

REPLACED SHEETS		SG
☒	Description : paged to paged 29	
☒	Claim(s) : paged 30 to paged 32	
☒	Drawing(s) : paged 1/2 to paged 7/8	
☒	Abstract : paged 33 to paged -	
DATE OF RECEIPT		

9603872-4

IPCS

1

DNA ENCODING 2-ACYLTRANSFERASES

5 This invention relates to modified plants. In particular, the invention relates to plants modified such that at least part of the plant (for example seeds of the plant) is capable of yielding a commercially useful oil.

10 Plants have long been a commercially valuable source of oil. Nutritional uses of plant-derived oils have hitherto been dominant, but attention is now turning additionally to plants as a source of industrially useful oils, for example as replacements for or improvements on mineral oils. Oil seeds, such as from rape, have a variety of lipids in them (Hildish & Williams, "Chemical
15 Composition of Natural Lipids", Chapman Hall, London, 1964). There is now considerable interest in altering lipid composition by the use of recombinant DNA technology (Knauf, *TIBtech*, February 1987, 40-47), but by
20 no means all of the goals have been realised to date for a variety of reasons, in spite of the ever-increasing sophistication of the technology.

25 Success in tailoring the lipid content of plant-derived oils requires a firm understanding of the biochemistry and genes involved. Broadly, two approaches are available. First, plants may be modified to permit the synthesis of fatty acids which are new (for the plant); so, for example, laurate and/or stearate may be synthesised in rape. Secondly, the pattern and/or extent
30 of incorporation of fatty acids into the glycerol backbone of the lipid may be altered. It is with this latter approach that the present invention is concerned, although the former approach may additionally be used.

This Sheet Has Been Replaced



Lipids are formed in plants by the addition of fatty acid moieties onto the glycerol backbone by a series of acyl transferase enzymes. There are three positions on the glycerol molecule at which fatty acid (acyl) moieties may be substituted, and the substitution reached at each position is catalysed by a position-specific enzyme: the enzymes are known as 1-, 2- and 3-acyltransferases, respectively.

One, but not the only, current aim of "lipid engineering" in plants is to provide oils including lipids with a high content of erucic (22:1) acid. Erucic acid-containing lipids are commercially desirable for a number of purposes, particularly as replacements to or supplements for mineral oils in certain circumstances, as alluded to above. In the case of oil seed rape (*Brassica napus*), one of the most significant oil producing crops in cultivation today, the specificity of the 2-acyltransferase enzyme positively discriminates against the incorporation of erucic acid at position 2. So, even in those cultivars of rape which are able to incorporate erucic acid at positions 1 and 3, where there is no (or at least reduced) discrimination against erucic acid, only a maximum 66% of the fatty acids incorporated into triacyl glycerols can be erucic acid. Such varieties of rape are known as HEAR (high erucic acid rape) varieties.

It would therefore be desirable to increase the erucic acid content of conventional oil seed rape, as well as HEAR varieties; the same can be said of oils of other vegetable oil crops such as maize, sunflower and soya, to name but a few examples. While in principle it may be thought possible to introduce into a desired plant DNA encoding a 2-acyltransferase of different fatty acid



specificity, for example from a different plant, in practice there are a number of problems.

5 First, 2-acyltransferase and 3-acyltransferase are membrane bound, and therefore insoluble, enzymes. They have not been purified. This makes working with them difficult and rules out the use of many conventional DNA cloning procedures. This difficulty does not, paradoxically, lie in the way of cloning the gene (or at
10 least cDNA) encoding the 1-acyltransferase enzyme, which is soluble: in fact, recombinant DNA work has already been undertaken on this enzyme for a completely different purpose, namely the enhancement of chilling resistance in tobacco plant leaves, by Murata et al (*Nature* 356 710-713
15 (1992)).

Secondly, very little is known about the 2- and 3-acyltransferases. There is no idea of their size or how they are targeted to membranes. No nucleotide or amino
20 acid sequence data are available and no antibodies have been raised against them.

Although there has been discussion, therefore, of the desirability of modifying 2-acyltransferase specificity, for example by importing a gene coding for the
25 corresponding enzyme, but of different specificity, from another species, there is a pressing need in the art for the key which enables this work to be done. The present invention provides such a key, in the form of a DNA sequence (in the specific case, a cDNA sequence) encoding
30 a 2-acyltransferase. The DNA sequence in Figure 1 from nucleotides 130 to 1254 encodes the 2-acyltransferase from maize (*Zea mays*), including the stop codon.



According to a first aspect of the invention, therefore, there is provided a recombinant or isolated DNA sequence, preferably encoding an enzyme having membrane-bound acyltransferase activity, and selected from:

5

(i) a DNA sequence comprising the DNA sequence of Figure 1 encoding at least from MET₄₄ to Stop₄₁₈ (SEQ ID: 2) or its complementary strand,

10

(ii) nucleic acid sequences hybridising to the DNA sequence of Figure 1 (SEQ ID: 1), or its complementary strand, under stringent conditions, and

15

(iii) nucleic acid sequences which would hybridise to the DNA sequence of Figure 1 (SEQ ID: 1), or its complementary strand, but for the degeneracy of the genetic code.

20

Fragments of the above DNA sequences, for example of at least 15, 20, 30, 40 or 60 nucleotides in length, are also within the scope of the invention.

25

Suitable stringent conditions include salt solutions of approximately 0.9 molar, at temperatures of from 35°C to 65°C. More particularly, stringent hybridisation conditions include 6 x SSC, 5 x Denhardt's solution, 0.5% SDS, 0.5% tetrasodium pyrophosphate and 50µg/ml denatured herring sperm DNA; washing may be for 2 x 30 minutes at 65°C in 1 x SSC, 0.1% SDS and 1 x 30 minutes in 0.2 x

30

SSC, 0.1% SDS at 65°C.

Nucleic acid sequences within the scope of the first aspect of the invention will generally encode a protein



having 2-acyltransferase activity, as that is the activity of the enzyme encoded by the Figure 1 nucleic acid sequence. Nucleic acid sequences not encoding a protein having enzymic activity (or the relevant enzymic activity) but otherwise conforming to the first aspect of the invention as set out above may be useful for other purposes (and are therefore also encompassed by the invention); for example they may be useful as probes, which is a utility shared by the nucleic acid sequences of the first aspect of the invention, including the Figure 1 sequence itself.

The probe utility arises as follows. As there is likely to be a high degree of homology between acyltransferases of different species (and particularly between 2-acyltransferases of different species) the Figure 1 sequence (or part of it, or other sequences within the invention) may be used to probe cDNA or genomic libraries of other species in order to clone DNA sequences encoding acyltransferases having desired specificities. For example, if it is desired to produce oil having a high content of erucic acid esterified to glycerol, a DNA library of any species which naturally makes erucic acid may be probed. Suitable plants include meadow foam (*Limnanthes* spp., especially *L. alba* and, particularly, *L. douglassi*) and *Crambe*. *Limnanthes douglassi* is the preferred species, as specificity studies show that there is positive discrimination towards incorporation of erucic acid into position 2 of the triacylglyceride. Libraries of organisms other than the higher plants may be probed; for example, certain bacteria may have an acyltransferase of the desired specificity.

This Sheet Has Been Replaced



DNA in accordance with the invention will in general have a higher degree of homology with at least part of the Figure 1 sequence than with known sequences.

5 Recombinant DNA in accordance with the invention may be in the form of a vector, which may have sufficient regulatory sequences (such as a promoter) to direct expression. Vectors which are not expression vectors are useful for cloning purposes (as expression vectors themselves may be). Host cells (such as bacteria and
10 plant cells) containing vectors in accordance with the invention themselves form part of the invention.

DNA sequences in accordance with the invention can be
15 used in another way in cloning a gene of interest from another species: if the DNA is coupled to a suitable promoter, for example on an expression vector in a suitable host organism, protein may be produced. Such protein may be used to generate polyclonal or monoclonal
20 antibodies, or other binding molecules, which may then be used to screen for expression of homologous proteins in other species, for example as part of a DNA library screening programme.

25 Suitable cDNA libraries of target species will generally be prepared when the gene of interest is likely to be expressed; so cDNA embryo libraries (prepared at the early lipid synthesis stage), for example of *Limnanthes* spp. will be preferred.

30 The invention therefore enables the cloning of a wide variety of genes (or, more generally, DNA sequences) encoding acyltransferases, and 2-acyltransferases in particular, using DNA sequences as described above.



Such acyltransferases, such as from *Limnanthes* spp. may also be cloned directly, for example using complementation studies, from a DNA library of the species in question. For example, if *E. coli* is used as the complementation host, a mutant is chosen which is defective in the relevant enzyme (for example 2-acyltransferase); the DNA library from the target species (such as *L. douglassi*) is cloned into the mutant complementation host; host cells incorporating the target acyltransferase gene in their genome can readily be selected using appropriate selective media. *E. coli* mutant JC201 is a suitable host for use in complementation studies relating to 2-acyltransferase.

Cloning the acyltransferase gene of choice into a microbial host, such as a bacterium like *E. coli*, in such a way that the gene can be expressed has a particularly advantage in that the substrate specificity of the acyltransferase gene can be assessed in the microbial host before transformed plants are prepared, thereby saving considerably on research time. Such an assessment may be made by competitive substrate assays, in which differently detectably labelled candidate substrates for the enzyme compete with each other for incorporation into the glyceride. For example, ^{14}C -erucyl CoA and ^3H -oleoyl CoA can be used as competitive substrates for 2-acyltransferase, and the relative amounts of ^{14}C or tritium uptake into glyceride can be measured. (As 2-acyltransferases have acceptor, glycerol-based, substrates and donor, fatty acid-based, substrates, the experiment can be carried out with different acceptors, such as 1-erucyl-glycerol-3-phosphate and 1-oleoyl-glycerol-3-phosphate.) A gene coding for an enzyme which preferentially donates erucic acid to the acceptor



(particularly 1-erucyl-glycerol-3-phosphate) may by this means be identified as a DNA sequence of choice for further use in the invention as described below.

5 In a second aspect of the invention, there is provided a plant having one or more insoluble acyltransferase enzymes having a substrate specificity which differs from the native enzyme of the plant.

10 While site-directed mutagenesis and/or other protein engineering techniques may be used to alter the specificity of an enzyme native to the plant, it is preferred that the plant be transgenic and incorporate an expressible acyltransferase gene encoding an enzyme of
15 the desired specificity from another species. 2-acyltransferases are the enzymes of choice. For example, as described above, a 2-acyltransferase enzyme which has an enhanced specificity for, or at least no discrimination against, erucic acid, may be made by this
20 means to express in a plant which would not normally incorporate erucic acid into triacylglycerides. An important embodiment of the invention relates to genetically engineered plants which have higher levels of erucic acid incorporated into triacylglycerols than in
25 corresponding non-engineered plants. Preferable though this embodiment may be, though, the invention is not limited to the enhancement of erucic acid incorporation into glycerides: other acids may be desired in other circumstances.

30

For the acyltransferase transgene to be expressible, a promoter has to be operatively coupled to it. Because at the present state of the art it is difficult precisely to regulate the site of incorporation of a transgene into



the host genome, it is preferred that the transgene be coupled to its promoter prior to transformation of the plant. Promoters useful in the invention may be temporal- and/or seed-specific, but there is no need for them to be so: constitutive promoters, such as the CaMV 35S promoter, may in fact be preferred because they are usually strong promoters. Other tissues are unlikely to be adversely affected if the transgene encoding the acyltransferase enzyme is expressed in them, as the availability of the fatty acid CoA substrates is effectively limited to the seed.

The promoter-transgene construct, once prepared, is introduced into plant cells by any suitable means. The invention extends to such plant cells. Preferably, DNA is transformed into plant cells using a disarmed Ti-plasmid vector and carried by *Agrobacterium* by procedures known in the art, for example as described in EP-A-0116718 and EP-A-0270822. Alternatively, the foreign DNA could be introduced directly into plant cells using an electrical discharge apparatus. This method is preferred where *Agrobacterium* is ineffective, for example where the recipient plant is monocotyledonous. Any other method that provides for the stable incorporation of the DNA within the nuclear DNA of any plant cell of any species would also be suitable. This includes species of plant which are not currently capable of genetic transformation.

Preferably DNA in accordance with the invention also contains a second chimeric gene (a "marker" gene) that enables a transformed plant or tissue culture containing the foreign DNA to be easily distinguished from other plants or tissue culture that do not contain the foreign



DNA. Examples of such a marker gene include antibiotic resistance (Herrera-Estrella et al, *EMBO J.* 2(6) 987-95 (1983) and Herrera-Estrella et al, *Nature* 303 209-13 (1983)), herbicide resistance (EP-A-0242246) and glucuronidase (GUS) expression (EP-A-0344029). Expression of the marker gene is preferably controlled by a second promoter which allows expression in cells other than the tapetum, thus allowing selection of cells or tissue containing the marker at any stage of regeneration of the plant. The preferred second promoter is derived from the gene which encodes the 35S subunit of Cauliflower Mosaic Virus (CaMV) coat protein. However any other suitable second promoter could be used.

A whole plant can be regenerated from a single transformed plant cell, and the invention therefore provides transgenic plants (or parts of them, such as propagating material) including DNA in accordance with the invention as described above. The regeneration can proceed by known methods.

In one embodiment of the invention, the transgenic plant's native acyltransferase gene which corresponds to the transgene may be rendered at least partially inoperative or removed. So, if the transgene encodes a 2-acyltransferase, the plant's native 2-acyltransferase may be rendered inoperative by, for example, antisense or ribozyme techniques, as is known in the art.

By means of the invention, plants generating oil with a tailored lipid content may be produced. For example, the lipid composition of triacylglycerides in a plant may be substantially altered to produce triacylglycerides with a desired fatty acid (for example erucic acid) content



higher than has hitherto been possible. For example, oil seed rape (*B. napus*) may be transformed to produce oil whose triacylglyceride has an erucic acid content of over 70%.

5

It can readily be seen that plants with increased lipid levels may be produced by means of the invention. However, the invention is also useful for producing plants with decreased lipid levels, which may be desired if elevated protein and/or starch levels are required. Decreased lipid levels may be achieved by interfering with the proper functioning of a gene encoding a 2-acyltransferase, for example by antisense or ribozyme technology. (Such reduced-lipid plants may if desired be further engineered for higher protein and/or starch content, if wished.)

10

15

20

Promoters which naturally drive 2-acyltransferases may also be obtained by hybridisation and/or restriction and/or sequencing studies using the Figure 1 sequence.

25

30

The invention enables the production of protein encoded by DNA of the first aspect of the invention, should that be desired. The protein may be expressed by host cells harbouring DNA in the form of an expression vector. The protein, which may be an enzyme having 2-acyltransferase activity, may have an amino acid sequence which is identical to or homologous with the Figure 1 sequence. The degree of homology will generally be greater than that of known proteins, and may be at least 40, 50, 60, 70, 80, 90, 95 or 99%.

Preferred features of each aspect of the invention are as for each other aspect *mutatis mutandis*.

