Asn Phe
245 250 255
Thr Pro Gln Arg Val Ile Arg Gly Leu Gly Pro Trp Gly Pro Gly Leu
260 265 270
Val Gln Arg Tyr Thr Ser Ala Arg Phe Gly Thr Ser Ser Thr Gly Glu
275 280 285
Leu Leu Thr Asp Glu Glu Ser Ala Leu Met Thr Asp Tyr Met Tyr His
290 295 300
Thr Leu Ala Ala Lys Ala Ser Gly Glu Leu Cys Leu Lys Tyr Ile Phe
305 310 315 320
Ser Phe Gly Ala Phe Ala Arg Lys Pro Leu Leu Gln Cys Ala Ser Asp
325 330 335
Trp Lys Val Pro Thr Thr Phe Ile Tyr Gly Gln Gln Asp Trp Met Asn
340 345 350

pctau2012001598-seql

Tyr Gln Gly Ala Gln Gln Ala Arg Lys Asp Met Lys Val Pro Cys Glu
355 360 365

Ile Ile Arg Val Pro Gln Gly Gly His Phe Val Phe Ile Asp Asn Pro
370 375 380

Ser Gly Phe His Ser Ala Val Phe Tyr Ala Cys Arg Asn Leu Leu Ser
385 390 395 400

Val Asn Gly Glu Glu Gly Phe Thr Phe Pro Asp Gly Leu Ile Ser Ala
405 410 415

<210> 433

<211> 414

<212> PRT

<213> Sorghum bicolor

<400> 433

Met Arg Leu Ala Ala Val Ala Arg Thr Thr Arg Met Ala Ala Glu Glu
1 5 10 15

Met Arg Arg Ala Ser Ala Ser Ala Ala Val Ala Ala Thr Thr Glu Ala
20 25 30

Ala Pro Ala Pro Ala Gln Ala Gly Ser Arg Trp Ala Arg Val Trp Pro
35 40 45

Ser Ala Leu Arg Trp Ile Pro Thr Ser Thr Asp Arg Ile Ile Ala Ala
50 55 60

Glu Lys Arg Leu Leu Ser Ile Val Lys Thr Gly Tyr Val Gln Glu Gln
65 70 75 80

Val Asn Ile Gly Ser Ala Pro Pro Gly Ser Lys Val Arg Trp Phe Arg
85 90 95

Ser Ala Ser Asp Glu Pro Arg Phe Ile Asn Thr Val Thr Phe Asp Gly
100 105 110

Lys Glu Asn Ala Pro Thr Leu Val Met Val His Gly Tyr Gly Ala Ser
115 120 125

Gln Gly Phe Phe Phe Arg Asn Phe Asp Ala Leu Ala Ser Arg Phe Arg
130 135 140

Val Ile Ala Ile Asp Gln Leu Gly Trp Gly Gly Ser Ser Arg Pro Asp
145 150 155 160

Phe Thr Cys Lys Ser Thr Glu Glu Thr Glu Ala Trp Phe Ile Asp Ser
165 170 175

pctau2012001598-seql

Phe Glu Glu Trp Arg Lys Ala Lys Asn Leu Ser Asn Phe Ile Leu Leu
180 185 190

Gly His Ser Phe Gly Gly Tyr Val Ala Ala Lys Tyr Ala Leu Lys His
195 200 205

Pro Glu His Ile Gln His Leu Val Leu Val Gly Pro Ala Gly Phe Ser
210 215 220

Ser Glu Thr Asp His Ser Ser Glu Trp Leu Thr Lys Phe Arg Ala Thr
225 230 235 240

Trp Lys Gly Met Leu Val Asn His Leu Trp Glu Ser Asn Phe Thr Pro
245 250 255

Gln Arg Val Ile Arg Gly Leu Gly Pro Trp Gly Pro Gly Leu Val Gln
260 265 270

Arg Tyr Thr Ser Ala Arg Phe Gly Thr Arg Ser Thr Gly Asp Ile Leu
275 280 285

Thr Asp Gln Glu Ser Thr Leu Leu Thr Asp Tyr Ile Tyr His Thr Leu
290 295 300

Ala Ala Lys Ala Ser Gly Glu Leu Cys Leu Lys Tyr Ile Phe Ser Phe
305 310 315 320

Gly Ala Phe Ala Arg Lys Pro Leu Leu Gln Cys Ala Ser Asp Trp Lys
325 330 335

Val Pro Thr Thr Phe Ile Tyr Gly Gln Glu Asp Trp Met Asn Tyr Gln
340 345 350

Gly Ala Gln Gln Ala Arg Lys Asp Met Lys Val Pro Cys Glu Ile Ile
355 360 365

Arg Val Pro Gln Ser Gly His Phe Val Phe Ile Asp Asn Pro Ser Gly
370 375 380

Phe His Ser Ala Val Phe Tyr Ala Cys Arg Asn Leu Leu Ser Gln Asn
385 390 395 400

Gly Glu Glu Gly Phe Thr Phe Pro Gly Gly Leu Ile Ser Ala
405 410

<210> 434
<211> 413
<212> PRT
<213> Ricinus communis

<400> 434

Met Lys Leu Leu Arg Leu Arg Leu Leu Arg Val Ser Ser Ser Ser Ile
Page 551

pctau2012001598-seql

```

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

```

275

280

285

Lys Glu Glu Glu Ser Lys Leu Leu Thr Asp Tyr Val Tyr His Thr Leu
 290 295 300

Ala Ala Lys Ala Ser Gly Glu Leu Cys Leu Lys Tyr Ile Phe Ser Phe
 305 310 315 320

Gly Ala Phe Ala Arg Met Pro Leu Leu Gln Ser Ala Ser Gln Trp Lys
 325 330 335

Val Pro Thr Thr Phe Ile Tyr Gly Met Gln Asp Trp Met Asn Tyr Gln
 340 345 350

Gly Ala Gln Arg Ala Arg Lys Asp Met Asn Val Pro Cys Glu Ile Ile
 355 360 365

Arg Val Pro Gln Gly Gly His Phe Val Phe Ile Asp Asn Pro Thr Gly
 370 375 380

Phe His Ser Ala Val Leu Tyr Ala Cys Arg Arg Phe Leu Ser Pro Asp
 385 390 395 400

Pro Asp Asn Glu Ser Leu Pro Glu Gly Leu Ile Ser Ala
 405 410

<210> 435

<211> 386

<212> PRT

<213> Medicago truncatula

<400> 435

Met Ala Glu Glu Ile Arg Gln Lys Asp Asp Val Asp Ser Ser Ser Lys
 1 5 10 15

Ser Lys Ser Phe Trp Ser Ser Leu Arg Trp Ile Pro Thr Ser Thr Asp
 20 25 30

His Ile Ile Ala Ala Glu Lys Arg Leu Leu Ser Ile Ile Lys Thr Gly
 35 40 45

Tyr Ala Gln Glu His Val Asn Ile Gly Ser Gly Pro Pro Gly Ser Lys
 50 55 60

Val Arg Trp Phe Arg Ser Thr Ser Asn Glu Pro Arg Phe Leu Asn Thr
 65 70 75 80

Val Thr Phe Asp Ser Lys Pro Asp Ser Pro Thr Leu Val Met Val His
 85 90 95

Gly Tyr Ala Ala Ser Gln Gly Phe Phe Phe Arg Asn Phe Asp Ala Leu
 100 105 110

pctau2012001598-seql

Ala Ser Arg Phe Arg Ile Ile Ala Val Asp Gln Leu Gly Trp Gly Gly
 115 120 125
 Ser Ser Arg Pro Asp Phe Thr Cys Lys Ser Thr Glu Glu Thr Glu Ala
 130 135 140
 Trp Phe Ile Asp Ser Phe Glu Glu Trp Arg Lys Ala Lys Asn Leu Thr
 145 150 155 160
 Asn Phe Ile Leu Leu Gly His Ser Phe Gly Gly Tyr Val Ala Ser Lys
 165 170 175
 Tyr Ala Leu Lys His Pro Gln His Val Gln His Leu Ile Leu Val Gly
 180 185 190
 Pro Ala Gly Phe Thr Glu Glu Thr Asp Pro Lys Thr Glu Phe Val Thr
 195 200 205
 Lys Phe Arg Ala Thr Trp Lys Gly Ala Val Leu Asn His Leu Trp Glu
 210 215 220
 Ser Asn Phe Thr Pro Gln Lys Ile Val Arg Gly Leu Gly Pro Trp Gly
 225 230 235 240
 Pro Asn Met Val Arg Lys Tyr Thr Ser Ala Arg Phe Gly Thr His Ser
 245 250 255
 Thr Gly Gln Lys Leu Ile Asp Glu Glu Ser Ser Leu Leu Thr Asp Tyr
 260 265 270
 Val Tyr His Thr Leu Ala Ala Lys Ala Ser Gly Glu Leu Cys Leu Lys
 275 280 285
 Tyr Ile Phe Ala Phe Gly Ala Phe Ala Arg Met Pro Leu Leu Gln Ser
 290 295 300
 Ala Gln Glu Trp Lys Val Pro Thr Thr Phe Ile Tyr Gly Tyr Glu Asp
 305 310 315 320
 Trp Met Asn Tyr Glu Gly Ala Gln Glu Ala Arg Lys His Met Lys Val
 325 330 335
 Pro Cys Glu Ile Ile Arg Val Pro Lys Ala Gly His Phe Val Phe Ile
 340 345 350
 Asp Asn Pro Ser Gly Phe His Ser Ala Val Phe Tyr Ala Cys Arg Arg
 355 360 365
 Phe Leu Thr Pro Asn Ser Asp Asn Glu Ser Leu Pro Glu Gly Leu Ser
 370 375 380