This Sheet Has Been Replaced



The invention is illustrated by the following examples. The examples refer to the accompanying drawings, in which:

- 5 FIGURE 1 shows the cDNA sequence derived in Example 1 (SEQ ID: 1) and its derived protein sequence (SEQ ID: 2).
- 10 FIGURE 2 shows a sequence alignment of part of the gene products of *plsB* (SEQ ID: 3) and *plsC* (SEQ ID: 4) with part of the sequence shown in Figure 1 (SEQ ID: 5), showing a conserved motif. *plsB* is the *E. coli* sn-glycerol-3-phosphate acyltransferase gene and *plsC* the 1-acyl-sn-glycerol-3-phosphate acyltransferase gene of *E. coli*. Double points indicate exact matches between two sequences and a single point conservative amino acid substitutions. Stars indicate identical amino acids in all three sequences and residues conserved in two out of the
- 15 three sequences are marked by a + symbol.
- 20 FIGURES 3A, 3B and 3C: Membrane phospholipids from *E. coli* strains were extracted into chloroform and separated by 2-dimensional thin layer chromatography. The first dimension (ascending) was
- 25 developed using chloroform:methanol:water (65:25:4) and the second dimension (left to right) developed with chloroform:methanol:acetic acid (65:25:10). Phospholipids were visualised by autoradiography for
- 30 16 hours at -70°C using Fuji RX film. The *E. coli* strains used were (Figure 3A): JC201 which carries a thermosensitive mutation in the 1-acyl-sn-glycerol-3-phosphate acyltransferase gene; (Figure 3B): JC201 containing the plasmid pPLSC, which



encodes the *E. coli* 1-acyl-sn-glycerol-3-phosphate acyltransferase gene; (Figure 3C): JC201 containing the plasmid whose cDNA insert sequence is shown in Figure 1. LPA, lysophosphatidic acid; PE, phosphatidylethanolamine; CL, cardiolipin; PG, phosphatidylglycerol; PA, phosphatidic acid; O, origin. 20% of the ^{32}P is incorporated in LPA in JC201 and all of the corresponding label is incorporated in PE in both of the other two strains.

FIGURE 4: Acyltransferase assays were performed using ^{32}P -labelled lysophosphatidic acid which had been extracted from the *E. coli* strain JC201 and oleoyl CoA as an acyl donor. Phospholipids present in the reaction mixtures were extracted into chloroform and separated using silica gel thin layer chromatography. Chloroform:methanol:acetic acid:water (25:15:4:2) was used to develop the plates. The phospholipids were visualised by autoradiography for 16 hours at -70°C using Fuji RX film. The *E. coli* strains used were: JC201 which carries a thermosensitive mutation in the 1-acyl-sn-glycerol-3-phosphate acyltransferase gene; JC201 containing the plasmid pPLSC which encodes the *E. coli* 1-acyl-sn-glycerol-3-phosphate acyltransferase gene; JC201 containing the plasmid whose maize cDNA insert sequence is shown in Figure 1. LPA, lysophosphatidic acid; PA, phosphatidic acid.

FIGURE 5 shows a comparison of the protein sequence shown in Figure 1 (SEQ ID: 6) with that derived (SEQ ID: 7) from a *B. napus* seed cDNA insert which was isolated by DNA hybridisation to the maize cDNA sequence. The sequences were aligned with the FastA



align program (1988). Double points signify identical amino acids and single points conservative amino acid substitutions.

5 Maize = 374 aa vs.rape = 311 aa
 51.5% identity; Optimised score: 705

EXAMPLE 1

Derivation of the Figure 1 cDNA sequence

10 Complementation studies using a maize cDNA expression
library transferred into the *E. coli* mutant JC201 allowed
the isolation of a plasmid encoding a 2-acyltransferase
enzyme from maize. The cDNA insert of this plasmid is
1.6kb in size, and includes a poly A tail of 70bp. The
15 insert was sequenced to give the data shown and
translation of the sequence revealed the presence of only
one large open reading frame. This is shown on Figure 1
with proposed start methionine and stop codon boxed. The
2-acyltransferase is 374 amino acids in size and
20 sequencing upstream of open reading frame showed that the
protein is expressed as part of a fusion protein in
E. coli. This consists of 10 amino acids of the β -
galactosidase protein, 43 amino acids (shown in sequence)
corresponding to the 5' untranslated region of the mRNA
25 and the 374 amino acid protein. Protein sequence
comparisons of the large open reading frame with the 2-
acyltransferase of *E. coli* show little overall identity
but there is a stretch of 80 residues which has a high
level of conservative substitution and contains some
30 amino acids that are conserved in the 2-acyltransferase,
1-acyltransferase and N-acetyl glucosamine
acyltransferase of *E. coli*.



EXAMPLE 2**Incorporation of ^{32}P into total phospholipids**

E. coli strains were grown in minimal medium containing ^{32}P orthophosphate. Total glycerolipids were extracted into organic solvents and separated by 2D thin layer chromatography (Figure 3) (Lysophosphatidic acid (LPA) is the substrate for 2-acyltransferase (2-AT)).

As can be seen in Figure 3A, the accumulation of ^{32}P -labelled LPA in the mutant JC 201 illustrates the absence of a fully functional 2-AT. Addition of a plasmid carrying either the native *E. coli* gene (Figure 3B), or the maize clone given in Figure 1 (Figure 3C) restores 2-AT activity to the cells, allowing LPA to be removed and further metabolised. (Lysophosphatidic acid.)

These data indicate that the DNA sequence given in Figure 1 codes for 2-AT.

EXAMPLE 3**Over expression of the cDNA**

The cDNA region specifying the protein sequence given in Figure 1 was cloned into the *E. coli* overexpression vector pET11d (Studier et al, *Meth. Enzymol.* 185 60-89 (1990)). Increased 2-acyltransferase activity following induction of expression from the plasmid insert confirmed that the sequence in Figure 1 is that of 2-AT.

EXAMPLE 4**Localisation of 2-AT activity in *E. coli* cells containing the maize clone**

2-acyltransferase assays were carried out using membranes isolated from the mutant strain JC.201 which lacks 2-AT and from JC.201 containing the maize plasmid (Figure 4).



2-AT activity was not detected in membrane fractions from JC.201. The addition of a plasmid carrying the native *E. coli* gene or the sequence given in Figure 1, to JC.201 resulted in restoration of 2-AT activity to the membranes.

EXAMPLE 5

Using the maize cDNA as a heterologous probe to obtain cDNA from oilseed rape

A seed cDNA library from *Brassica napus* was screened with the sequence given in Figure 1, using standard techniques (Sambrook et al "Molecular Cloning - A Laboratory Manual", 2nd Edition, Cold Spring Harbor Laboratory Press, 1989).

Conditions

For the hybridisation of the maize cDNA insert to the rape library: hybridisation was in 6xSSC, 5xDenhardts solution, 0.5%SDS, 0.5% tetrasodium pyrophosphate and 50ugml⁻¹ denatured herring sperm DNA. The filters were washed 2 x 30 minutes at 65°C in 1xSSC, 0.1%SDS and 1 x 30 minutes in 0.2xSSC, 0.1%SDS at 65°C.

A hybridising clone was sequenced and a protein sequence derived for the large ORF. Alignment of this protein sequence with that derived from the maize cDNA clone given in Figure 1, is shown in Figure 5.

The strong identity between these sequences illustrates the potential of using the sequence given in Figure 1 to obtain other 2-ATS.



EXAMPLE 6**Transgenic plants**

The sequence given in Figure 1 can be cloned, alongside a suitable promoter, into a suitable vector for expression in plants. The vector can be used to transform plants and the resulting plants expressing the 2-AT can be analysed for lipid content. Lipid metabolism is expected to be upregulated and elevated lipid levels were detectable in seeds.

EXAMPLE 7**Antisense**

The sequence given in Figure 1 may be cloned, alongside a suitable promoter, in the antisense orientation into a suitable vector for expression in plants. The vector can be used to transform plants and the resulting plants expressing the 2-AT can be analysed for protein and starch content. Elevated levels of starch and protein are expected to be detectable in seeds.

EXAMPLE 8**Down-regulation of Native 2-AT**

The DNA sequence of a 2-AT derived from *L. douglassii* (obtained as described in Example 5) can be introduced into oilseed rape (OSR) under the expression of a suitable promoter, using vectors and plant transformation methods well known in the art. A second sequence, comprising antisense or ribozymes against the rape cDNA (Example 5) can be introduced for simultaneous expression. The resultant transformed plant is expected to have 2-AT activity corresponding to that of *L. douglassii*, with concurrent down regulation of the native rape 2-AT gene.



The modified OSR plant thus obtained had higher levels of erucic acid in position 2 of its triacylglycerols than wild type plants. In addition higher levels of trierucin are found in the seed oil.

5

EXAMPLE 9**Genomic library screening**

10

The sequence given in Figure 1 is used to screen a genomic library of *Arabidopsis* and a hybrid using clone obtained. Using standard techniques, a promoter may be derived from this clone. The promoter may be used to drive expression in plant cell membranes.

This Sheet Has Been Replaced



SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

(A) NAME: Nickerson BIOCEM Limited
(non-US only)
(B) STREET: Cambridge Science Park, Milton Road
(C) CITY: Cambridge
(E) COUNTRY: United Kingdom
(F) POSTAL CODE (ZIP): CB4 4GZ

(A) NAME: SLABAS, Antoni Ryszard
(US only)
(B) STREET: 8 Telford Close, High Shinclffe
(C) CITY: Durham
(E) COUNTRY: United Kingdom
(F) POSTAL CODE (ZIP): DH1 2YJ

(A) NAME: BROWN, Adrian Paul
(US only)
(B) STREET: 4 Church Villas, Shadforth
(C) CITY: Durham
(E) COUNTRY: United Kingdom
(F) POSTAL CODE (ZIP): DH6 1LQ

(ii) TITLE OF INVENTION: DNA ENCODING 2-ACYLTRANSFERASES

(iii) NUMBER OF SEQUENCES: 7

(iv) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

(v) CURRENT APPLICATION DATA:

APPLICATION NUMBER: WO PCT/GB93/_____

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1514 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION: 130..1254

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

CCCCGTCCCTC CTCGTCGCCG GCGGAGCCGC CTACTATCGC CTGGAGAAGG AGCGCCGCGG	60
GGAGCTTTTC CCACTGCCGA CTGCCGTCTG ACCCTCCGAG ATCGGAAGCG GCGCCGGCGC	120
CGGCCGGCG ATG GCG ATC CCG CTC GTG CTC GTC GTG CTC CCG CTC GGC	168
Met Ala Ile Pro Leu Val Leu Val Val Leu Pro Leu Gly	
1 5 10	

This Sheet Has Been Replaced



CTG	CTC	TTC	CTC	CTG	TCC	GGC	CTC	ATC	GTC	AAC	GCC	ATC	CAG	GCC	GTC	216
Leu	Leu	Phe	Leu	Leu	Ser	Gly	Leu	Ile	Val	Asn	Ala	Ile	Gln	Ala	Val	
15						20					25					
CTA	TTT	GTG	ACG	ATA	AGG	CCC	TTT	TCG	AAG	AGC	TTC	TAC	CGT	CGG	ATC	264
Leu	Phe	Val	Thr	Ile	Arg	Pro	Phe	Ser	Lys	Ser	Phe	Tyr	Arg	Arg	Ile	
30					35					40					45	
AAC	AGA	TTC	TTG	GCC	GAG	CTG	CTG	TGG	CTT	CAG	CTT	GTC	TGG	GTG	GTG	312
Asn	Arg	Phe	Leu	Ala	Glu	Leu	Leu	Trp	Leu	Gln	Leu	Val	Trp	Val	Val	
				50					55					60		
GAC	TGG	TGG	GCA	GGT	GTT	AAG	GTA	CAA	CTG	CAT	GCA	GAT	GAG	GAA	ACT	360
Asp	Trp	Trp	Ala	Gly	Val	Lys	Val	Gln	Leu	His	Ala	Asp	Glu	Glu	Thr	
			65					70					75			
TAC	AGA	TCA	ATG	GGT	AAA	GAG	CAT	GCA	CTC	ATC	ATA	TCA	AAT	CAT	CGG	408
Tyr	Arg	Ser	Met	Gly	Lys	Glu	His	Ala	Leu	Ile	Ile	Ser	Asn	His	Arg	
		80					85					90				
AGT	GAT	ATT	GAT	TGG	CTC	ATT	GGA	TGG	ATA	TTG	GCC	CAG	CGT	TCA	GGG	456
Ser	Asp	Ile	Asp	Trp	Leu	Ile	Gly	Trp	Ile	Leu	Ala	Gln	Arg	Ser	Gly	
	95					100					105					
TGC	CTT	GGA	AGT	ACA	CTT	GCT	GTC	ATG	AAG	AAG	TCA	TCC	AAG	TTC	CTT	504
Cys	Leu	Gly	Ser	Thr	Leu	Ala	Val	Met	Lys	Lys	Ser	Ser	Lys	Phe	Leu	
110					115					120					125	
CCA	GTT	ATT	GGC	TGG	TCA	ATG	TGG	TTT	GCA	GAG	TAC	CTC	TTT	TTG	GAA	552
Pro	Val	Ile	Gly	Trp	Ser	Met	Trp	Phe	Ala	Glu	Tyr	Leu	Phe	Leu	Glu	
				130					135					140		
AGG	AGC	TGG	GCC	AAG	GAT	GAA	AAG	ACA	CTA	AAG	TGG	GGT	CTC	CAA	AGG	600
Arg	Ser	Trp	Ala	Lys	Asp	Glu	Lys	Thr	Leu	Lys	Trp	Gly	Leu	Gln	Arg	
			145					150					155			
TTG	AAA	GAC	TTC	CCT	AGA	CCA	TTT	TGG	CTA	GCT	CTT	TTC	GTC	GAG	GGT	648
Leu	Lys	Asp	Phe	Pro	Arg	Pro	Phe	Trp	Leu	Ala	Leu	Phe	Val	Glu	Gly	
		160					165					170				
ACT	CGC	TTT	ACT	CCA	GCA	AAG	CTT	CTC	GCA	GCT	CAG	GAA	TAT	GCG	GCC	696
Thr	Arg	Phe	Thr	Pro	Ala	Lys	Leu	Leu	Ala	Ala	Gln	Glu	Tyr	Ala	Ala	
		175				180					185					
TCC	CAG	GGC	TTA	CCG	GCT	CCT	AGA	AAT	GTA	CTT	ATT	CCA	CGT	ACC	AAG	744
Ser	Gln	Gly	Leu	Pro	Ala	Pro	Arg	Asn	Val	Leu	Ile	Pro	Arg	Thr	Lys	
190					195					200					205	
GGA	TTT	GTA	TCT	GCT	GTA	AGT	ATT	ATG	CGA	GAT	TTT	GTT	CCA	GCC	ATT	792
Gly	Phe	Val	Ser	Ala	Val	Ser	Ile	Met	Arg	Asp	Phe	Val	Pro	Ala	Ile	
				210					215					220		
TAT	GAT	ACA	ACT	GTA	ATA	GTC	CCT	AAA	GAT	TCC	CCT	CAA	CCA	ACA	ATG	840
Tyr	Asp	Thr	Thr	Val	Ile	Val	Pro	Lys	Asp	Ser	Pro	Gln	Pro	Thr	Met	
			225					230					235			
CTG	CGG	ATT	TTG	AAA	GGG	CAA	TCA	TCA	GTG	ATA	CAT	GTC	CGC	ATG	AAA	888
Leu	Arg	Ile	Leu	Lys	Gly	Gln	Ser	Ser	Val	Ile	His	Val	Arg	Met	Lys	
		240					245					250				
CGT	CAT	GCA	ATG	AGT	GAG	ATG	CCA	AAA	TCA	GAT	GAG	GAT	GTT	TCA	AAA	936
Arg	His	Ala	Met	Ser	Glu	Met	Pro	Lys	Ser	Asp	Glu	Asp	Val	Ser	Lys	
	255					260					265					



TGG	TGT	AAA	GAC	ATT	TTT	GTG	GCA	AAG	GAT	GCC	TTA	CTG	GAC	AAG	CAT	984
Trp	Cys	Lys	Asp	Ile	Phe	Val	Ala	Lys	Asp	Ala	Leu	Leu	Asp	Lys	His	
270					275					280					285	
TTG	GCA	ACA	GGC	ACT	TTC	GAT	GAG	GAG	ATT	AGA	CCT	ATT	GGC	CGT	CCA	1032
Leu	Ala	Thr	Gly	Thr	Phe	Asp	Glu	Glu	Ile	Arg	Pro	Ile	Gly	Arg	Pro	
				290					295					300		
GTG	AAA	TCA	TTG	CTG	GTG	ACC	CTG	TTC	TGG	TCG	TGC	CTC	CTG	CTG	TTT	1080
Val	Lys	Ser	Leu	Leu	Val	Thr	Leu	Phe	Trp	Ser	Cys	Leu	Leu	Leu	Phe	
			305					310					315			
GGC	GCC	ATC	GAG	TTC	TTC	AAG	TGG	ACA	CAG	CTT	CTG	TCG	ACG	TGG	AGG	1128
Gly	Ala	Ile	Glu	Phe	Phe	Lys	Trp	Thr	Gln	Leu	Leu	Ser	Thr	Trp	Arg	
		320					325					330				
GGT	GTG	GCG	TTC	ACT	GCC	GCA	GGG	ATG	GCG	CTT	GTG	ACG	GGT	GTC	ATG	1176
Gly	Val	Ala	Phe	Thr	Ala	Ala	Gly	Met	Ala	Leu	Val	Thr	Gly	Val	Met	
	335					340					345					
CAT	GTC	TTC	ATC	ATG	TTC	TCC	CAG	GCT	GAG	CGG	TCG	AGC	TCA	GCC	AGG	1224
His	Val	Phe	Ile	Met	Phe	Ser	Gln	Ala	Glu	Arg	Ser	Ser	Ser	Ala	Arg	
350					355					360					365	
GCG	GCA	CGG	AAC	CGG	GTC	AAG	AAG	GAA	TGAAAAATGG	AGGGTGGAGA						1271
Ala	Ala	Arg	Asn	Arg	Val	Lys	Lys	Glu								
			370					375								
TGAGGTTCTC	GTGGGGTTTG	TTATGGGCAA	CCTTCAAAAG	GACTCTCCAT	TCATATTAGT											1331
ATTAATTCAT	ATATATGCAG	CGCCAAATTC	CAGACATTGA	TATGCTCTCA	AATAGGATGT											1391
TCTGCTCCCC	TCTTGATTTT	GTATGCAGGA	AAGGGTTTGT	AGGGAGTTTA	CCCCCCCCCC											1451
CCCCCCCCCC	GCCTTTCTTT	GGGGAAGAAA	GACATATTCT	GGAAGCCTTC	CAGTAGTTCA											1511
AAA																1514

This Sheet Has Been Replaced





(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 374 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```

Met Ala Ile Pro Leu Val Leu Val Val Leu Pro Leu Gly Leu Leu Phe
 1             5             10             15
Leu Leu Ser Gly Leu Ile Val Asn Ala Ile Gln Ala Val Leu Phe Val
          20             25             30
Thr Ile Arg Pro Phe Ser Lys Ser Phe Tyr Arg Arg Ile Asn Arg Phe
          35             40             45
Leu Ala Glu Leu Leu Trp Leu Gln Leu Val Trp Val Val Asp Trp Trp
          50             55             60
Ala Gly Val Lys Val Gln Leu His Ala Asp Glu Glu Thr Tyr Arg Ser
          65             70             75             80
Met Gly Lys Glu His Ala Leu Ile Ile Ser Asn His Arg Ser Asp Ile
          85             90             95
Asp Trp Leu Ile Gly Trp Ile Leu Ala Gln Arg Ser Gly Cys Leu Gly
          100            105            110
Ser Thr Leu Ala Val Met Lys Lys Ser Ser Lys Phe Leu Pro Val Ile
          115            120            125
Gly Trp Ser Met Trp Phe Ala Glu Tyr Leu Phe Leu Glu Arg Ser Trp
          130            135            140
Ala Lys Asp Glu Lys Thr Leu Lys Trp Gly Leu Gln Arg Leu Lys Asp
          145            150            155            160
Phe Pro Arg Pro Phe Trp Leu Ala Leu Phe Val Glu Gly Thr Arg Phe
          165            170            175
Thr Pro Ala Lys Leu Leu Ala Ala Gln Glu Tyr Ala Ala Ser Gln Gly
          180            185            190
Leu Pro Ala Pro Arg Asn Val Leu Ile Pro Arg Thr Lys Gly Phe Val
          195            200            205
Ser Ala Val Ser Ile Met Arg Asp Phe Val Pro Ala Ile Tyr Asp Thr
          210            215            220
Thr Val Ile Val Pro Lys Asp Ser Pro Gln Pro Thr Met Leu Arg Ile
          225            230            235            240
Leu Lys Gly Gln Ser Ser Val Ile His Val Arg Met Lys Arg His Ala
          245            250            255
Met Ser Glu Met Pro Lys Ser Asp Glu Asp Val Ser Lys Trp Cys Lys
          260            265            270
Asp Ile Phe Val Ala Lys Asp Ala Leu Leu Asp Lys His Leu Ala Thr
          275            280            285

```

Gly	Thr	Phe	Asp	Glu	Glu	Ile	Arg	Pro	Ile	Gly	Arg	Pro	Val	Lys	Ser
290						295					300				
Leu	Leu	Val	Thr	Leu	Phe	Trp	Ser	Cys	Leu	Leu	Leu	Phe	Gly	Ala	Ile
305					310					315					320
Glu	Phe	Phe	Lys	Trp	Thr	Gln	Leu	Leu	Ser	Thr	Trp	Arg	Gly	Val	Ala
				325					330					335	
Phe	Thr	Ala	Ala	Gly	Met	Ala	Leu	Val	Thr	Gly	Val	Met	His	Val	Phe
			340					345					350		
Ile	Met	Phe	Ser	Gln	Ala	Glu	Arg	Ser	Ser	Ser	Ala	Arg	Ala	Ala	Arg
		355					360					365			
Asn	Arg	Val	Lys	Lys	Glu										
370															

This Sheet Has Been Replaced



(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

Tyr	Phe	Val	Glu	Gly	Gly	Arg	Ser	Arg	Thr	Gly	Arg	Leu	Leu	Asp
1				5					10					15

This Sheet Has Been Replaced

(2) INFORMATION FOR SEQ ID NO: 4:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 15 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

Met	Phe	Pro	Glu	Gly	Thr	Arg	Ser	Arg	Gly	Arg	Gly	Leu	Leu	Pro
1				5					10					15

This Sheet Has Been Replaced





(2) INFORMATION FOR SEQ ID NO: 5:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 15 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

Leu Phe Val Glu Gly Thr Arg Phe Thr Pro Ala Lys Leu Leu Ala
 1 5 10 15

(2) INFORMATION FOR SEQ ID NO: 6:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 374 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

Met Ala Ile Pro Leu Val Leu Val Val Leu Pro Leu Gly Leu Leu Phe
 1 5 10 15
 Leu Leu Ser Gly Leu Ile Val Asn Ala Ile Gln Ala Val Leu Phe Val
 20 25 30
 Thr Ile Arg Pro Phe Ser Lys Ser Phe Tyr Arg Arg Ile Asn Arg Phe
 35 40 45
 Leu Ala Glu Leu Leu Trp Leu Gln Leu Val Trp Val Val Asp Trp Trp
 50 55 60
 Ala Gly Val Lys Val Gln Leu His Ala Asp Glu Glu Thr Tyr Arg Ser
 65 70 75 80
 Met Gly Lys Glu His Ala Leu Ile Ile Ser Asn His Arg Ser Asp Ile
 85 90 95
 Asp Trp Leu Ile Gly Trp Ile Leu Ala Gln Arg Ser Gly Cys Leu Gly
 100 105 110
 Ser Thr Leu Ala Val Met Lys Lys Ser Ser Lys Phe Leu Pro Val Ile
 115 120 125
 Gly Trp Ser Met Trp Phe Ala Glu Tyr Leu Phe Leu Glu Arg Ser Trp
 130 135 140
 Ala Lys Asp Glu Lys Thr Leu Lys Trp Gly Leu Gln Arg Leu Lys Asp
 145 150 155 160
 Phe Pro Arg Pro Phe Trp Leu Ala Leu Phe Val Glu Gly Thr Arg Phe
 165 170 175
 Thr Pro Ala Lys Leu Leu Ala Ala Gln Glu Tyr Ala Ala Ser Gln Gly
 180 185 190

Leu Pro Ala Pro Arg Asn Val Leu Ile Pro Arg Thr Lys Gly Phe Val
 195 200 205
 Ser Ala Val Ser Ile Met Arg Asp Phe Val Pro Ala Ile Tyr Asp Thr
 210 215 220
 Thr Val Ile Val Pro Lys Asp Ser Pro Gln Pro Thr Met Leu Arg Ile
 225 230 235 240
 Leu Lys Gly Gln Ser Ser Val Ile His Val Arg Met Lys Arg His Ala
 245 250 255
 Met Ser Glu Met Pro Lys Ser Asp Glu Asp Val Ser Lys Trp Cys Lys
 260 265 270
 Asp Ile Phe Val Ala Lys Asp Ala Leu Leu Asp Lys His Leu Ala Thr
 275 280 285
 Gly Thr Phe Asp Glu Glu Ile Arg Pro Ile Gly Arg Pro Val Lys Ser
 290 295 300
 Leu Leu Val Thr Leu Phe Trp Ser Cys Leu Leu Leu Phe Gly Ala Ile
 305 310 315 320
 Glu Phe Phe Lys Trp Thr Gln Leu Leu Ser Thr Trp Arg Gly Val Ala
 325 330 335
 Phe Thr Ala Ala Gly Met Ala Leu Val Thr Gly Val Met His Val Phe
 340 345 350
 Ile Met Phe Ser Gln Ala Glu Arg Ser Ser Ser Ala Arg Ala Ala Arg
 355 360 365
 Asn Arg Val Lys Lys Glu
 370

This Sheet Has Been Replaced





(2) INFORMATION FOR SEQ ID NO: 7:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 295 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