pctau2012001598-seql

Ser Ala
385

<210> 436
<211> 371
<212> PRT
<213> Oryza sativa
<400> 436

Met Ala Ala Glu Gly Met Ser Thr Ala Ala Ala Ala Glu Ala Thr
1 5 10 15

Ala Thr Ala Ala Pro Ala Ala Gly Ser Arg Trp Gly Arg Ala Trp Pro
20 25 30

Ser Ala Leu Arg Trp Ile Pro Thr Ser Thr Asp Arg Ile Ile Ala Ala
35 40 45

Glu Lys Arg Leu Leu Ser Ile Val Lys Thr Gly Tyr Val Gln Glu Gln
50 55 60

Val Asn Ile Gly Ser Ser Pro Pro Gly Ser Lys Val Arg Trp Phe Arg
65 70 75 80

Ser Ser Ser Asp Glu Pro Arg Phe Ile Asn Thr Val Thr Phe Asp Ser
85 90 95

Glu Glu Asn Ala Pro Thr Leu Val Met Val His Gly Tyr Gly Ala Ser
100 105 110

Gln Gly Phe Phe Phe Arg Asn Phe Asp Ala Leu Ala Ser Arg Phe Arg
115 120 125

Val Ile Ala Ile Asp Gln Leu Gly Trp Gly Gly Ser Ser Arg Pro Asp
130 135 140

Phe Thr Cys Lys Ser Thr Glu Glu Thr Glu Ala Trp Phe Ile Asp Ser
145 150 155 160

Phe Glu Glu Trp Arg Lys Ala Lys Asn Leu Ser Asn Phe Ile Leu Leu
165 170 175

Gly His Ser Phe Gly Gly Tyr Val Ala Ala Lys Tyr Ala Leu Gln His
180 185 190

Pro Glu His Val Gln His Leu Ile Leu Val Gly Pro Ala Gly Phe Ser
195 200 205

Ser Glu Thr Glu His Ser Ser Glu Trp Leu Thr Lys Phe Arg Ala Thr
210 215 220

pctau2012001598-seql

Trp Lys Gly Met Leu Val Asn His Leu Trp Glu Ser Asn Phe Thr Pro
225 230 235 240

Gln Arg Ile Val Arg Gly Leu Gly Pro Trp Gly Pro Gly Leu Val Gln
245 250 255

Arg Tyr Thr Ser Ala Arg Phe Gly Ser His Ser Thr Gly Glu Leu Leu
260 265 270

Thr Glu Gln Glu Ser Thr Leu Leu Thr Asp Tyr Ile Tyr His Thr Leu
275 280 285

Ala Ala Lys Ala Ser Gly Glu Leu Cys Leu Lys His Ile Phe Ser Phe
290 295 300

Gly Ala Phe Ala Arg Lys Pro Leu Leu Gln Arg Ser Tyr Glu His Pro
305 310 315 320

Gly Asp Leu Glu Ser Ile Cys Thr Ile Gly Gly His Phe Val Phe Ile
325 330 335

Asp Asn Pro Ser Gly Phe His Ser Ala Val Phe His Ala Cys Arg Lys
340 345 350

Phe Leu Ser Gly Asp Gly Glu Glu Gly Leu Ser Leu Pro Glu Gly Leu
355 360 365

Thr Ser Ala
370

<210> 437
<211> 1317
<212> DNA
<213> Arabidopsis thaliana

<400> 437
gatctctctc cctctctctc tctctctctc cgaggaaaat ggataacttc ttaccctttc 60
cctcttctaa cgcaaactct gtccaagaac tctctatgga tcctaacaac aatcgctcgc 120
acttcacaac agtccctact tatgatcatc atcaggctca gcctcatcac ttcttgccctc 180
cgttttcata cccggtggag cagatggcgg cggtgatgaa tcctcagccg gtttacttat 240
cggagtgtta tcctcagatc ccggttacgc aaaccggaag tgaattcggg tctctggttg 300
gtaatccttg tttgtggcaa gagagaggtg gttttcttga tccgcgatg acgaagatgg 360
caaggatcaa caggaaaaac gccatgatga gatcaagaaa caactctagc cctaattcta 420
gtccaagtga gttgggttgat tcaaagagac agctgatgat gcttaacttg aaaaataacg 480
tgcagatctc cgacaagaaa gatagctacc aacagtccac atttgataac aagaagctta 540
gggttttgtg tgagaaggaa ttgaagaaca gcgatgttgg gtcactcggg aggatagttc 600
taccaaagag agatgcagaa gcaaattctc cgaagctatc tgataaagaa ggaatcgttg 660

pctau2012001598-seq1

tacagatgag agatgttttc tctatgcagt cttggtcttt caaatacaag ttttgggtcca	720
ataacaagag cagaatgtat gtcctcgaga acacaggaga atttgtgaag caaaatggag	780
ctgagatagg agacttttta acaatatacg aggacgaaag caagaatctc tacttcgcca	840
tgaatggaaa ttcgggaaaa caaatgaag gaagagaaaa tgagtcgagg gaaaggaacc	900
actacgaaga ggcaatgctt gattacatac caagagacga agaggaagct tccattgcaa	960
tgctcatcgg aaatctaaac gatcactatc ccatccctaa cgatctcatg gacctcacca	1020
ctgaccttca gcaccatcaa gccacgtcct catcaatgcc acctgaggat cacgcgtacg	1080
tgggttcacg cgatgatcag gtgagcttta acgactttga gtgggtggtga tatgggtggtg	1140
gaagtcttca agttcataac cccctttatg aaaatagacc ttaagatata caaaagagat	1200
taaaagaaaa aaaagttagt atatttcacg atatctctca ttgaagatga gatttatatc	1260
tataattggt taatagtgtt tttattactt ttctatcaat atattaaagt ttttaatt	1317

<210> 438
 <211> 1439
 <212> DNA
 <213> *Medicago truncatula*

<400> 438	
tttcatcctt acatattttg catattgaaa cacgtaggat ggaataagat tgataacaaa	60
aattgcattg tttgcatatt gaaaacatgg gacaattgca tgggttcacg tgcttcatta	120
taagccacac attaggaaac acaggttgat attcaccact atttaacata agaatatctc	180
atgtgtaagc attcatacaa ataccacaat tgaattaaaa accaaagaaa tgtcttcctc	240
taacttctct tgtatcctat ccatctcctt aacattcttc atcttgctac tgaacaagg	300
gaattcagca gaaacaactt ccttttccat cacaaaattt gtcccagatc aaaagaatct	360
catcttccaa ggcgatgcga aaactgcctc aacaggaag ttagaactct ccaaggcagt	420
caagaactct attggttagag ctctttattc cgcccctatt cacatttggg atagcaaaac	480
cggtagtggt gctaactttc aaactacctt cacctttaca ataacggcgc ctaatactta	540
taatgttgca gacggtcttg cattcttcat tgcaccaatt gatactaagc cgaaatcaat	600
tcacatgga ggataccttg gagttttcga tagcaaaact tacaaaaaat caattcaaac	660
tggtgcagtt gaaattgaca ctttctataa tgctcaatgg gatccaaatc ccggaaatat	720
aagtagcact ggtcgacata ttggaatcga tgtaaactct atcaaatcaa taagcaccgt	780
gccgtggagt ttggaaaaca ataaaaaggc taatgttgcg atagggttta atggtgcaac	840
aaatgtgttg agtgttgatg tggaatatcc tttgattcgt cattataccc taagtcatgt	900
tgtgcctttg aaggatgttg ttcctgagtg ggtaaggatt ggtttctctt cttctactgg	960
agccgaatat tcagcacatg atattttatc gtgggtctttt gattcaaagt tgaacctagg	1020
ttttgagaac aatatcaatg ccaatgtttc aagctctact caagctgcat agttgaaaac	1080
ttatccatta tgtatgtgtg agtgtaacca accagtctaa gaaaactata ataagatacc	1140

pctau2012001598-seq1

tgaaataatg gttcattatc gtgtagtaga aatatggtca caccatatct tctttttttt	1200
ttaataaatt atggaataat gctattttctc gcgagagtta tgtttcggaa agattcatga	1260
atagatgtta atcaattaga tctatatata tatatatata tatatatata tatatatata	1320
tagcattttc ttaaattatg catatgtaat atcgtgtaat gctattgttt atatcaatga	1380
atgggtgtttt gtagtcacat aattcgtaat ttctctccat gagaacagcg aaccaatta	1439

<210> 439
 <211> 1393
 <212> DNA
 <213> Brassica napus

<400> 439	
gagatgggta tccctataag gtgcagcatc gaaccatctg caacattttg actcgttttc	60
ttttgtgttt ttataacatc tgtctcttct tctctgatc tccctctctt ttctttttca	120
atctcccca cgaacctccc ttcataactc tctttctctc cccgggaaat atggataact	180
tcttgccctt ttcctcttct aacgcaaact ctgtccaaga actctccatg gatcttaaca	240
agaatcgctc gcacttctcc atggcgcagc ctcagcactt gttgccgcct tactcgtacg	300
ttgcatgtcc ggcacttgat cagacgggga ccatgaatca tcagcctctt cactcatcgg	360
atgcttttcc tcagatcccg gttgtacaaa ccggagggtga attcggctat ttggtttgta	420
agcccgggtg gaggcaggaa cgagggtgat ttcttgatcc acactccact aagatggcta	480
ggatcaacag gaagaaggcg atgctaagat caagaaacaa ctctaaccct aattctagtt	540
cgaatgagtt ggttgattca aggagacaag tggctcttac catgaaaaat aatgccgaga	600
ttgctgctag aaaagatttt tatcgattct cctcattcga taacaagaaa cttagggttt	660
tgttggtgaa gcacttgaag aacagcgatg ttgggtcact tggaaggatt gttctaccaa	720
agagagaagc agaaggaaat cttccggagc tatctgataa agaaggaatg gtattagaga	780
tgagagatgt tgactctgtg cagtcttggt ctttcaaata caagtactgg tccaataaca	840
agagcagaat gtatgtcctc gaaaacacag gagaatttgt gaagaaaaat ggagtattga	900
tgggagacta tctaacaatc tacgaggacg aaagcaagaa tctctacttc tccatcagaa	960
agcaccaca caaacaaaat gatggaagag aggatgagtc gatggaagtt atcgagatga	1020
acttctatga agatataatg tttgattaca taccaaatga tgaagacgat tccattgcaa	1080
tgctcctcgg aaatctaaac gagcactatc cctacccaaa tgatcttatg gatctcactg	1140
tcaatcttga tcagcatcag caagccacct cctcgctgcc acctgctgat cacatgagct	1200
cgaacgattt cttatggtga tgtgatggac gttgatatgg attccctttg agatgatata	1260
caagggatga aaagaaaaga gtatcatatt catatccata tttgtttgat aaaatgtgtt	1320
tgttccaat ctattattta tgaaaaactt atttgtgttt aactccagat taattaaata	1380
tttttcattt gac	1393

<210> 440
 <211> 1755

pctau2012001598-seq1

<212> DNA

<213> Arabidopsis thaliana

<400> 440

```

atgaactcga tgaataactg gttaggcttc tctctctctc ctcatgatca aaatcatcac      60
cgtacggatg ttgactcctc caccaccaga accgccgtag atgttgccgg aggggtactgt      120
tttgatctgg ccgctccctc cgatgaatct tctgccgttc aaacatcttt tctttctcct      180
ttcgggtgtca ccctcgaagc tttcaccaga gacaataata gtcactcccg agattgggac      240
atcaatggtg gtgcatgcaa taacattaac aataacgaac aaaatggacc aaagcttgag      300
aatttcctcg gccgcaccac cacgatttac aataccaacg agaccgttgt agatggaaat      360
ggcgattgtg gaggaggaga cgggtggtgg ggcggtctac taggcctttc gatgataaaa      420
acatggctga gtaatcattc ggttgcta atgctaatc aagacaatgg taacggtgca      480
cgaggcttgt ccctctctat gaattcatct actagtata gcaacaacta caacaacaat      540
gatgatgtcg tccaagagaa gactattgtt gatgtcgtag aaactacacc gaagaaaact      600
attgagagtt ttggacaaag gacgtctata taccgcggtg ttacaaggca tcggtggaca      660
ggtagatacg aggcacatth atgggacaat agttgcaaaa gagaaggcca gactcgcaaa      720
ggaagacaag tttatctggg aggttatgac aaagaagaaa aagcagctag ggcttacgat      780
ttagccgcac taaagtattg gggaaccacc actactacta acttcccctt gagtgaatat      840
gagaaagagg tagaagagat gaagcacatg acgaggcaag agtatgttgc ctctctgcgc      900
aggaaaagta gtggtttctc tcgtggtgca tcgatttatc gaggagtaac aaggcatcac      960
caacatggaa ggtggcaagc taggatcgga agagtcgccg gtaacaaaga cctctacttg     1020
ggaactttcg gcacacagga agaggctgct gaggcttatg acattgcagc cattaaattc     1080
agaggattaa gcgcagtgac taacttcgac atgaacagat acaatgttaa agcaatcctc     1140
gagagcccga gtctacctat tggtagttct gcgaaacgtc tcaaggacgt taataatccg     1200
gttccagcta tgatgattag taataacgtt tcagagagtg caaataatgt tagcggttgg     1260
caaaacactg cgtttcagca tcatcaggga atggatttga gcttattgca gcaacagcag     1320
gagaggtacg ttggttatta caatggagga aacttgtcta ccgagagtac tagggtttgt     1380
ttcaaacaag aggaggaaca acaacacttc ttgagaaact cgccgagtca catgactaat     1440
gttgatcatc atagctcgac ctctgatgat tctgttaccg tttgtggaaa tggtgttagt     1500
tatggtggtt atcaaggatt cgcaatccct gttggaacat cggttaatta cgatcccttt     1560
actgctgctg agattgctta caacgcaaga aatcattatt actatgctca gcatcagcaa     1620
caacagcaga ttcagcagtc gccgggagga gatthttccg tggcgatttc gaataaccat     1680
agctctaaca tgtactttca cggggaaggt ggtggagaag gggctccaac gttttcagtt     1740
tggaacgaca cttag                                           1755