```

Met Ala Met Ala Ala Ala Val Ile Val Pro Leu Gly Ile Leu Phe Phe
1           5           10           15
Ile Ser Gly Leu Val Val Asn Leu Leu Gln Arg Ser Gly Cys Leu Gly
20           25           30
Ser Ala Leu Ala Val Met Lys Lys Ser Ser Lys Phe Leu Pro Val Ile
35           40           45
Gly Trp Ser Met Trp Phe Ser Glu Tyr Leu Phe Leu Glu Arg Asn Trp
50           55           60
Ala Lys Asp Glu Ser Thr Leu Lys Ser Gly Leu Gln Arg Leu Asn Asp
65           70           75           80
Phe Pro Arg Pro Phe Trp Leu Ala Leu Phe Val Glu Gly Thr Arg Phe
85           90           95
Thr Glu Ala Lys Leu Lys Ala Ala Gln Glu Tyr Ala Ala Ser Ser Glu
100          105          110
Leu Pro Val Pro Arg Asn Val Leu Ile Pro Arg Thr Lys Gly Phe Val
115          120          125
Ser Ala Val Ser Asn Met Arg Ser Phe Val Pro Ala Ile Tyr Asp Met
130          135          140
Thr Val Ala Ile Pro Lys Thr Ser Pro Pro Pro Thr Met Leu Arg Leu
145          150          155          160
Phe Lys Gly Gln Pro Ser Val Val His Val His Ile Lys Cys His Ser
165          170          175
Met Lys Asp Leu Pro Glu Ser Glu Asp Glu Ile Ala Gln Trp Cys Arg
180          185          190
Asp Gln Phe Val Thr Lys Asp Ala Leu Leu Asp Lys His Ile Ala Ala
195          200          205
Asp Thr Phe Ala Gly Gln Lys Glu Gln Asn Ile Gly Arg Pro Ile Lys
210          215          220
Ser Leu Ala Val Val Leu Ser Trp Ala Cys Leu Leu Thr Leu Gly Ala
225          230          235          240
Met Lys Phe Leu His Trp Ser Asn Leu Phe Ser Ser Trp Lys Gly Ile
245          250          255
Ala Leu Ser Ala Leu Gly Leu Gly Ile Ile Thr Leu Cys Met Gln Ile
260          265          270
Leu Ile Arg Ser Ser Gln Ser Glu Arg Ser Thr Pro Ala Lys Val Ala
275          280          285

```

Pro Ala Lys Pro Lys Asp Asn
290 295

This Sheet Has Been Replaced



CLAIMS

- 5 1. A recombinant or isolated DNA sequence selected from
- (i) a DNA sequence comprising the DNA sequence of Figure 1 encoding at least from Met₄₄ to Stop₄₁₈ (SEQ ID: 2) or its complementary strand;
- 10 (ii) nucleic acid sequences hybridising to the DNA sequence of Figure 1 (SEQ ID: 1) or its complementary strand under stringent conditions; and
- 15 (iii) nucleic acid sequences which would hybridise to the DNA sequence of Figure 1 (SEQ ID: 1), or its complementary strand, but for the degeneracy of the genetic code.
- 20 2. A DNA sequence as claimed in claim 1 which encodes an enzyme having membrane-bound acyltransferase activity.
3. A DNA sequence as claimed in claim 2, in which the acyltransferase activity is 2-acyltransferase activity.
- 25 4. An isolated protein which is the expression product of a DNA sequence as claimed in claim 1, 2 or 3.
5. An antibody capable of specifically binding to a protein as claimed in claim 4.
- 30 6. A plant having one or more insoluble acyltransferase enzymes having a substrate specificity which differs from the native enzyme of the plant.





7. A plant as claimed in claim 6, which is transgenic for an insoluble acyltransferase enzyme from another species.
- 5 8. A plant as claimed in claim 6 or 7, which is rape, maize, sunflower or soya.
9. A plant as claimed in claim 7, wherein the said other species is a species of the genus *Limnanthes* (such as *L. alba* or, preferably, *L. douglassi*) or *Crambe*.
- 10 10. A plant as claimed in any one of claims 6 to 9, wherein the insoluble acyltransferase enzyme is 2-acyltransferase.
- 15 11. A plant as claimed in claim 10, wherein the 2-acyltransferase has a higher specificity for erucic acid than the native enzyme of the plant.
- 20 12. A plant as claimed in any one of claims 6 to 11, wherein the said native enzyme is at least partially rendered inoperative or removed, for example by a ribozyme or by antisense nucleic acid.
- 25 13. A method of generating oil, the method comprising cultivating a plant as claimed in anyone of claims 6 to 12 and harvesting oil produced by the plant or a part (particularly seeds) thereof.
- 30 14. A microbial host capable of expressing a DNA sequence as claimed in claim 1, 2 or 3.
15. A protein which is substantially homologous to a protein as claimed in claim 4.



16. A fragment of a DNA sequence as claimed in claim 1, 2 or 3, comprising at least 15 nucleotides.

5 17. DNA encoding RNA which is antisense to RNA encoded by DNA as claimed in claim 1, 2 or 3.

18. DNA encoding a ribozyme specific to RNA encoded by DNA as claimed in claim 1, 2 or 3.

10 19. Isolated or recombinant DNA containing a promoter which naturally drives expression of a gene to produce a protein as claimed in claim 4 or 15.

This Sheet Has Been Replaced



* C L M S *

ABSTRACTDNA ENCODING 2-ACYLTRANSFERASES

5 Plants, particularly transgenic plants, may be produced
 having a 2-acyltransferase enzyme or other insoluble
 acyltransferase enzyme with an altered substrate
 specificity compared to the native enzyme. For example,
 oil seed rape (*Brassica napus*) may contain a 2-
10 acyltransferase transgene derived from *Limnanthes*
 douglassi in order to increase the erucic acid content of
 the oil. The cDNA sequence of maize (*Zea mays*) 2-
 acyltransferase is disclosed and is useful for cloning
 acyltransferase genes and/or cDNAs from other organisms,
15 including *L. douglassi*.

This Sheet Has Been Replaced





1/8

FIG. 1 (1/3)

1/1	CCC CGT CCT CCT CGT CGC CGG CGG AGC CGC CTA CTA TCG CCT GGA GAA GGA GCG CCG CGG	31/11	CGG CTA CTA TCG CCT GGA GAA GGA GCG CCG CGG	CGG
61/21	GGA GCT TTT CCC ACT GCC GAC TGC CGT CTG ACC CTC CGA GAT CGG AAG CGG CGC CGG CGC	91/31	CGA GAT CGG AAG CGG CGC CGC CGG CGC	CGC
121/41	CGG CCG GCG ATG GCG ATC CCG CTC GTG CTC GTG CTC GTC GTG CTC CCG CTC GGC CTG CTC TTC CTC	151/51	CTC GTC GTG CTC CCG CTC GGC CTG CTC TTC CTC	CTC
181/61	CTG TCC GGC CTC ATC GTC AAC GCC ATC CAG GCC GTC CTA TTT GTG ACG ATA AGG CCC TTT	211/71	CTA TTT GTG ACG ATA AGG CCC TTT	TTT
241/81	TCG AAG AGC TTC TAC CGT CGG ATC AAC AGA TTC TTG GCC GAG CTG CTG TGG CTT CAG CTT	271/91	GCC GAG CTG CTG TGG CTT CAG CTT	CTT
301/101	GTC TGG GTG GAC TGG TGG GCA GGT GTT AAG GTA CAA CTG CAT GCA GAT GAG GAA ACT	331/111	CAA CTG CAT GCA GAT GAG GAA ACT	ACT
361/121	TAC AGA TCA ATG GGT AAA GAG CAT GCA CTC ATC ATA TCA AAT CAT CGG AGT GAT ATT GAT	391/131	TCA AAT CAT CGG AGT GAT ATT GAT	GAT
421/141	TGG CTC ATT GGA TGG ATA TTG GCC CAG CGT TCA GGG TGC CTT GGA AGT ACA CTT GCT GTC	451/151	TGC CTT GGA AGT ACA CTT GCT GTC	GTC
481/161	ATG AAG AAG TCC AAG TTC CTT CCA GTT ATT GGC TGG TCA ATG TGG TTT GCA GAG TAC	511/171	TGG TCA ATG TGG TTT GCA GAG TAC	TAC
541/181	CTC TTT TTG GAA AGG AGC TGG GCC AAG GAT GAA AAG ACA CTA AAG TGG GGT CTC CAA AGG	571/191	CTA AAG TGG GGT CTC CAA AGG	AGG



2/8

FIG. 1
(2/3)

601/201	TTG AAA	GAC TTC	CCT AGA	CCA TTT	TGG CTA	GCT CTT	631/211	TTC GTC	GAG GGT	ACT ACT	CGC TTT	ACT
L K	D F	P R	P F	L W	L L	A L		F V	E G	T T	R F	T
661/221	CCA GCA	AAG CTT	CTC GCA	GCT CAG	GAA TAT	GCG GCC	691/231	TCC CAG	GGC TTA	CCG GCT	CCT AGA	
P A	K L	L A	A Q	E Y	A A	A A		S Q	G L	P A	P R	
721/241	AAT GTA	CTT ATT	CCA CGT	ACC AAG	GGA TTT	GTA TCT	751/251	GCT GTA	AGT ATT	ATG CGA	GAT TTT	
N V	L I	P R	T K	G F	V S	A A		A V	S I	M R	D F	
781/261	GTT CCA	GCC ATT	TAT GAT	ACA ACT	GTA ATA	GTC CCT	811/271	AAA GAT	TCC CCT	CAA CCA	ACA ATG	
V P	A I	Y D	T T	V I	V P	K D		S D	S P	Q P	T M	
841/281	CTG CGG	ATT TTG	AAA GGG	CAA TCA	GTG ATA	CAT CAT	871/291	GTC CGC	ATG AAA	CGT CAT	GCA ATG	
L R	I L	K G	Q S	S V	I H	V V		R M	K R	H A	M	
901/301	AGT GAG	ATG CCA	AAA TCA	GAT GAT	GTT TCA	AAA TGG	931/311	TGT AAA	GAC ATT	TTT GTG	GCA	
S E	M P	K S	D E	D V	S K	W C		K D	I F	V A		
961/321	AAG GAT	GCC TTA	CTG GAC	AAG CAT	TTG GCA	ACA GGC	991/331	TTC GAT	GAG GAG	ATT AGA	CCT	
K D	A L	L D	K H	L A	T G	T F		D E	E I	R P		
1021/341	ATT GGC	CGT CCA	GTG AAA	TCA TTG	CTG GTG	ACC CTG	1051/351	TGG TCG	TGC CTC	CTG CTG	TTT	
I G	R P	V K	S L	L V	T L	F W		S C	L L	L F		
1081/361	GGC GCC	ATC GAG	TTT TTC	AAG TGG	ACA CAG	CTT CTG	1111/371	TCG ACG	TGG AGG	GGT GTG	GCG TTC	
G A	I E	F F	K W	T Q	L L	S T		W R	G V	A F		



3/8

1141/381
 ACT GCC GCA GGG ATG GCG CTT GTG ACG GGT GTC ATG CAT GTC TTC ATC ATG TTC TCC CAG
 T A A G M A L V T G V M H V F I M F S Q
 1201/401
 GCT GAG CGG TCG AGC TCA GCC AGG GCG GCA CGG AAC CGG GTC AAG AAG GAA TGA AAA ATG
 A E R S S S A R A A R N R V K K E * K M
 1261/421
 GAG GGT GGA GAT GAG GTT CTC GTG GGG TTT GTT ATG GGC AAC CTT CAA AAG GAC TCT CCA
 E G G D E V L V G F V M G N L Q K D S P
 1321/441
 TTC ATA TTA GTA TTA ATT CAT ATA TAT GCA GCG CCA AAT TCC AGA CAT TGA TAT GCT CTC
 F I L V L I H I Y A A P N S R H * Y A L
 1381/461
 AAA TAG GAT GTT CTG CTC CCC TCT TGT ATT TGT ATG CAG GAA AGG GTT TGT AGG GAG TTT
 K * D V L L P S C I C M Q E R V C R E F
 1441/481
 ACC CCC CCC CCC CCC GGC TTT CTT TGG GGA AGA AAG ACA TAT TCT GGA AGC CTT
 T P P P P P A F L W G R K T Y S G S L
 1501/501
 CCA GTA GTT CAA AA
 P V V Q

FIG. 1 (3/3)



4/8

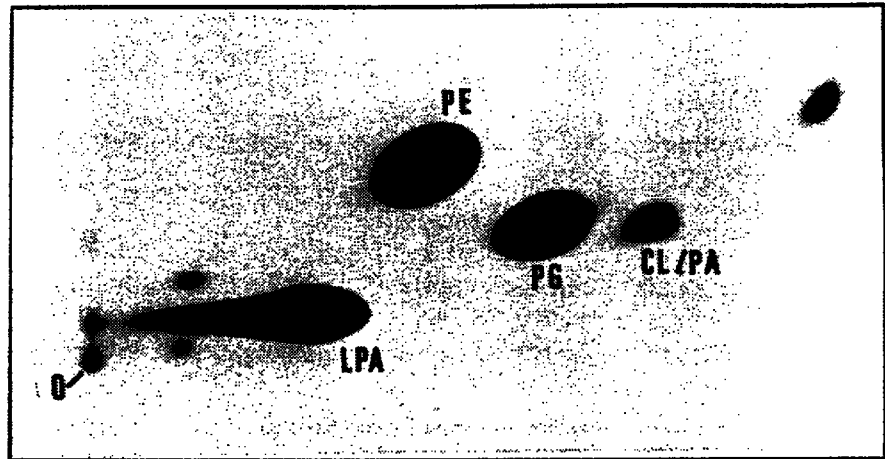
FIG. 2

plsb	-	Y	F	V	E	G	G	R	S	R	T	G	R	L	L	D	-
			.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
plsc	-	M	F	P	E	G	T	R	S	R	G	R	G	L	L	P	-
			.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
maize	-	L	F	V	E	G	T	R	F	T	P	A	K	L	L	A	-
			*	+	*	*	+	*	+	+				*	*	*	*

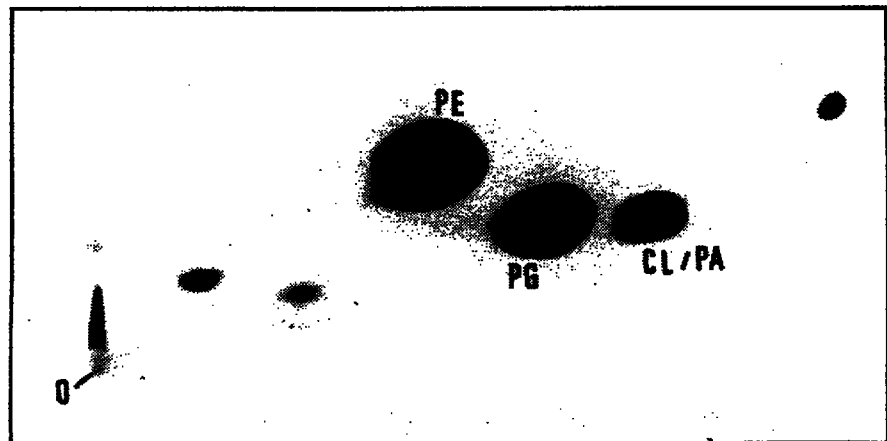


5 / 8
FIG. 3

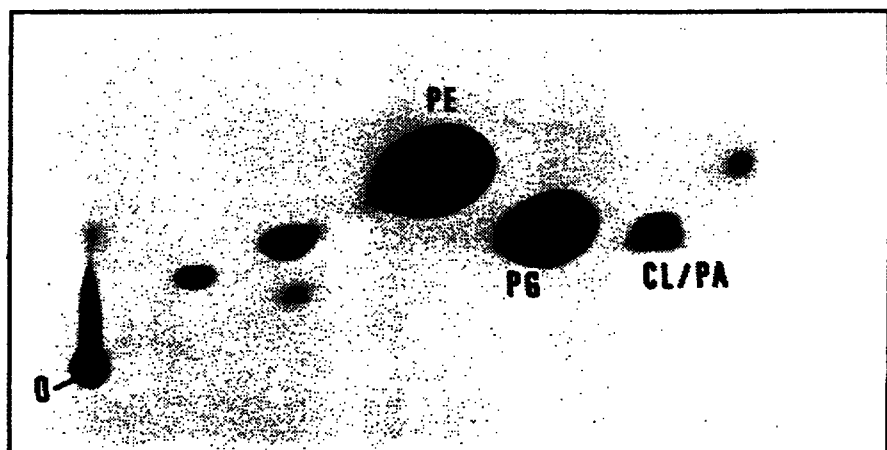
A : mutant
JC201



B : mutant +
pPLSC



C : mutant +
maize
cDNA



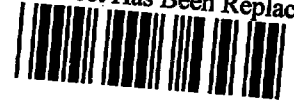
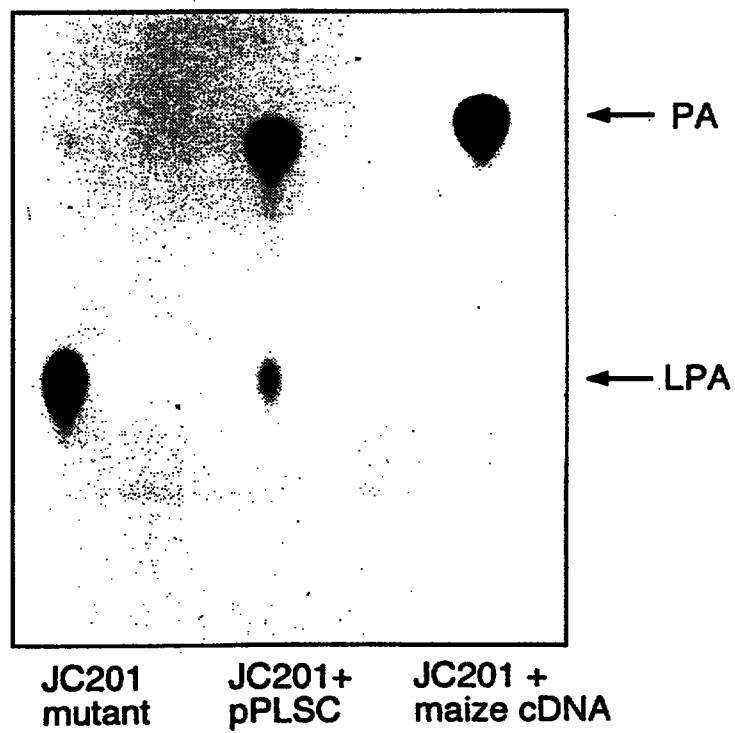


FIG. 4





7 / 8

FIG. 5 (1/2)

Maize = 374 aa vs.rape = 311 aa

51.5% identity; Optimized score: 705

	10	20	30	40	50	60
maize	MAIPLVLVPLGLLFLSGLIVNAIQAVLFVTIRPFSSKSFYRRINRFLAELLWLQLVWV					
	::: . :::::::::::::::					
rape	MAMA-AAVIVPLGILFFISGLVNVN-----					
	10	20				
	70	80	90	100	110	120
maize	VDWWAGVKVQLHADEETYSRSMGKEHALIISNHRSDIDWLIGWILAQRSGLGSTLAVMKK					
	: :::::::::::::::					
rape	-----LLQMSGCLGSALAVMKK					
				30		40
	130	140	150	160	170	180
maize	SSKFLPVIGSMWFAEYLFLEERSWAKDEKTLKWGLQRLKDFPRPFWLALFVEGTRFTPAK					
	::::::::::::::::::::: ::::::::::::::: ::::::::::::::: ::					
rape	SSKFLPVIGSMWFSEYLFLEERNWAKDESTLKSGLQRLNDFPRPFWLALFVEGTRFTEAK					
	50	60	70	80	90	100



88 / 88

	190	200	210	220	230	240
maize	LLAAQEAASQGLPAPRNVLI	PRTKGFVSAVSIMRDFVPAIYD	TTVIVPKDSPQPTMLRI			
	:	:	:	:	:	:
rape	LKAAQEYAASELPPVRNVLI	PRTKGFVSAVSNMRSFVPAIYD	MTVAIPKTSPPPTMLRL			
	110	120	130	140	150	160

	250	260	270	280	290
maize	LKGQSSVHVVRMKRRHAMSE	MPKSDVDVSKWCKDIFVAKD	ALLDKHLATGTF-DEEIRPIG		
	:	:	:	:	:
rape	FKGQPSVVHVHIKCHSMKDL	PESEDEIAQWCRDQFVTKD	ALLDKHIAADTFAGQKEQNIG		
	170	180	190	200	210

	300	310	320	330	340	350
maize	RPVKSLVTLFWSCLLLF	GAEFFKWTQLLSTWRGVA	FTAGMALVTGVMHVFIMFSQAE			
	:	:	:	:	:	:
rape	RPIKSLAVVLSWACLLTL	GAMKFLHWSNLFSSWK	GIALSALGLGIITLCMQILIRSSQSE			
	230	240	250	260	270	280

	360	370
maize	RSSSARAARNRVKKE	
	:	:
rape	RSTPAKVAPAKPKDN	
		290

FIG. 5(2/2)

REPLACED SHEETS	SG
☐ Description : paged to paged	
☒ Claim(s) : paged ...30... to paged ...33....	
☐ Drawing(s) : paged to paged	
☒ Abstract : paged ...33... to paged ...-.....	
DATE OF RECEIPT	

1.

REPLACEMENT SHEETS	SG
☒ Description : paged ...1. to paged ...21.	
☒ Claim(s) : paged ...30. to paged ...32.	
☒ Drawing(s) : paged ...1/8. to paged ...7/8.	
☒ Abstract : paged ...33 to paged ...-.	
DATE OF RECEIPT ...24/9/96.....	

DNA ENCODING 2-ACYLTRANSFERASES

5 This invention relates to modified plants. In particular, the invention relates to plants modified such that at least part of the plant (for example seeds of the plant) is capable of yielding a commercially useful oil.

10 Plants have long been a commercially valuable source of oil. Nutritional uses of plant-derived oils have hitherto been dominant, but attention is now turning additionally to plants as a source of industrially useful oils, for example as replacements for or improvements on mineral oils. Oil seeds, such as from rape, have a variety of lipids in them (Hildish & Williams, "Chemical
15 Composition of Natural Lipids", Chapman Hall, London, 1964). There is now considerable interest in altering lipid composition by the use of recombinant DNA technology (Knauf, TIBtech, February 1987, 40-47), but by
20 no means all of the goals have been realised to date for a variety of reasons, in spite of the ever-increasing sophistication of the technology.

25 Success in tailoring the lipid content of plant-derived oils requires a firm understanding of the biochemistry and genes involved. Broadly, two approaches are available. First, plants may be modified to permit the synthesis of fatty acids which are new (for the plant); so, for example, laurate and/or stearate may be synthesised in rape. Secondly, the pattern and/or extent
30 of incorporation of fatty acids into the glycerol backbone of the lipid may be altered. It is with this latter approach that the present invention is concerned, although the former approach may additionally be used.

This is the replacement sheet

Date: 24 SEP 1996

Lipids are formed in plants by the addition of fatty acid moieties onto the glycerol backbone by a series of acyl transferase enzymes. There are three positions on the glycerol molecule at which fatty acid (acyl) moieties may be substituted, and the substitution reached at each position is catalysed by a position-specific enzyme: the enzymes are known as 1-, 2- and 3-acyltransferases, respectively.

One, but not the only, current aim of "lipid engineering" in plants is to provide oils including lipids with a high content of erucic (22:1) acid. Erucic acid-containing lipids are commercially desirable for a number of purposes, particularly as replacements to or supplements for mineral oils in certain circumstances, as alluded to above. In the case of oil seed rape (*Brassica napus*), one of the most significant oil producing crops in cultivation today, the specificity of the 2-acyltransferase enzyme positively discriminates against the incorporation of erucic acid at position 2. So, even in those cultivars of rape which are able to incorporate erucic acid at positions 1 and 3, where there is no (or at least reduced) discrimination against erucic acid, only a maximum 66% of the fatty acids incorporated into triacyl glycerols can be erucic acid. Such varieties of rape are known as HEAR (high erucic acid rape) varieties.

It would therefore be desirable to increase the erucic acid content of conventional oil seed rape, as well as HEAR varieties; the same can be said of oils of other vegetable oil crops such as maize, sunflower and soya, to name but a few examples. While in principle it may be thought possible to introduce into a desired plant DNA encoding a 2-acyltransferase of different fatty acid

This is the replacement sheet

Date: 24 SEP 1996

specificity, for example from a different plant, in practice there are a number of problems.

5 First, 2-acyltransferase and 3-acyltransferase are
membrane bound, and therefore insoluble, enzymes. They
have not been purified. This makes working with them
difficult and rules out the use of many conventional DNA
10 cloning procedures. This difficulty does not,
paradoxically, lie in the way of cloning the gene (or at
least cDNA) encoding the 1-acyltransferase enzyme, which
is soluble: in fact, recombinant DNA work has already
been undertaken on this enzyme for a completely different
purpose, namely the enhancement of chilling resistance in
tobacco plant leaves, by Murata et al (Nature 356 710-713
15 (1992)).

Secondly, very little is known about the 2- and 3-
acyltransferases. There is no idea of their size or how
they are targeted to membranes. No nucleotide or amino