```

<210> 441

<211> 2061

<212> DNA

<213> *Medicago truncatula*

<400> 441

atgaacttgt taggtttctc tctatctcca caagaacaac atccatcaac acaagatcaa	60
acggtggcct cccgttttgg gttcaaccct aatgaaatct caggctctga tgttcaagga	120
gatcactgct atgatctctc ttctcacaca actcctcatc attcactcaa cttttctcat	180
cctttttcca tttatgaagc tttccacaca aataacaaca ttcacaccac tcaagattgg	240
aaggagaact acaacaacca aaacctacta ttgggaacat catgcatgaa ccaaaatgtg	300
aacaacaaca accaacaagc acaaccaaag ctagaaaact tcctcgggtg acactctttc	360
accgaccatc aagaatacgg tggtagcaac tcatactctt cattacacct cccacctcat	420
cagccggaag catcctgtgg cgggtggtgat ggtagtacia gtaacaataa ctcaataggt	480
ttatctatga taaaaacatg gctcagaaac caaccaccac caccagaaaa caacaacaat	540
aacaacaatg aaagtgtgac acgtgtgcag acactatcac tttctatgag tactggctca	600
cagtcaagtt catctgtgcc tcttctcaat gcaaatgtga tgagtgggtga gatttcctca	660
tcggaaaaca aacaaccacc cacaactgca gttgtacttg atagcaacca aacaagtgtc	720
gttgaaagtg ctgtgcctag aaaatccggt gatacatttg gacaaagaac ttccattttac	780
cgtggtgtaa caaggcatag atggacaggg agatatgaag ctcacctttg ggataatagt	840
tgtagaagag aggggcagac tcgcaaagga aggcaagttt acttgggagg ttatgacaaa	900
gaagaaaaag cagctagagc ctatgatttg gcagcactaa aatattgggg aacaactact	960
acaacaaatt ttccaattag ccattatgaa aaagaagtgg aagaaatgaa gcatatgaca	1020
aggcaagagt acgttgctgc attgagaagg aaaagtagtg gtttttcacg aggtgcatcc	1080
atttaccgag gagtaacaag acatcatcaa catggtagat ggcaagctag gattggaaga	1140
gttgcaggca acaagatct ctacctagga actttcagca ctcaagaaga ggcagcagag	1200
gcatatgatg tggcagcaat aaaattcaga ggactgagtg cagttacaaa ctttgacatg	1260
agcagatatg atgtcaaaac catacttgag agcagcacat taccaattgg tgggtgctgca	1320
aagcgtttta aagacatgga gcaagttgaa ttgaatcatg tgaatgttga tattagccat	1380
agaactgaac aagatcatag catcatcaac aacacttccc atttaacaga acaagccatc	1440
tatgcagcaa caaatgcatc taattggcat gcactttcat tccaacatca acaaccacat	1500
catcattaca atgccaacaa catgcagtta cagaattatc cttatggaac tcaaactcaa	1560
aagcttttgt gcaacaaga acaagattct gatgatcata gtacttatac tactgtctact	1620
gatattcatc aactacagtt aggaataat aataacaata ctcacaattt ctttggttta	1680
caaaatatca tgagtatgga ttctgcttcc atggataata gttctggatc taattctgtt	1740
gtttatggtg gtggagatca tgggtggttat ggaggaaatg gtggatatat gattccaatg	1800
gctattgcaa atgatggtaa ccaaaatcca agaagcaaca acaatttttg tgagagttag	1860
attaaaggat ttggttatga aaatgttttt gggactacta ctgacctta tcatgcacag	1920
gcagcaagga acttgtacta tcagccacaa caattatctg ttgatcaagg atcaaattgg	1980

pctau2012001598-seql

gttccaactg ctattccaac acttgctcca aggactacca atgtctctct atgtcctcct 2040
 ttcactttgt tgcataaata g 2061

<210> 442
 <211> 363
 <212> PRT
 <213> Arabidopsis thaliana

<400> 442

Met Asp Asn Phe Leu Pro Phe Pro Ser Ser Asn Ala Asn Ser Val Gln
 1 5 10 15

Glu Leu Ser Met Asp Pro Asn Asn Asn Arg Ser His Phe Thr Thr Val
 20 25 30

Pro Thr Tyr Asp His His Gln Ala Gln Pro His His Phe Leu Pro Pro
 35 40 45

Phe Ser Tyr Pro Val Glu Gln Met Ala Ala Val Met Asn Pro Gln Pro
 50 55 60

Val Tyr Leu Ser Glu Cys Tyr Pro Gln Ile Pro Val Thr Gln Thr Gly
 65 70 75 80

Ser Glu Phe Gly Ser Leu Val Gly Asn Pro Cys Leu Trp Gln Glu Arg
 85 90 95

Gly Gly Phe Leu Asp Pro Arg Met Thr Lys Met Ala Arg Ile Asn Arg
 100 105 110

Lys Asn Ala Met Met Arg Ser Arg Asn Asn Ser Ser Pro Asn Ser Ser
 115 120 125

Pro Ser Glu Leu Val Asp Ser Lys Arg Gln Leu Met Met Leu Asn Leu
 130 135 140

Lys Asn Asn Val Gln Ile Ser Asp Lys Lys Asp Ser Tyr Gln Gln Ser
 145 150 155 160

Thr Phe Asp Asn Lys Lys Leu Arg Val Leu Cys Glu Lys Glu Leu Lys
 165 170 175

Asn Ser Asp Val Gly Ser Leu Gly Arg Ile Val Leu Pro Lys Arg Asp
 180 185 190

Ala Glu Ala Asn Leu Pro Lys Leu Ser Asp Lys Glu Gly Ile Val Val
 195 200 205

Gln Met Arg Asp Val Phe Ser Met Gln Ser Trp Ser Phe Lys Tyr Lys
 210 215 220

pctau2012001598-seql

Phe Trp Ser Asn Asn Lys Ser Arg Met Tyr Val Leu Glu Asn Thr Gly
225 230 235 240

Glu Phe Val Lys Gln Asn Gly Ala Glu Ile Gly Asp Phe Leu Thr Ile
245 250 255

Tyr Glu Asp Glu Ser Lys Asn Leu Tyr Phe Ala Met Asn Gly Asn Ser
260 265 270

Gly Lys Gln Asn Glu Gly Arg Glu Asn Glu Ser Arg Glu Arg Asn His
275 280 285

Tyr Glu Glu Ala Met Leu Asp Tyr Ile Pro Arg Asp Glu Glu Glu Ala
290 295 300

Ser Ile Ala Met Leu Ile Gly Asn Leu Asn Asp His Tyr Pro Ile Pro
305 310 315 320

Asn Asp Leu Met Asp Leu Thr Thr Asp Leu Gln His His Gln Ala Thr
325 330 335

Ser Ser Ser Met Pro Pro Glu Asp His Ala Tyr Val Gly Ser Ser Asp
340 345 350

Asp Gln Val Ser Phe Asn Asp Phe Glu Trp Trp
355 360

<210> 443
<211> 280
<212> PRT
<213> Medicago truncatula

<400> 443

Met Ser Ser Ser Asn Phe Ser Cys Ile Leu Ser Ile Ser Leu Thr Phe
1 5 10 15

Phe Ile Leu Leu Leu Asn Lys Val Asn Ser Ala Glu Thr Thr Ser Phe
20 25 30

Ser Ile Thr Lys Phe Val Pro Asp Gln Lys Asn Leu Ile Phe Gln Gly
35 40 45

Asp Ala Lys Thr Ala Ser Thr Gly Lys Leu Glu Leu Ser Lys Ala Val
50 55 60

Lys Asn Ser Ile Gly Arg Ala Leu Tyr Ser Ala Pro Ile His Ile Trp
65 70 75 80

Asp Ser Lys Thr Gly Ser Val Ala Asn Phe Gln Thr Thr Phe Thr Phe
85 90 95

pctau2012001598-seql

Thr Ile Thr Ala Pro Asn Thr Tyr Asn Val Ala Asp Gly Leu Ala Phe
100 105 110

Phe Ile Ala Pro Ile Asp Thr Lys Pro Lys Ser Ile His His Gly Gly
115 120 125

Tyr Leu Gly Val Phe Asp Ser Lys Thr Tyr Lys Lys Ser Ile Gln Thr
130 135 140

Val Ala Val Glu Ile Asp Thr Phe Tyr Asn Ala Gln Trp Asp Pro Asn
145 150 155 160

Pro Gly Asn Ile Ser Ser Thr Gly Arg His Ile Gly Ile Asp Val Asn
165 170 175

Ser Ile Lys Ser Ile Ser Thr Val Pro Trp Ser Leu Glu Asn Asn Lys
180 185 190

Lys Ala Asn Val Ala Ile Gly Phe Asn Gly Ala Thr Asn Val Leu Ser
195 200 205

Val Asp Val Glu Tyr Pro Leu Ile Arg His Tyr Thr Leu Ser His Val
210 215 220

Val Pro Leu Lys Asp Val Val Pro Glu Trp Val Arg Ile Gly Phe Ser
225 230 235 240

Ser Ser Thr Gly Ala Glu Tyr Ser Ala His Asp Ile Leu Ser Trp Ser
245 250 255

Phe Asp Ser Lys Leu Asn Leu Gly Phe Glu Asn Asn Ile Asn Ala Asn
260 265 270

Val Ser Ser Ser Thr Gln Ala Ala
275 280

<210> 444
<211> 349
<212> PRT
<213> Brassica napus

<400> 444

Met Asp Asn Phe Leu Pro Phe Ser Ser Ser Asn Ala Asn Ser Val Gln