20 acid sequence data are available and no antibodies have
been raised against them.

Although there has been discussion, therefore, of the
desirability of modifying 2-acyltransferase specificity,
25 for example by importing a gene coding for the
corresponding enzyme, but of different specificity, from
another species, there is a pressing need in the art for
the key which enables this work to be done. The present
invention provides such a key, in the form of a DNA
30 sequence (in the specific case, a cDNA sequence) encoding
a 2-acyltransferase. The DNA sequence in Figure 1 from
nucleotides 130 to 1254 encodes the 2-acyltransferase
from maize (*Zea mays*), including the stop codon.

This is the replacement sheet

Date: 24 SEP 1996

According to a first aspect of the invention, therefore, there is provided a recombinant or isolated DNA sequence, preferably encoding an enzyme having membrane-bound acyltransferase activity, and selected from:

5

(i) a DNA sequence comprising the DNA sequence of Figure 1 encoding at least from MET₄₄ to Stop₄₁₈ (SEQ ID: 2) or its complementary strand,

10

(ii) nucleic acid sequences hybridising to the DNA sequence of Figure 1 (SEQ ID: 1), or its complementary strand, under stringent conditions, and

15

(iii) nucleic acid sequences which would hybridise to the DNA sequence of Figure 1 (SEQ ID: 1), or its complementary strand, but for the degeneracy of the genetic code.

20

Fragments of the above DNA sequences, for example of at least 15, 20, 30, 40 or 60 nucleotides in length, are also within the scope of the invention.

25

Suitable stringent conditions include salt solutions of approximately 0.9 molar at temperatures of from 35°C to 65°C. More particularly, stringent hybridisation conditions include 6 x SSC, 5 x Denhardt's solution, 0.5% SDS, 0.5% tetrasodium pyrophosphate and 50µg/ml denatured herring sperm DNA; washing may be for 2 x 30 minutes at 65°C in 1 x SSC, 0.1% SDS and 1 x 30 minutes in 0.2 x SSC, 0.1% SDS at 65°C.

30

Nucleic acid sequences within the scope of the first aspect of the invention will generally encode a protein

This is the replacement sheet

Date: 24 SEP 1996

having 2-acyltransferase activity, as that is the activity of the enzyme encoded by the Figure 1 nucleic acid sequence. Nucleic acid sequences not encoding a protein having enzymic activity (or the relevant enzymic activity) but otherwise conforming to the first aspect of the invention as set out above may be useful for other purposes (and are therefore also encompassed by the invention); for example they may be useful as probes, which is a utility shared by the nucleic acid sequences of the first aspect of the invention, including the Figure 1 sequence itself.

The probe utility arises as follows. As there is likely to be a high degree of homology between acyltransferases of different species (and particularly between 2-acyltransferases of different species) the Figure 1 sequence (or part of it, or other sequences within the invention) may be used to probe cDNA or genomic libraries of other species in order to clone DNA sequences encoding acyltransferases having desired specificities. For example, if it is desired to produce oil having a high content of erucic acid esterified to glycerol, a DNA library of any species which naturally makes erucic acid may be probed. Suitable plants include meadow foam (*Limnanthes* spp., especially *L. alba* and, particularly, *L. douglassi*) and *Crambe*. *Limnanthes douglassi* is the preferred species, as specificity studies show that there is positive discrimination towards incorporation of erucic acid into position 2 of the triacylglyceride. Libraries of organisms other than the higher plants may be probed; for example, certain bacteria may have an acyltransferase of the desired specificity.

This is the replacement sheet

Date: 2.4 SEP 1996

DNA in accordance with the invention will in general have a higher degree of homology with at least part of the Figure 1 sequence than with known sequences.

5 Recombinant DNA in accordance with the invention may be in the form of a vector, which may have sufficient regulatory sequences (such as a promoter) to direct expression. Vectors which are not expression vectors are useful for cloning purposes (as expression vectors themselves may be).
10 Host cells (such as bacteria and plant cells) containing vectors in accordance with the invention themselves form part of the invention.

15 DNA sequences in accordance with the invention can be used in another way in cloning a gene of interest from another species: if the DNA is coupled to a suitable promoter, for example on an expression vector in a suitable host organism, protein may be produced. Such protein may be used to generate polyclonal or monoclonal
20 antibodies, or other binding molecules, which may then be used to screen for expression of homologous proteins in other species, for example as part of a DNA library screening programme.

25 Suitable cDNA libraries of target species will generally be prepared when the gene of interest is likely to be expressed; so cDNA embryo libraries (prepared at the early lipid synthesis stage), for example of *Limnanthes* spp. will be preferred.

30

The invention therefore enables the cloning of a wide variety of genes (or, more generally, DNA sequences) encoding acyltransferases, and 2-acyltransferases in particular, using DNA sequences as described above.

This is the replacement sheet

Date: 24 SEP 1996

Such acyltransferases, such as from *Limnanthes* spp. may also be cloned directly, for example using complementation studies, from a DNA library of the species in question. For example, if *E. coli* is used as the complementation host, a mutant is chosen which is defective in the relevant enzyme (for example 2-acyltransferase); the DNA library from the target species (such as *L. douglassi*) is cloned into the mutant complementation host; host cells incorporating the target acyltransferase gene in their genome can readily be selected using appropriate selective media. *E. coli* mutant JC201 is a suitable host for use in complementation studies relating to 2-acyltransferase.

Cloning the acyltransferase gene of choice into a microbial host, such as a bacterium like *E. coli*, in such a way that the gene can be expressed has a particularly advantage in that the substrate specificity of the acyltransferase gene can be assessed in the microbial host before transformed plants are prepared, thereby saving considerably on research time. Such an assessment may be made by competitive substrate assays, in which differently detectably labelled candidate substrates for the enzyme compete with each other for incorporation into the glyceride. For example, ^{14}C -erucyl CoA and ^3H -oleoyl CoA can be used as competitive substrates for 2-acyltransferase, and the relative amounts of ^{14}C or tritium uptake into glyceride can be measured. (As 2-acyltransferases have acceptor, glycerol-based, substrates and donor, fatty acid-based, substrates, the experiment can be carried out with different acceptors, such as 1-erucyl-glycerol-3-phosphate and 1-oleoyl-glycerol-3-phosphate.) A gene coding for an enzyme which preferentially donates erucic acid to the acceptor

This is the replacement sheet

Date: 24 SEP 1996

(particularly 1-erucyl-glycerol-3-phosphate) may by this means be identified as a DNA sequence of choice for further use in the invention as described below.

5 In a second aspect of the invention, there is provided a plant having one or more insoluble acyltransferase enzymes having a substrate specificity which differs from the native enzyme of the plant.

10 While site-directed mutagenesis and/or other protein engineering techniques may be used to alter the specificity of an enzyme native to the plant, it is preferred that the plant be transgenic and incorporate an expressible acyltransferase gene encoding an enzyme of
15 the desired specificity from another species. 2-acyltransferases are the enzymes of choice. For example, as described above, a 2-acyltransferase enzyme which has an enhanced specificity for, or at least no discrimination against, erucic acid, may be made by this
20 means to express in a plant which would not normally incorporate erucic acid into triacylglycerides. An important embodiment of the invention relates to genetically engineered plants which have higher levels of erucic acid incorporated into triacylglycerols than in
25 corresponding non-engineered plants. Preferable though this embodiment may be, though, the invention is not limited to the enhancement of erucic acid incorporation into glycerides: other acids may be desired in other circumstances.

30

For the acyltransferase transgene to be expressible, a promoter has to be operatively coupled to it. Because at the present state of the art it is difficult precisely to regulate the site of incorporation of a transgene into

This is the replacement sheet

Date: 24 SEP 1996

the host genome, it is preferred that the transgene be coupled to its promoter prior to transformation of the plant. Promoters useful in the invention may be temporal- and/or seed-specific, but there is no need for them to be so: constitutive promoters, such as the CaMV 35S promoter, may in fact be preferred because they are usually strong promoters. Other tissues are unlikely to be adversely affected if the transgene encoding the acyltransferase enzyme is expressed in them, as the availability of the fatty acid CoA substrates is effectively limited to the seed.

The promoter-transgene construct, once prepared, is introduced into plant cells by any suitable means. The invention extends to such plant cells. Preferably, DNA is transformed into plant cells using a disarmed Ti-plasmid vector and carried by *Agrobacterium* by procedures known in the art, for example as described in EP-A-0116718 and EP-A-0270822. Alternatively, the foreign DNA could be introduced directly into plant cells using an electrical discharge apparatus. This method is preferred where *Agrobacterium* is ineffective, for example where the recipient plant is monocotyledonous. Any other method that provides for the stable incorporation of the DNA within the nuclear DNA of any plant cell of any species would also be suitable. This includes species of plant which are not currently capable of genetic transformation.

Preferably DNA in accordance with the invention also contains a second chimeric gene (a "marker" gene) that enables a transformed plant or tissue culture containing the foreign DNA to be easily distinguished from other plants or tissue culture that do not contain the foreign

This is the replacement sheet

Date: 24 SEP 1996

DNA. Examples of such a marker gene include antibiotic resistance (Herrera-Estrella et al, *EMBO J.* 2(6) 987-95 (1983) and Herrera-Estrella et al, *Nature* 303 209-13 (1983)), herbicide resistance (EP-A-0242246) and glucuronidase (GUS) expression (EP-A-0344029). Expression of the marker gene is preferably controlled by a second promoter which allows expression in cells other than the tapetum, thus allowing selection of cells or tissue containing the marker at any stage of regeneration of the plant. The preferred second promoter is derived from the gene which encodes the 35S subunit of Cauliflower Mosaic Virus (CaMV) coat protein. However any other suitable second promoter could be used.

A whole plant can be regenerated from a single transformed plant cell, and the invention therefore provides transgenic plants (or parts of them, such as propagating material) including DNA in accordance with the invention as described above. The regeneration can proceed by known methods.

In one embodiment of the invention, the transgenic plant's native acyltransferase gene which corresponds to the transgene may be rendered at least partially inoperative or removed. So, if the transgene encodes a 2-acyltransferase, the plant's native 2-acyltransferase may be rendered inoperative by, for example, antisense or ribozyme techniques, as is known in the art.

By means of the invention, plants generating oil with a tailored lipid content may be produced. For example, the lipid composition of triacylglycerides in a plant may be substantially altered to produce triacylglycerides with a desired fatty acid (for example erucic acid) content

This is the replacement sheet

Date: 24 SEP 1996

higher than has hitherto been possible. For example, oil seed rape (*B. napus*) may be transformed to produce oil whose triacylglyceride has an erucic acid content of over 70%.

5

It can readily be seen that plants with increased lipid levels may be produced by means of the invention. However, the invention is also useful for producing plants with decreased lipid levels, which may be desired if elevated protein and/or starch levels are required. Decreased lipid levels may be achieved by interfering with the proper functioning of a gene encoding a 2-acyltransferase, for example by antisense or ribozyme technology. (Such reduced-lipid plants may if desired be further engineered for higher protein and/or starch content, if wished.)

10
15

Promoters which naturally drive 2-acyltransferases may also be obtained by hybridisation and/or restriction and/or sequencing studies using the Figure 1 sequence.

20

The invention enables the production of protein encoded by DNA of the first aspect of the invention, should that be desired. The protein may be expressed by host cells harbouring DNA in the form of an expression vector. The protein, which may be an enzyme having 2-acyltransferase activity, may have an amino acid sequence which is identical to or homologous with the Figure 1 sequence. The degree of homology will generally be greater than that of known proteins, and may be at least 40, 50, 60, 70, 80, 90, 95 or 99%.

25
30

Preferred features of each aspect of the invention are as for each other aspect *mutatis mutandis*.

This is the replacement sheet

Date: 2.4 SEP 1996

The invention is illustrated by the following examples. The examples refer to the accompanying drawings, in which:

5 FIGURE 1 shows the cDNA sequence derived in Example 1 (SEQ ID: 1) and its derived protein sequence (SEQ ID: 2).

10 FIGURE 2 shows a sequence alignment of part of the gene products of *plsB* (SEQ ID: 3) and *plsC* (SEQ ID: 4) with part of the sequence shown in Figure 1 (SEQ ID: 5), showing a conserved motif. *plsB* is the *E. coli* sn-glycerol-3-phosphate acyltransferase gene and *plsC* the 1-acyl-sn-glycerol-3-phosphate acyltransferase gene of *E. coli*. Double points indicate exact matches between two sequences and a single point conservative amino acid substitutions. Stars indicate identical amino acids in all three sequences and residues conserved in two out of the
15 three sequences are marked by a + symbol.
20

25 FIGURES 3A, 3B and 3C: Membrane phospholipids from *E. coli* strains were extracted into chloroform and separated by 2-dimensional thin layer chromatography. The first dimension (ascending) was developed using chloroform:methanol:water (65:25:4) and the second dimension (left to right) developed with chloroform:methanol:acetic acid (65:25:10). Phospholipids were visualised by autoradiography for
30 16 hours at -70°C using Fuji RX film. The *E. coli* strains used were (Figure 3A): JC201 which carries a thermosensitive mutation in the 1-acyl-sn-glycerol-3-phosphate acyltransferase gene; (Figure 3B): JC201 containing the plasmid pPLSC, which

This is the replacement sheet

24 SEP 1996

Date:

encodes the *E. coli* 1-acyl-sn-glycerol-3-phosphate acyltransferase gene; (Figure 3C): JC201 containing the plasmid whose cDNA insert sequence is shown in Figure 1. LPA, lysophosphatidic acid; PE, phosphatidylethanolamine; CL, cardiolipin; PG, phosphatidylglycerol; PA, phosphatidic acid; O, origin. 20% of the ^{32}P is incorporated in LPA in JC201 and all of the corresponding label is incorporated in PE in both of the other two strains.

FIGURE 4: Acyltransferase assays were performed using ^{32}P -labelled lysophosphatidic acid which had been extracted from the *E. coli* strain JC201 and oleoyl CoA as an acyl donor. Phospholipids present in the reaction mixtures were extracted into chloroform and separated using silica gel thin layer chromatography. Chloroform:methanol:acetic acid:water (25:15:4:2) was used to develop the plates. The phospholipids were visualised by autoradiography for 16 hours at -70°C using Fuji RX film. The *E. coli* strains used were: JC201 which carries a thermosensitive mutation in the 1-acyl-sn-glycerol-3-phosphate acyltransferase gene; JC201 containing the plasmid pPLSC which encodes the *E. coli* 1-acyl-sn-glycerol-3-phosphate acyltransferase gene; JC201 containing the plasmid whose maize cDNA insert sequence is shown in Figure 1. LPA, lysophosphatidic acid; PA, phosphatidic acid.

FIGURE 5 shows a comparison of the protein sequence shown in Figure 1 (SEQ ID: 6) with that derived (SEQ ID: 7) from a *B. napus* seed cDNA insert which was isolated by DNA hybridisation to the maize cDNA sequence. The sequences were aligned with the FastA

This is the replacement sheet

Date: 24 SEP 1996

align program (1988). Double points signify identical amino acids and single points conservative amino acid substitutions.

5 Maize = 374 aa vs. rape = 311 aa
51.5% identity; Optimised score: 705

EXAMPLE 1

Derivation of the Figure 1 cDNA sequence

10 Complementation studies using a maize cDNA expression library transferred into the *E. coli* mutant JC201 allowed the isolation of a plasmid encoding a 2-acyltransferase enzyme from maize. The cDNA insert of this plasmid is 1.6kb in size, and includes a poly A tail of 70bp. The
15 insert was sequenced to give the data shown and translation of the sequence revealed the presence of only one large open reading frame. This is shown on Figure 1 with proposed start methionine and stop codon boxed. The 2-acyltransferase is 374 amino acids in size and
20 sequencing upstream of open reading from showed that the protein is expressed as part of a fusion protein in *E. coli*. This consists of 10 amino acids of the β -galactosidase protein, 43 amino acids (shown in sequence) corresponding to the 5' untranslated region of the mRNA and the 374 amino acid protein. Protein sequence
25 comparisons of the large open reading frame with the 2-acyltransferase of *E. coli* show little overall identity but there is a stretch of 80 residues which has a high level of conservative substitution and contains some
30 amino acids that are conserved in the 2-acyltransferase, 1-acyltransferase and N-acetyl glucosamine acyltransferase of *E. coli*.

This is the replacement sheet

Date: 24 SEP 1996

EXAMPLE 2**Incorporation of ^{32}P into total phospholipids**

E. coli strains were grown in minimal medium containing ^{32}P orthophosphate. Total glycerolipids were extracted into organic solvents and separated by 2D thin layer chromatography (Figure 3) (Lysophosphatidic acid (LPA) is the substrate for 2-acyltransferase (2-AT)).

As can be seen in Figure 3A, the accumulation of ^{32}P -labelled LPA in the mutant JC 201 illustrates the absence of a fully functional 2-AT. Addition of a plasmid carrying either the native *E. coli* gene (Figure 3B), or the maize clone given in Figure 1 (Figure 3C) restores 2-AT activity to the cells, allowing LPA to be removed and further metabolised. (Lysophosphatidic acid.)

These data indicate that the DNA sequence given in Figure 1 codes for 2-AT.

EXAMPLE 3**Over expression of the cDNA**

The cDNA region specifying the protein sequence given in Figure 1 was cloned into the *E. coli* overexpression vector pET11d (Studier et al, Meth. Enzymol. 185 60-89 (1990)). Increased 2-acyltransferase activity following induction of expression from the plasmid insert confirmed that the sequence in Figure 1 is that of 2-AT.

EXAMPLE 4**Localisation of 2-AT activity in *E. coli* cells containing the maize clone**

2-acyltransferase assays were carried out using membranes isolated from the mutant strain JC.201 which lacks 2-AT and from JC.201 containing the maize plasmid (Figure 4).

This is the replacement sheet

Date: 24 SEP 1996

2-AT activity was not detected in membrane fractions from JC.201. The addition of a plasmid carrying the native *E. coli* gene or the sequence given in Figure 1, to JC.201 resulted in restoration of 2-AT activity to the membranes.

EXAMPLE 5

Using the maize cDNA as a heterologous probe to obtain cDNA from oilseed rape

A seed cDNA library from *Brassica napus* was screened with the sequence given in Figure 1, using standard techniques (Sambrook et al "Molecular Cloning - A Laboratory Manual", 2nd Edition, Cold Spring Harbor Laboratory Press, 1989).

Conditions

For the hybridisation of the maize cDNA insert to the rape library: hybridisation was in 6xSSC, 5xDenhardt's solution, 0.5%SDS, 0.5% tetrasodium pyrophosphate and 50ugml⁻¹ denatured herring sperm DNA. The filters were washed 2 x 30 minutes at 65°C in 1xSSC, 0.1%SDS and 1 x 30 minutes in 0.2xSSC, 0.1%SDS at 65°C.

A hybridising clone was sequenced and a protein sequence derived for the large ORF. Alignment of this protein sequence with that derived from the maize cDNA clone given in Figure 1, is shown in Figure 5.

The strong identity between these sequences illustrates the potential of using the sequence given in Figure 1 to obtain other 2-ATS.

This is the replacement sheet

Date: 24 SEP 1996

EXAMPLE 6**Transgenic plants**

The sequence given in Figure 1 can be cloned, alongside a suitable promoter, into a suitable vector for expression in plants. The vector can be used to transform plants and the resulting plants expressing the 2-AT can be analysed for lipid content. Lipid metabolism is expected to be upregulated and elevated lipid levels were detectable in seeds.

EXAMPLE 7**Antisense**

The sequence given in Figure 1 may be cloned, alongside a suitable promoter, in the antisense orientation into a suitable vector for expression in plants. The vector can be used to transform plants and the resulting plants expressing the 2-AT can be analysed for protein and starch content. Elevated levels of starch and protein are expected to be detectable in seeds.

EXAMPLE 8**Down-regulation of Native 2-AT**

The DNA sequence of a 2-AT derived from *L. douglassii* (obtained as described in Example 5) can be introduced into oilseed rape (OSR) under the expression of a suitable promoter, using vectors and plant transformation methods well known in the art. A second sequence, comprising antisense or ribozymes against the rape cDNA (Example 5) can be introduced for simultaneous expression. The resultant transformed plant is expected to have 2-AT activity corresponding to that of *L. douglassii*, with concurrent down regulation of the native rape 2-AT gene.

This is the replacement sheet

Date: 24 SEP 1996

The modified OSR plant thus obtained had higher levels of erucic acid in position 2 of its triacylglycerols than wild type plants. In addition higher levels of trierucin are found in the seed oil.

5

EXAMPLE 9**Genomic library screening**

10

The sequence given in Figure 1 is used to screen a genomic library of *Arabidopsis* and a hybrid using clone obtained. Using standard techniques, a promoter may be derived from this clone. The promoter may be used to drive expression in plant cell membranes.

This is the replacement sheet

Date: 24 SEP 1996

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

(A) NAME: Nickerson BIOCEM Limited
(non-US only)
(B) STREET: Cambridge Science Park, Milton Road
(C) CITY: Cambridge
(E) COUNTRY: United Kingdom
(F) POSTAL CODE (ZIP): CB4 4GZ

(A) NAME: SLABAS, Antoni Ryszard
(US only)
(B) STREET: 8 Telford Close, High Shincliffe
(C) CITY: Durham
(E) COUNTRY: United Kingdom
(F) POSTAL CODE (ZIP): DH1 2YJ

(A) NAME: BROWN, Adrian Paul
(US only)
(B) STREET: 4 Church Villas, Shadforth
(C) CITY: Durham
(E) COUNTRY: United Kingdom
(F) POSTAL CODE (ZIP): DH6 1LQ

(ii) TITLE OF INVENTION: DNA ENCODING 2-ACYLTRANSFERASES

(iii) NUMBER OF SEQUENCES: 7

(iv) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

(v) CURRENT APPLICATION DATA:

APPLICATION NUMBER: WO PCT/GB93/_____

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1514 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION: 130..1254

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

CCCCGTCCCTC CTCGTGCGCCG GCGGAGCCGC CTACTATCGC CTGGAGAAGG AGCGCCGCGG	60
GGAGCTTTTC CCACTGCCGA CTGCCGTCTG ACCCTCCGAG ATCGGAAGCG GCGCCGGCGC	120
CGGCCGGCG ATG GCG ATC CCG CTC GTG CTC GTC GTG CTC CCG CTC GGC	168
Met Ala Ile Pro Leu Val Leu Val Val Leu Pro Leu Gly	
1 5 10	

This is the replacement sheet

Date: 24 SEP 1996

CTG	CTC	TTC	CTC	CTG	TCC	GGC	CTC	ATC	GTC	AAC	GCC	ATC	CAG	GCC	GTC	216
Leu	Leu	Phe	Leu	Leu	Ser	Gly	Leu	Ile	Val	Asn	Ala	Ile	Gln	Ala	Val	
15						20					25					
CTA	TTT	GTG	ACG	ATA	AGG	CCC	TTT	TCG	AAG	AGC	TTC	TAC	CGT	CGG	ATC	264
Leu	Phe	Val	Thr	Ile	Arg	Pro	Phe	Ser	Lys	Ser	Phe	Tyr	Arg	Arg	Ile	
30					35					40					45	
AAC	AGA	TTC	TTG	GCC	GAG	CTG	CTG	TGG	CTT	CAG	CTT	GTC	TGG	GTG	GTG	312
Asn	Arg	Phe	Leu	Ala	Glu	Leu	Leu	Trp	Leu	Gln	Leu	Val	Trp	Val	Val	
				50				55						60		
GAC	TGG	TGG	GCA	GGT	GTT	AAG	GTA	CAA	CTG	CAT	GCA	GAT	GAG	GAA	ACT	360
Asp	Trp	Trp		Gly	Val	Lys	Val	Gln	Leu	His	Ala	Asp	Glu	Glu	Thr	
			65					70					75			
TAC	AGA	TCA	ATG	GGT	AAA	GAG	CAT	GCA	CTC	ATC	ATA	TCA	AAT	CAT	CGG	408
Tyr	Arg	Ser	Met	Gly	Lys	Glu	His	Ala	Leu	Ile	Ile	Ser	Asn	His	Arg	
	80					85						90				
AGT	GAT	ATT	GAT	TGG	CTC	ATT	GGA	TGG	ATA	TTG	GCC	CAG	CGT	TCA	GGG	456
Ser	Asp	Ile	Asp	Trp	Leu	Ile	Gly	Trp	Ile	Leu	Ala	Gln	Arg	Ser	Gly	
95						100					105					
TGC	CTT	GGA	AGT	ACA	CTT	GCT	GTC	ATG	AAG	AAG	TCA	TCC	AAG	TTC	CTT	504
Cys	Leu	Gly	Ser	Thr	Leu	Ala	Val	Met	Lys	Lys	TCA	TCC	AAG	TTC	CTT	
110					115						120				125	
CCA	GTT	ATT	GGC	TGG	TCA	ATG	TGG	TTT	GCA	GAG	TAC	CTC	TTT	TTG	GAA	552
Pro	Val	Ile	Gly	Trp	Ser	Met	Trp	Phe	Ala	Glu	Tyr	Leu	Phe	Leu	Glu	
				130					135					140		
AGG	AGC	TGG	GCC	AAG	GAT	GAA	AAG	ACA	CTA	AAG	TGG	GGT	CTC	CAA	AGG	600
Arg	Ser	Trp	Ala	Lys	Asp	Glu	Lys	Thr	Leu	Lys	Trp	Gly	Leu	Gln	Arg	
			145					150					155			
TTG	AAA	GAC	TTC	CCT	AGA	CCA	TTT	TGG	CTA	GCT	CTT	TTC	GTC	GAG	GGT	648
Leu	Lys	Asp	Phe	Pro	Arg	Pro	Phe	Trp	Leu	Ala	Leu	Phe	Val	Glu	Gly	
		160					165					170				
ACT	CGC	TTT	ACT	CCA	GCA	AAG	CTT	CTC	GCA	GCT	CAG	GAA	TAT	GCG	GCC	696
Thr	Arg	Phe	Thr	Pro	Ala	Lys	Leu	Leu	Ala	Ala	Gln	Glu	Tyr	Ala	Ala	
	175					180					185					
TCC	CAG	GGC	TTA	CCG	GCT	CCT	AGA	AAT	GTA	CTT	ATT	CCA	CGT	ACC	AAG	744
Ser	Gln	Gly	Leu	Pro	Ala	Pro	Arg	Asn	Val	Leu	Ile	Pro	Arg	Thr	Lys	
190					195					200					205	
GGA	TTT	GTA	TCT	GCT	GTA	AGT	ATT	ATG	CGA	GAT	TTT	GTT	CCA	GCC	ATT	792
Gly	Phe	Val	Ser	Ala	Val	Ser	Ile	Met	Arg	Asp	Phe	Val	Pro	Ala	Ile	
				210					215					220		
TAT	GAT	ACA	ACT	GTA	ATA	GTC	CCT	AAA	GAT	TCC	CCT	CAA	CCA	ACA	ATG	840
Tyr	Asp	Thr	Thr	Val	Ile	Val	Pro	Lys	Asp	Ser	Pro	Gln	Pro	Thr	Met	
				225				230					235			
CTG	CGG	ATT	TTG	AAA	GGG	CAA	TCA	TCA	GTG	ATA	CAT	GTC	CGC	ATG	AAA	888
Leu	Arg	Ile	Leu	Lys	Gly	Gln	Ser	Ser	Val	Ile	His	Val	Arg	Met	Lys	
		240					245					250				
CGT	CAT	GCA	ATG	AGT	GAG	ATG	CCA	AAA	TCA	GAT	GAG	GAT	GTT	TCA	AAA	936
Arg	His	Ala	Met	Ser	Glu	Met	Pro	Lys	Ser	Asp	Glu	Asp	Val	Ser	Lys	
	255					260					265					

This is the replacement sheet

Date: 24 SEP 1996

TGG	TGT	AAA	GAC	ATT	TTT	GTG	GCA	AAG	GAT	GCC	TTA	CTG	GAC	AAG	CAT	984
Trp	Cys	Lys	Asp	Ile	Phe	Val	Ala	Lys	Asp	Ala	Leu	Leu	Asp	Lys	His	
270					275					280					285	
TTG	GCA	ACA	GGC	ACT	TTC	GAT	GAG	GAG	ATT	AGA	CCT	ATT	GGC	CGT	CCA	1032
Leu	Ala	Thr	Gly	Thr	Phe	Asp	Glu	Glu	Ile	Arg	Pro	Ile	Gly	Arg	Pro	
				290					295					300		
GTG	AAA	TCA	TTG	CTG	GTG	ACC	CTG	TTC	TGG	TCG	TGC	CTC	CTG	CTG	TTT	1080
Val	Lys	Ser	Leu	Leu	Val	Thr	Leu	Phe	Trp	Ser	Cys	Leu	Leu	Leu	Phe	
			305					310					315			
GGC	GCC	ATC	GAG	TTC	TTC	AAG	TGG	ACA	CAG	CTT	CTG	TCG	ACG	TGG	AGG	1128
Gly	Ala	Ile	Glu	Phe	Phe	Lys	Trp	Thr	Gln	Leu	Leu	Ser	Thr	Trp	Arg	
			320				325					330				
GGT	GTG	GCG	TTC	ACT	GCC	GCA	GGG	ATG	GCG	CTT	GTG	ACG	GGT	GTC	ATG	1176
Gly	Val	Ala	Phe	Thr	Ala	Ala	Gly	Met	Ala	Leu	Val	Thr	Gly	Val	Met	
	335					340					345					
CAT	GTC	TTC	ATC	ATG	TTC	TCC	CAG	GCT	GAG	CGG	TCG	AGC	TCA	GCC	AGG	1224
His	Val	Phe	Ile	Met	Phe	Ser	Gln	Ala	Glu	Arg	Ser	Ser	Ser	Ala	Arg	
350					355					360					365	
GCG	GCA	CGG	AAC	CGG	GTC	AAG	AAG	GAA	TGAAAAATGG	AGGGTGGAGA						1271
Ala	Ala	Arg	Asn	Arg	Val	Lys	Lys	Glu								
				370				375								
TGAGGTTCTC	GTGGGGTTTG	TTATGGGCAA	CCTTCAAAG	GACTCTCCAT	TCATATTAGT											1331
ATTAATTCAT	ATATATGCAG	CGCCAAATTC	CAGACATTGA	TATGCTCTCA	AATAGGATGT											1391
TCTGCTCCCC	TCTTGTATTT	GTATGCAGGA	AAGGGTTTGT	AGGGAGTTTA	CCCCCCCCCC											1451
CCCCCCCCCC	GCCTTTCTTT	GGGGAAGAAA	GACATATTCT	GGAAGCCTTC	CAGTAGTTCA											1511
AAA																1514

This is the replacement sheet

Date: 24 SEP 1996

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 374 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```

Met Ala Ile Pro Leu Val Leu Val Val Leu Pro Leu Gly Leu Leu Phe
 1             5             10             15
Leu Leu Ser Gly Leu Ile Val Asn Ala Ile Gln Ala Val Leu Phe Val
          20             25             30
Thr Ile Arg Pro Phe Ser Lys Ser Phe Tyr Arg Arg Ile Asn Arg Phe
          35             40             45
Leu Ala Glu Leu Leu Trp Leu Gln Leu Val Trp Val Val Asp Trp Trp
          50             55             60
Ala Gly Val Lys Val Gln Leu His Ala Asp Glu Glu Thr Tyr Arg Ser
          65             70             75             80
Met Gly Lys Glu His Ala Leu Ile Ile Ser Asn His Arg Ser Asp Ile
          85             90             95
Asp Trp Leu Ile Gly Trp Ile Leu Ala Gln Arg Ser Gly Cys Leu Gly
          100            105            110
Ser Thr Leu Ala Val Met Lys Lys Ser Ser Lys Phe Leu Pro Val Ile
          115            120            125
Gly Trp Ser Met Trp Phe Ala Glu Tyr Leu Phe Leu Glu Arg Ser Trp
          130            135            140
Ala Lys Asp Glu Lys Thr Leu Lys Trp Gly Leu Gln Arg Leu Lys Asp
          145            150            155            160
Phe Pro Arg Pro Phe Trp Leu Ala Leu Phe Val Glu Gly Thr Arg Phe
          165            170            175
Thr Pro Ala Lys Leu Leu Ala Ala Gln Glu Tyr Ala Ala Ser Gln Gly
          180            185            190
Leu Pro Ala Pro Arg Asn Val Leu Ile Pro Arg Thr Lys Gly Phe Val
          195            200            205
Ser Ala Val Ser Ile Met Arg Asp Phe Val Pro Ala Ile Tyr Asp Thr
          210            215            220
Thr Val Ile Val Pro Lys Asp Ser Pro Gln Pro Thr Met Leu Arg Ile
          225            230            235            240
Leu Lys Gly Gln Ser Ser Val Ile His Val Arg Met Lys Arg His Ala
          245            250            255
Met Ser Glu Met Pro Lys Ser Asp Glu Asp Val Ser Lys Trp Cys Lys
          260            265            270
Asp Ile Phe Val Ala Lys Asp Ala Leu Leu Asp Lys His Leu Ala Thr
          275            280            285