1 5 10 15

Glu Leu Ser Met Asp Leu Asn Lys Asn Arg Ser His Phe Ser Met Ala
20 25 30

Gln Pro Gln His Leu Leu Pro Pro Tyr Ser Tyr Val Ala Cys Pro Ala
35 40 45

Leu Asp Gln Thr Gly Thr Met Asn His Gln Pro Leu His Ser Ser Asp
Page 563

pctau2012001598-seql

50

55

60

Ala Phe Pro Gln Ile Pro Val Val Gln Thr Gly Gly Glu Phe Gly Tyr
65 70 75 80

Leu Val Cys Lys Pro Gly Val Arg Gln Glu Arg Gly Gly Phe Leu Asp
85 90 95

Pro His Ser Thr Lys Met Ala Arg Ile Asn Arg Lys Lys Ala Met Leu
100 105 110

Arg Ser Arg Asn Asn Ser Asn Pro Asn Ser Ser Ser Asn Glu Leu Val
115 120 125

Asp Ser Arg Arg Gln Val Ala Leu Thr Met Lys Asn Asn Ala Glu Ile
130 135 140

Ala Ala Arg Lys Asp Phe Tyr Arg Phe Ser Ser Phe Asp Asn Lys Lys
145 150 155 160

Leu Arg Val Leu Leu Val Lys His Leu Lys Asn Ser Asp Val Gly Ser
165 170 175

Leu Gly Arg Ile Val Leu Pro Lys Arg Glu Ala Glu Gly Asn Leu Pro
180 185 190

Glu Leu Ser Asp Lys Glu Gly Met Val Leu Glu Met Arg Asp Val Asp
195 200 205

Ser Val Gln Ser Trp Ser Phe Lys Tyr Lys Tyr Trp Ser Asn Asn Lys
210 215 220

Ser Arg Met Tyr Val Leu Glu Asn Thr Gly Glu Phe Val Lys Lys Asn
225 230 235 240

Gly Val Leu Met Gly Asp Tyr Leu Thr Ile Tyr Glu Asp Glu Ser Lys
245 250 255

Asn Leu Tyr Phe Ser Ile Arg Lys His Pro His Lys Gln Asn Asp Gly
260 265 270

Arg Glu Asp Glu Ser Met Glu Val Ile Glu Met Asn Phe Tyr Glu Asp
275 280 285

Ile Met Phe Asp Tyr Ile Pro Asn Asp Glu Asp Asp Ser Ile Ala Met
290 295 300

Leu Leu Gly Asn Leu Asn Glu His Tyr Pro Tyr Pro Asn Asp Leu Met
305 310 315 320

Asp Leu Thr Val Asn Leu Asp Gln His Gln Gln Ala Thr Ser Ser Ser
Page 564

Pro Pro Ala Asp His Met Ser Ser Asn Asp Phe Leu Trp
 340 345

<210> 445
 <211> 584
 <212> PRT
 <213> Arabidopsis thaliana

<400> 445

Met Asn Ser Met Asn Asn Trp Leu Gly Phe Ser Leu Ser Pro His Asp
 1 5 10 15

Gln Asn His His Arg Thr Asp Val Asp Ser Ser Thr Thr Arg Thr Ala
 20 25 30

Val Asp Val Ala Gly Gly Tyr Cys Phe Asp Leu Ala Ala Pro Ser Asp
 35 40 45

Glu Ser Ser Ala Val Gln Thr Ser Phe Leu Ser Pro Phe Gly Val Thr
 50 55 60

Leu Glu Ala Phe Thr Arg Asp Asn Asn Ser His Ser Arg Asp Trp Asp
 65 70 75 80

Ile Asn Gly Gly Ala Cys Asn Asn Ile Asn Asn Asn Glu Gln Asn Gly
 85 90 95

Pro Lys Leu Glu Asn Phe Leu Gly Arg Thr Thr Thr Ile Tyr Asn Thr
 100 105 110

Asn Glu Thr Val Val Asp Gly Asn Gly Asp Cys Gly Gly Gly Asp Gly
 115 120 125

Gly Gly Gly Gly Ser Leu Gly Leu Ser Met Ile Lys Thr Trp Leu Ser
 130 135 140

Asn His Ser Val Ala Asn Ala Asn His Gln Asp Asn Gly Asn Gly Ala
 145 150 155 160

Arg Gly Leu Ser Leu Ser Met Asn Ser Ser Thr Ser Asp Ser Asn Asn
 165 170 175

Tyr Asn Asn Asn Asp Asp Val Val Gln Glu Lys Thr Ile Val Asp Val
 180 185 190

Val Glu Thr Thr Pro Lys Lys Thr Ile Glu Ser Phe Gly Gln Arg Thr
 195 200 205

Ser Ile Tyr Arg Gly Val Thr Arg His Arg Trp Thr Gly Arg Tyr Glu
 210 215 220

pctau2012001598-seql

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

pctau2012001598-seql

Asn Val Val Ser Tyr Gly Gly Tyr Gln Gly Phe Ala Ile Pro Val Gly
500 505 510

Thr Ser Val Asn Tyr Asp Pro Phe Thr Ala Ala Glu Ile Ala Tyr Asn
515 520 525

Ala Arg Asn His Tyr Tyr Tyr Ala Gln His Gln Gln Gln Gln Ile
530 535 540

Gln Gln Ser Pro Gly Gly Asp Phe Pro Val Ala Ile Ser Asn Asn His
545 550 555 560

Ser Ser Asn Met Tyr Phe His Gly Glu Gly Gly Glu Gly Ala Pro
565 570 575

Thr Phe Ser Val Trp Asn Asp Thr
580

<210> 446
<211> 686
<212> PRT
<213> Medicago truncatula

<400> 446