```

This is the replacement sheet

Date: 24 SEP 1996

Gly	Thr	Phe	Asp	Glu	Glu	Ile	Arg	Pro	Ile	Gly	Arg	Pro	Val	Lys	Ser
290						295					300				
Leu	Leu	Val	Thr	Leu	Phe	Trp	Ser	Cys	Leu	Leu	Leu	Phe	Gly	Ala	Ile
305					310					315					320
Glu	Phe	Phe	Lys	Trp	Thr	Gln	Leu	Leu	Ser	Thr	Trp	Arg	Gly	Val	Ala
				325					330					335	
Phe	Thr	Ala	Ala	Gly	Met	Ala	Leu	Val	Thr	Gly	Val	Met	His	Val	Phe
			340					345					350		
Ile	Met	Phe	Ser	Gln	Ala	Glu	Arg	Ser	Ser	Ser	Ala	Arg	Ala	Ala	Arg
		355					360					365			
Asn	Arg	Val	Lys	Lys	Glu										
370															

This is the replacement sheet

Date: 24 SEP 1996

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

Tyr	Phe	Val	Glu	Gly	Gly	Arg	Ser	Arg	Thr	Gly	Arg	Leu	Leu	Asp
1				5					10					15

This is the replacement sheet

Date: 24 SEP 1996

(2) INFORMATION FOR SEQ ID NO: 4:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 15 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

Met	Phe	Pro	Glu	Gly	Thr	Arg	Ser	Arg	Gly	Arg	Gly	Leu	Leu	Pro
1				5					10					15

This is the replacement sheet

Date: 24 SEP 1996

(2) INFORMATION FOR SEQ ID NO: 5:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 15 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

Leu Phe Val Glu Gly Thr Arg Phe Thr Pro Ala Lys Leu Leu Ala
 1 5 10 15

(2) INFORMATION FOR SEQ ID NO: 6:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 374 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

Met Ala Ile Pro Leu Val Leu Val Val Leu Pro Leu Gly Leu Leu Phe
 1 5 10 15
 Leu Leu Ser Gly Leu Ile Val Asn Ala Ile Gln Ala Val Leu Phe Val
 20 25 30
 Thr Ile Arg Pro Phe Ser Lys Ser Phe Tyr Arg Arg Ile Asn Arg Phe
 35 40 45
 Leu Ala Glu Leu Leu Trp Leu Gln Leu Val Trp Val Val Asp Trp Trp
 50 55 60
 Ala Gly Val Lys Val Gln Leu His Ala Asp Glu Glu Thr Tyr Arg Ser
 65 70 75 80
 Met Gly Lys Glu His Ala Leu Ile Ile Ser Asn His Arg Ser Asp Ile
 85 90 95
 Asp Trp Leu Ile Gly Trp Ile Leu Ala Gln Arg Ser Gly Cys Leu Gly
 100 105 110
 Ser Thr Leu Ala Val Met Lys Lys Ser Ser Lys Phe Leu Pro Val Ile
 115 120 125
 Gly Trp Ser Met Trp Phe Ala Glu Tyr Leu Phe Leu Glu Arg Ser Trp
 130 135 140
 Ala Lys Asp Glu Lys Thr Leu Lys Trp Gly Leu Gln Arg Leu Lys Asp
 145 150 155 160
 Phe Pro Arg Pro Phe Trp Leu Ala Leu Phe Val Glu Gly Thr Arg Phe
 165 170 175
 Thr Pro Ala Lys Leu Leu Ala Ala Gln Glu Tyr Ala Ala Ser Gln Gly
 180 185 190

This is the replacement sheet

Date: 24 SEP 1996

Leu Pro Ala Pro Arg Asn Val Leu Ile Pro Arg Thr Lys Gly Phe Val
 195 200 205
 Ser Ala Val Ser Ile Met Arg Asp Phe Val Pro Ala Ile Tyr Asp Thr
 210 215 220
 Thr Val Ile Val Pro Lys Asp Ser Pro Gln Pro Thr Met Leu Arg Ile
 225 230 235 240
 Leu Lys Gly Gln Ser Ser Val Ile His Val Arg Met Lys Arg His Ala
 245 250 255
 Met Ser Glu Met Pro Lys Ser Asp Glu Asp Val Ser Lys Trp Cys Lys
 260 265 270
 Asp Ile Phe Val Ala Lys Asp Ala Leu Leu Asp Lys His Leu Ala Thr
 275 280 285
 Gly Thr Phe Asp Glu Glu Ile Arg Pro Ile Gly Arg Pro Val Lys Ser
 290 295 300
 Leu Leu Val Thr Leu Phe Trp Ser Cys Leu Leu Leu Phe Gly Ala Ile
 305 310 315 320
 Glu Phe Phe Lys Trp Thr Gln Leu Leu Ser Thr Trp Arg Gly Val Ala
 325 330 335
 Phe Thr Ala Ala Gly Met Ala Leu Val Thr Gly Val Met His Val Phe
 340 345 350
 Ile Met Phe Ser Gln Ala Glu Arg Ser Ser Ser Ala Arg Ala Ala Arg
 355 360 365
 Asn Arg Val Lys Lys Glu
 370

This is the replacement sheet

Date: 24 SEP 1996

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 295 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

```

Met Ala Met Ala Ala Ala Val Ile Val Pro Leu Gly Ile Leu Phe Phe
1           5           10           15
Ile Ser Gly Leu Val Val Asn Leu Leu Gln Arg Ser Gly Cys Leu Gly
20           25           30
Ser Ala Leu Ala Val Met Lys Lys Ser Ser Lys Phe Leu Pro Val Ile
35           40           45
Gly Trp Ser Met Trp Phe Ser Glu Tyr Leu Phe Leu Glu Arg Asn Trp
50           55           60
Ala Lys Asp Glu Ser Thr Leu Lys Ser Gly Leu Gln Arg Leu Asn Asp
65           70           75           80
Phe Pro Arg Pro Phe Trp Leu Ala Leu Phe Val Glu Gly Thr Arg Phe
85           90           95
Thr Glu Ala Lys Leu Lys Ala Ala Gln Glu Tyr Ala Ala Ser Ser Glu
100          105          110
Leu Pro Val Pro Arg Asn Val Leu Ile Pro Arg Thr Lys Gly Phe Val
115          120          125
Ser Ala Val Ser Asn Met Arg Ser Phe Val Pro Ala Ile Tyr Asp Met
130          135          140
Thr Val Ala Ile Pro Lys Thr Ser Pro Pro Pro Thr Met Leu Arg Leu
145          150          155          160
Phe Lys Gly Gln Pro Ser Val Val His Val His Ile Lys Cys His Ser
165          170          175
Met Lys Asp Leu Pro Glu Ser Glu Asp Glu Ile Ala Gln Trp Cys Arg
180          185          190
Asp Gln Phe Val Thr Lys Asp Ala Leu Leu Asp Lys His Ile Ala Ala
195          200          205
Asp Thr Phe Ala Gly Gln Lys Glu Gln Asn Ile Gly Arg Pro Ile Lys
210          215          220
Ser Leu Ala Val Val Leu Ser Trp Ala Cys Leu Leu Thr Leu Gly Ala
225          230          235          240
Met Lys Phe Leu His Trp Ser Asn Leu Phe Ser Ser Trp Lys Gly Ile
245          250          255
Ala Leu Ser Ala Leu Gly Leu Gly Ile Ile Thr Leu Cys Met Gln Ile
260          265          270
Leu Ile Arg Ser Ser Gln Ser Glu Arg Ser Thr Pro Ala Lys Val Ala
275          280          285

```

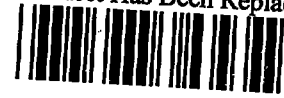
This is the replacement sheet

Date: 24 SEP 1996

Pro Ala Lys Pro Lys Asp Asn
290 295

This is the replacement sheet

Date: 24 SEP 1996

CLAIMS

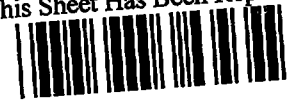
- 5 1. A recombinant or isolated DNA sequence selected from
- (i) a DNA sequence comprising the DNA sequence of Figure 1 encoding at least from Met₄₄ to Stop₄₁₈ (SEQ ID: 2) or its complementary strand;
- 10 (ii) nucleic acid sequences hybridising to the DNA sequence of Figure 1 (SEQ ID: 1) or its complementary strand under stringent conditions; and
- 15 (iii) nucleic acid sequences which would hybridise to the DNA sequence of Figure 1 (SEQ ID: 1), or its complementary strand, but for the degeneracy of the genetic code.
- 20 2. A DNA sequence as claimed in claim 1 which encodes an enzyme having membrane-bound acyltransferase activity.
3. A DNA sequence as claimed in claim 2, in which the acyltransferase activity is 2-acyltransferase activity.
- 25 4. An isolated protein which is the expression product of a DNA sequence as claimed in claim 1, 2 or 3.
- 30 5. An antibody capable of specifically binding to a protein as claimed in claim 4.
6. A plant having one or more insoluble acyltransferase enzymes having a substrate specificity which differs from the native enzyme of the plant.



* C L M S *

This is the replacement sheet

Date: 24 SEP 1996



7. A plant as claimed in claim 6, which is transgenic for an insoluble acyltransferase enzyme from another species.

5 8. A plant as claimed in claim 6 or 7, which is rape, maize, sunflower or soya.

9. A plant as claimed in claim 7, wherein the said other species is a species of the genus *Limnanthes* (such
10 as *L. alba* or, preferably, *L. douglassi*) or *Crambe*.

10. A plant as claimed in any one of claims 6 to 9, wherein the insoluble acyltransferase enzyme is 2-acyltransferase.

15 11. A plant as claimed in claim 10, wherein the 2-acyltransferase has a higher specificity for erucic acid than the native enzyme of the plant.

20 12. A plant as claimed in any one of claims 6 to 11, wherein the said native enzyme is at least partially rendered inoperative or removed, for example by a ribozyme or by antisense nucleic acid.

25 13. A method of generating oil, the method comprising cultivating a plant as claimed in anyone of claims 6 to 12 and harvesting oil produced by the plant or a part (particularly seeds) thereof.

30 14. A microbial host capable of expressing a DNA sequence as claimed in claim 1, 2 or 3.

15. A protein which is substantially homologous to a protein as claimed in claim 4.





16. A fragment of a DNA sequence as claimed in claim 1, 2 or 3, comprising at least 15 nucleotides.
- 5 17. DNA encoding RNA which is antisense to RNA encoded by DNA as claimed in claim 1, 2 or 3.
18. DNA encoding a ribozyme specific to RNA encoded by DNA as claimed in claim 1, 2 or 3.
- 10 19. Isolated or recombinant DNA containing a promoter which naturally drives expression of a gene to produce a protein as claimed in claim 4 or 15.



This is the replacement sheet

Date: 24 SEP 1996

ABSTRACTDNA ENCODING 2-ACYLTRANSFERASES

5 Plants, particularly transgenic plants, may be produced
 having a 2-acyltransferase enzyme or other insoluble
 acyltransferase enzyme with an altered substrate
 specificity compared to the native enzyme. For example,
 oil seed rape (*Brassica napus*) may contain a 2-
10 acyltransferase transgene derived from *Limnanthes*
 douglassi in order to increase the erucic acid content of
 the oil. The cDNA sequence of maize (*Zea mays*) 2-
 acyltransferase is disclosed and is useful for cloning
 acyltransferase genes and/or cDNAs from other organisms,
15 including *L. douglassi*.