Met Asn Leu Leu Gly Phe Ser Leu Ser Pro Gln Glu Gln His Pro Ser
1 5 10 15

Thr Gln Asp Gln Thr Val Ala Ser Arg Phe Gly Phe Asn Pro Asn Glu
20 25 30

Ile Ser Gly Ser Asp Val Gln Gly Asp His Cys Tyr Asp Leu Ser Ser
35 40 45

His Thr Thr Pro His His Ser Leu Asn Leu Ser His Pro Phe Ser Ile
50 55 60

Tyr Glu Ala Phe His Thr Asn Asn Asn Ile His Thr Thr Gln Asp Trp
65 70 75 80

Lys Glu Asn Tyr Asn Asn Gln Asn Leu Leu Leu Gly Thr Ser Cys Met
85 90 95

Asn Gln Asn Val Asn Asn Asn Asn Gln Gln Ala Gln Pro Lys Leu Glu
100 105 110

Asn Phe Leu Gly Gly His Ser Phe Thr Asp His Gln Glu Tyr Gly Gly
115 120 125

Ser Asn Ser Tyr Ser Ser Leu His Leu Pro Pro His Gln Pro Glu Ala
130 135 140

pctau2012001598-seql

Ser Cys Gly Gly Gly Asp Gly Ser Thr Ser Asn Asn Asn Ser Ile Gly
 145 150 155 160
 Leu Ser Met Ile Lys Thr Trp Leu Arg Asn Gln Pro Pro Pro Pro Glu
 165 170 175
 Asn Asn Asn Asn Asn Asn Asn Glu Ser Gly Ala Arg Val Gln Thr Leu
 180 185 190
 Ser Leu Ser Met Ser Thr Gly Ser Gln Ser Ser Ser Ser Val Pro Leu
 195 200 205
 Leu Asn Ala Asn Val Met Ser Gly Glu Ile Ser Ser Ser Glu Asn Lys
 210 215 220
 Gln Pro Pro Thr Thr Ala Val Val Leu Asp Ser Asn Gln Thr Ser Val
 225 230 235 240
 Val Glu Ser Ala Val Pro Arg Lys Ser Val Asp Thr Phe Gly Gln Arg
 245 250 255
 Thr Ser Ile Tyr Arg Gly Val Thr Arg His Arg Trp Thr Gly Arg Tyr
 260 265 270
 Glu Ala His Leu Trp Asp Asn Ser Cys Arg Arg Glu Gly Gln Thr Arg
 275 280 285
 Lys Gly Arg Gln Val Tyr Leu Gly Gly Tyr Asp Lys Glu Glu Lys Ala
 290 295 300
 Ala Arg Ala Tyr Asp Leu Ala Ala Leu Lys Tyr Trp Gly Thr Thr Thr
 305 310 315 320
 Thr Thr Asn Phe Pro Ile Ser His Tyr Glu Lys Glu Val Glu Glu Met
 325 330 335
 Lys His Met Thr Arg Gln Glu Tyr Val Ala Ser Leu Arg Arg Lys Ser
 340 345 350
 Ser Gly Phe Ser Arg Gly Ala Ser Ile Tyr Arg Gly Val Thr Arg His
 355 360 365
 His Gln His Gly Arg Trp Gln Ala Arg Ile Gly Arg Val Ala Gly Asn
 370 375 380
 Lys Asp Leu Tyr Leu Gly Thr Phe Ser Thr Gln Glu Glu Ala Ala Glu
 385 390 395 400
 Ala Tyr Asp Val Ala Ala Ile Lys Phe Arg Gly Leu Ser Ala Val Thr
 405 410 415

pctau2012001598-seql

Asn Phe Asp Met Ser Arg Tyr Asp Val Lys Thr Ile Leu Glu Ser Ser
 420 425 430
 Thr Leu Pro Ile Gly Gly Ala Ala Lys Arg Leu Lys Asp Met Glu Gln
 435 440 445
 Val Glu Leu Asn His Val Asn Val Asp Ile Ser His Arg Thr Glu Gln
 450 455 460
 Asp His Ser Ile Ile Asn Asn Thr Ser His Leu Thr Glu Gln Ala Ile
 465 470 475 480
 Tyr Ala Ala Thr Asn Ala Ser Asn Trp His Ala Leu Ser Phe Gln His
 485 490 495
 Gln Gln Pro His His His Tyr Asn Ala Asn Asn Met Gln Leu Gln Asn
 500 505 510
 Tyr Pro Tyr Gly Thr Gln Thr Gln Lys Leu Trp Cys Lys Gln Glu Gln
 515 520 525
 Asp Ser Asp Asp His Ser Thr Tyr Thr Thr Ala Thr Asp Ile His Gln
 530 535 540
 Leu Gln Leu Gly Asn Asn Asn Asn Asn Thr His Asn Phe Phe Gly Leu
 545 550 555 560
 Gln Asn Ile Met Ser Met Asp Ser Ala Ser Met Asp Asn Ser Ser Gly
 565 570 575
 Ser Asn Ser Val Val Tyr Gly Gly Gly Asp His Gly Gly Tyr Gly Gly
 580 585 590
 Asn Gly Gly Tyr Met Ile Pro Met Ala Ile Ala Asn Asp Gly Asn Gln
 595 600 605
 Asn Pro Arg Ser Asn Asn Asn Phe Gly Glu Ser Glu Ile Lys Gly Phe
 610 615 620
 Gly Tyr Glu Asn Val Phe Gly Thr Thr Thr Asp Pro Tyr His Ala Gln
 625 630 635 640
 Ala Ala Arg Asn Leu Tyr Tyr Gln Pro Gln Gln Leu Ser Val Asp Gln
 645 650 655
 Gly Ser Asn Trp Val Pro Thr Ala Ile Pro Thr Leu Ala Pro Arg Thr
 660 665 670
 Thr Asn Val Ser Leu Cys Pro Pro Phe Thr Leu Leu His Glu
 675 680 685

pctau2012001598-seql

<210> 447
 <211> 336
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> inducible promoter

<400> 447
 tcgatagttg tgatagttcc cacttgtccg tccgcatcgg catccgcagc tcgggatagt 60
 tccgacctag gattggatgc atgcggaacc gcacgagggc ggggcggaaa ttgacacacc 120
 actcctctcc acgcaccgtt caagaggtag gcgtatagag ccgtatagag cagagacgga 180
 gcactttctg gtactgtccg cacgggatgt ccgcacggag agccacaaac gagcggggcc 240
 ccgtacgtgc tctcctaccc caggatcgca tccccgcata gctgaacatc tatataaaga 300
 cccccaaggt tctcagtctc accaacaatca tcaacc 336

<210> 448
 <211> 2466
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> inducer

<400> 448
 atggccgaca ctagaagaag gcagaaccac tcttgtgacc catgccgtaa gggcaagaga 60
 agatgtgatg ctccagagaa ccgtaacgag gctaatagaga acggatgggt gtcattgtct 120
 aactgcaaga ggtggaacaa ggactgcacc ttcaactggc ttagctccca aaggctctaag 180
 gctaaggggtg ctgctccaag agctaggact aagaaggcta ggactgctac tactacctcc 240
 gagccttcta cttccgctgc tactattcca actcccgagt ccgataatca cgatgctcca 300
 ccagtgatca actcccacga tgctttgcca tcttgactc agggacttct ttctcacctt 360
 ggcgatctct tcgacttctc ccattctgct attccagcta acgctgagga tgctgctaac 420
 gtgcaatctg atgtccatt cccatgggat cttgctatcc caggcgattt ctctatggga 480
 cagcaacttg agaagcccct ctccccattg tctttccagg ctgttcttct tccaccacac 540
 tccccaaaca ctgatgatct cattcgtgag cttgaggaac agactaccga tccagattcc 600
 gtgactgaca ctaactccgt tcagcaagtt gctcaggatg gctctctttg gtctgatagg 660
 cagtctccac tcctcccaga aaacagtttg tgcatggctt ccgactctac cgctagaagg 720
 tatgctaggt ccacatgac caagaacctc atgaggatct accacgactc catggaaaac 780
 gccctttctt gctggcttac tgagcacaac tgcccatact ccgaccagat ttcttacctc 840
 ccaccaaagc aaagggctga gtggggacca aattggctta acaggatgtg cattaggggtg 900
 tgcaggctcg ataggggtgc aacttctctt agaggaaggg ctctctccgc tgaagaagat 960
 aaggctgctg ctagggcact tcaccttgct attgtggctt tcgcttctca gtggactcaa 1020
 catgtcaaaa ggggagctgg acttaacgtc ccagctgata ttgctgctga cgagcgttct 1080

pctau2012001598-seq1

attaggcgta acgcttggaa tgaggctagg catgcacttc agcacactac tggaatccca	1140
tccttcaggg tgatcttcgc caacatcatc ttcagcctca ctcagtcctg gctcgatgat	1200
gatgagcaac atggaatggg agctaggctc gataagcttc tcgagaatga tgggtgctcca	1260
gtgttcctcg agactgctaa taggcagctc tacaccttca ggcacaagtt cgctaggatg	1320
cagagaaggg gtaaggcttt caataggctt cctgggtggat ccgtggcttc tactttcgct	1380
ggaattttcg agactccac cccctcatct gagtctccac aacttgatcc agtgggtggct	1440
tctgaggaac acaggtctac tctgtctctc atgttctggc tcgggatcat gttcgacact	1500
ctgtctgctg ctatgtacca gaggccactt gttgtgtccg atgaggactc ccagatctct	1560
tctgcttctc caccaagaag aggtgccgag actcctatta accttgattg ctgggagcca	1620
ccaaggcagg tcccatctaa tcaagagaag tctgatgtgt ggggcgacct gttccttagg	1680
acttctgatt ctttgccga ccacgagtcc cacactcaaa tttctcaacc agctgctagg	1740
tggccatgca cttatgaaca agctgctgct gctctctcct ctgctactcc tgtaaggtg	1800
ttgctttaca ggcgtgtgac tcagctccag actttgttgt ataggggagc ttctccagct	1860
aggcttgagg ctgctattca gaggactctc tacgtgtaca accactggac tgctaagtac	1920
cagccattca tgcaggattg cgttgccaac catgagcttc tcccatccag gatccagtct	1980
tggtagctga tccttgatgg aactggcac cttgctgcta tgcttttggc tgatgtgctc	2040
gagtccatcg acagggattc ctactccgat atcaaccaca tcgacctcgt gactaagctc	2100
aggcttgata acgctcttgc tgtgtctgct ctcgctaggt catctcttag aggccaagaa	2160
ctcgatccag gcaaggcttc tccaatgtac aggcaattcc acgactccct tactgaggtt	2220
gcattccttg ttgagccatg gactgtggtg ctcatccact catttgctaa ggctgcttac	2280
atcctcctcg attgccttga tcttgatggc cagggaaacg ctctcgctgg ataccttcaa	2340
cttaggcaga actgcaacta ctgcatcagg gctctccagt tccttggccg taagtctgat	2400
atggctgctc tcgtggctaa ggatcttgag aggggactca acggaaaggt cgacagcttc	2460
ctctaa	2466