(No appropriate figure)

This Sheet Has Been Replaced



This is the replacement sheet

Date: 24 SEP 1996

FIG. 1 (1/3)

1/1	CCC CGT CCT CCT CGT CGC CGG CGG AGC CGC CTA TCG CCT GGA GAA GGA GCG CCG CGG	31/11
61/21	P R P P R R R R S R L L S P G E G A P R	
121/41	GGA GCT TTT CCC ACT GCC GAC TGC CGT CTG ACC CTC CGA GAT CGG AAG CGG CGC CGG CGC	91/31
181/61	G A F P T A D C R L T L R D R K R R R	
241/81	CGG CCG GCG ATG GCG ATC CCG CTC GTG CTC CCG CTC GGC CTC GTG CTC TTC CTC	151/51
301/101	R P A M A I P L V L V V L P L G L L F L	
361/121	181/61	211/71
421/141	CTG TCC GGC CTC ATC GTC AAC GCC ATC CAG GCC GTC CTA TTT GTG ACG ATA AGG CCC TTT	
481/161	L S G L I V N A I Q A V L F V T I R P F	
541/181	241/81	271/91
601/201	TCG AAG AGC TTC TAC CGT CGG ATC AAC AGA TTC TTG GCC GAG CTG CTG TGG CTT CAG CTT	
661/221	S K S F Y R R I N R F L A E L L W L Q L	
721/241	301/101	331/111
781/261	GTC TGG GTG GAC TGG TGG GCA GGT GTT AAG GTA CAA CTG CAT GCA GAT GAG GAA ACT	
841/281	V W V V D W W A G V K V Q L H A D E E T	
901/301	361/121	391/131
961/321	TAC AGA TCA ATG GGT AAA GAG CAT GCA CTC ATC ATA TCA AAT CAT CGG AGT GAT ATT GAT	
1021/341	Y R S M G K E H A L I I S N H R S D I D	
1081/361	421/141	451/151
1141/381	TGG CTC ATT GGA TGG ATA TTG GCC CAG CGT TCA GGG TGC CTT GGA AGT ACA CTT GCT GTC	
1201/401	W L I G W I L A Q R S G C L G S T L A V	
1261/421	481/161	511/171
1321/441	ATG AAG AAG TCA TCC AAG TTC CTT CCA GTT ATT GGC TGG TCA ATG TGG TTT GCA GAG TAC	
1381/461	M K K S S K F L P V I G W S M W F A E Y	
1441/481	541/181	571/191
1501/501	CTC TTT TTG GAA AGG AGC TGG GCC AAG GAT GAA AAG ACA CTA AAG TGG GGT CTC CAA AGG	
1561/521	L F L E R S W A K D E K T L K W G L Q R	

This is the replacement sheet

Date: 24 SEP 1996

FIG. 1
(2/3)

601/201	TTG AAA	GAC TTC	CCT AGA	CCA TTT	TGG CTA	GCT CTT	631/211	TTC GTC	GAG GGT	ACT ACT	CGC TTT	ACT
L K	D	F	P	P	F	W	L	F	V	E	G	T
661/221	CCA GCA	AAG CTT	CTC GCA	GCT CAG	GAA TAT	GCG GCC	691/231	TCC CAG	GGC TTA	CCG GCT	CCT AGA	AGA
P A	K	L	L	A	A	Q	E	S	Q	G	L	P
721/241	AAT GTA	CTT ATT	CCA CGT	ACC AAG	GGA TTT	GTA TCT	751/251	GCT GTA	AGT ATT	ATG ATG	CGA GAT	TTT
N V	L	I	P	R	T	K	G	A	V	S	I	M
781/261	GTT CCA	GCC ATT	TAT GAT	ACA ACT	GTA ATA	GTC CCT	811/271	AAA GAT	TCC CCT	CAA CCA	ACA ATG	ATG
V P	A	I	Y	D	T	T	V	K	D	S	P	Q
841/281	CTG CGG	ATT TTG	AAA GGG	CAA TCA	TCA GTG	ATA CAT	871/291	GTC CGC	ATG AAA	CGT CAT	GCA ATG	ATG
L R	I	L	K	G	Q	S	S	V	R	M	K	R
901/301	AGT GAG	ATG CCA	AAA TCA	GAT GAT	GAT GTT	TCA AAA	931/311	TGG TGT	AAA GAC	ATT TTT	GTG GCA	GCA
S E	M	P	K	S	D	E	D	V	C	K	D	I
961/321	AAG GAT	GCC TTA	CTG GAC	AAG CAT	TTG GCA	ACA GGC	991/331	ACT TTC	GAT GAG	GAG ATT	AGA CCT	CCT
K D	A	L	L	D	K	H	L	A	T	F	D	E
1021/341	ATT GGC	CGT CCA	GTG AAA	TCA TTG	CTG GTG	ACC CTG	1051/351	TTT TGG	TGC TGC	CTC CTC	CTG CTG	TTT
I G	R	P	V	K	S	L	L	V	T	L	F	S
1081/361	GGC GCC	ATC GAG	TTC TTC	AAG TGG	ACA CAG	CTT CTG	1111/371	TCG ACG	TGG AGG	GGT GTG	GCG TTC	TTC
G A	I	E	F	F	K	W	T	Q	L	S	T	W
												R
												G
												V
												A
												F

This is the replacement sheet

Date: 24 SEP 1961

1141/381
 ACT GCC GCA GGG ATG GCG CTT GTG ACG GGT GTC ATG CAT GTC TTC ATC ATG TTC TCC CAG
 T A A G M A L V T G V M H V F I M F S Q
 1201/401
 GCT GAG CGG TCG AGC TCA GCC AGG GCG GCA CGG AAC CGG GTC AAG AAG GAA TGA AAA ATG
 A E R S S S A R A A R N R V K K E * K M
 1261/421
 GAG GGT GGA GAT GAG GTT CTC GTG GGG TTT GTT ATG GGC AAC CTT CAA AAG GAC TCT CCA
 E G G D E V L V G F V M G N L Q K D S P
 1321/441
 TTC ATA TTA GTA TTA ATT CAT ATA TAT GCA GCG CCA AAT TCC AGA CAT TGA TAT GCT CTC
 F I L V L I H I Y A A P N S R H * Y A L
 1381/461
 AAA TAG GAT GTT CTG CTC CCC TCT TGT ATT TGT ATG CAG GAA AGG GTT TGT AGG GAG TTT
 K * D V L L P S C I C M Q E R V C R E F
 1441/481
 ACC CCC CCC CCC CCC TTT CTT TGG GGA AGA AAG ACA TAT TCT GGA AGC CTT
 T P P P P P A F L W G R K T Y S G S L
 1501/501
 CCA GTA GTT CAA AA
 P V V Q

FIG. 1 (3/3)

This is the replacement sheet

Date: 24 SEP 1996

FIG. 2

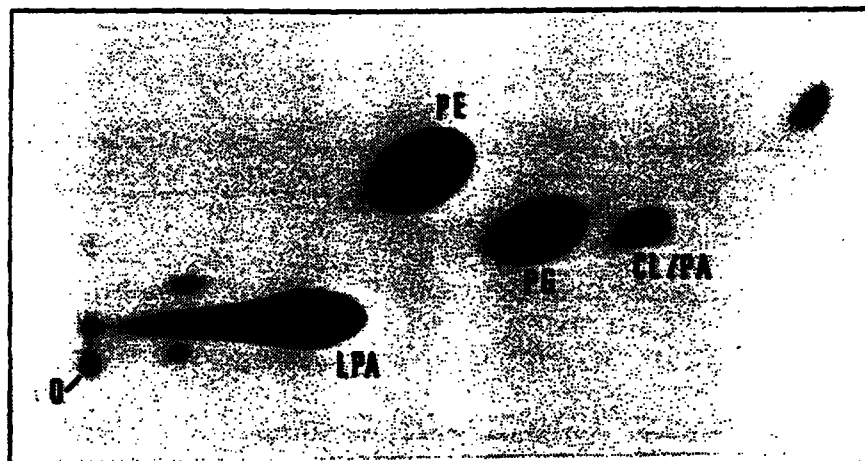
plsb	-	Y	F	V	E	G	G	R	S	R	T	G	R	L	L	D	-
			.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
plsc	-	M	F	P	E	G	T	R	S	R	G	R	G	L	L	P	-
			.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
maize	-	L	F	V	E	G	T	R	F	T	P	A	K	L	L	A	-
			*	+	*	*	+	*	+	+				*	*	*	*

This is the replacement sheet

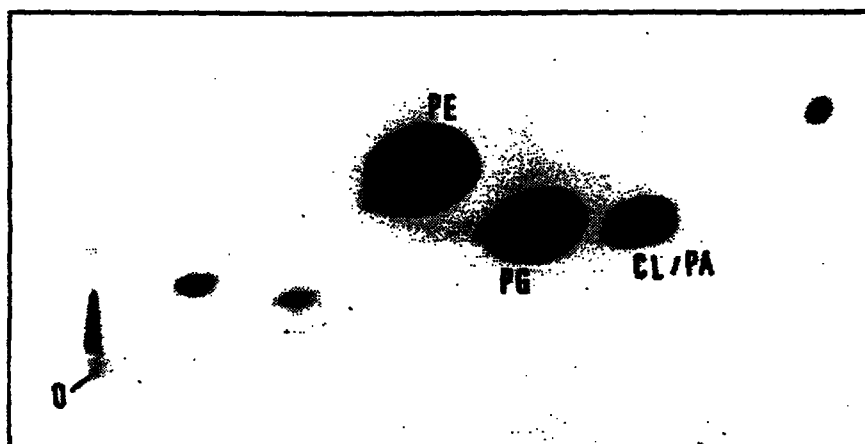
Date: 24 SEP 1996

5 / 8
FIG. 3

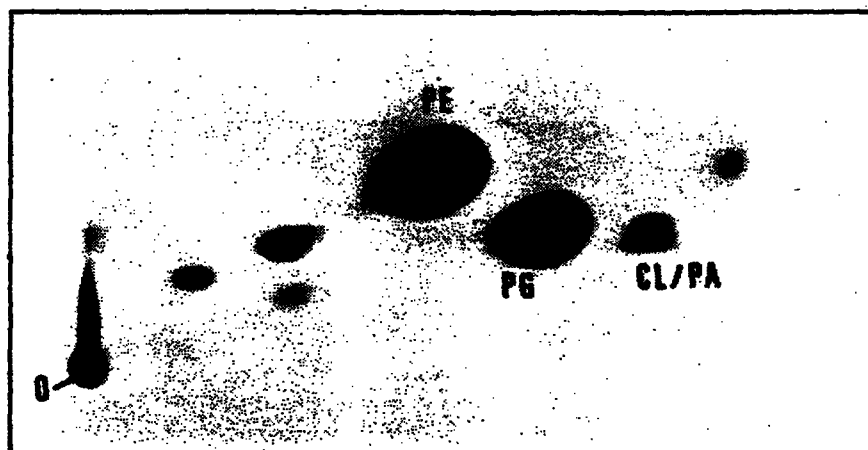
A : mutant
JC201



B : mutant +
pPLSC



C : mutant +
maize
cDNA



This is the replacement sheet

Date: 24 SEP 1996

FIG. 4

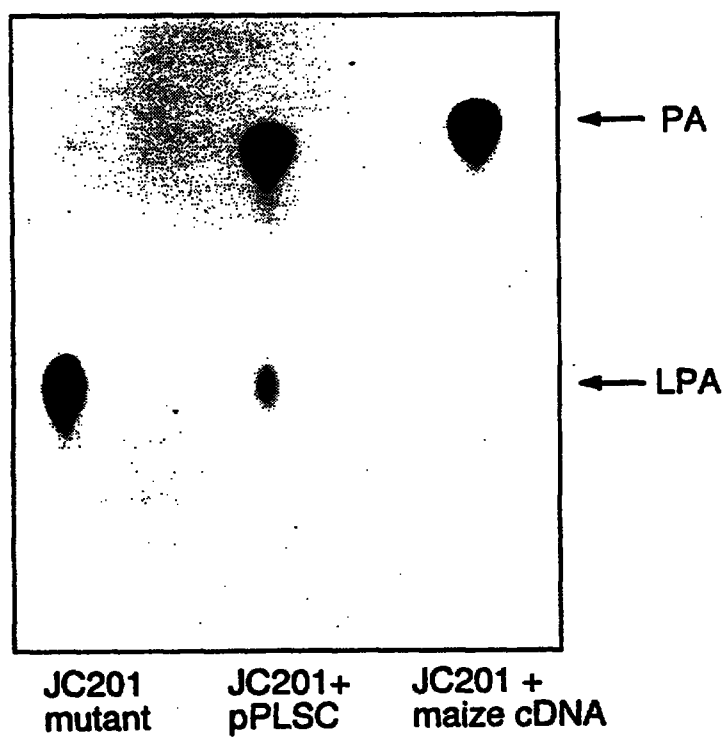


FIG. 5 (1/2)

Maize = 374 aa vs.rape = 311 aa

51.5% identity; Optimized score: 705

	10	20	30	40	50	60
maize	MAIPVLVVLPLGLLGLISGLIVNAIQAVLEFVTIRPFSSKSFYRRINRFLAELLWLQLVWV					

.....

rape	MAMA-AAVIVPLGILFFISGLVN-----					
	10	20				

	70	80	90	100	110	120
maize	VDWWAGVKVQLHADEETYSRMGKEHALIISNHRSDIDWLIGWILAQRSGLGSTLAVMKK					

.....

rape	-----LLQSGCLGSALAVMKK					
				30		40

	130	140	150	160	170	180
maize	SSKFLPVIGWSMWFELERSWAKDETKLKWGLQRLKDFPRFWLALFVEGTRFTPAK					

.....

rape	50	60	70	80	90	100
	SSKFLPVIGWSMWFSEYLFLEARNWAKDESTLKSGLQRLNDFPRFWLALFVEGTRFTPAK					

maize	190	200	210	220	230	240
	LLAAQEYAA	SOGLPAPRN	VLIPRTKGF	VSASIMRDF	VPAIYDTT	VIVPKDSPQPTMLRI
rape	110	120	130	140	150	160
	LKAAQEYAA	SSELVPVRN	VLIPRTKGF	VSASVSNMRS	FVPAIYDM	TVAIPKTSPPPTMLRL
maize	250	260	270	280	290	
	LKGQSSV	IHVMRKR	HAMSEMPK	SDVDVSK	WCKDIFVAK	DALLDKHLATGTF-DEEIRPIG
rape	170	180	190	200	210	220
	FKGQPSV	VHVHIKCH	SMKDLPE	SEDEIAQ	WCRDQF	VTKDALLDKHIAADTFAGQKEQNIG
maize	300	310	320	330	340	350
	RPVKSL	LVTLFWS	CLLLFGA	IEFFKWT	QLLSTWR	GVAFTAAGMALVTGVMHVFIMFSQAE
rape	230	240	250	260	270	280
	RPIKSL	AVVLSW	ACLLTL	GAMKFL	HWSNLF	SWSKGIALSALGLGITLTCMQILIRSSQSE
maize	360	370				
	RSSSARA	ARNRVKKE				
rape						
	RSTPAK	VAPAKPKDN				
		290				

FIG. 5(2/2)

This is the replacement sheet

Date:

24 SEP 1996

THE CLAIMS

REPLACEMENT SHEETS SG
☐ Description : paged to paged
☒ Claim(s) : paged 32. to paged 31.
☐ Drawing(s) : paged to paged
☒ Abstract : paged 32. to paged 31.

DATE OF RECEIPT 24/7/98

1. A recombinant or isolated DNA sequence selected from
 - (i) a DNA sequence including the DNA sequence of Figure 1 encoding at least from Met₄₄ to Stop₄₁₈ (SEQ ID: 2) or its complementary strand;
 - (ii) nucleic acid sequences hybridising to the DNA sequence of Figure 1 (SEQ ID: 1) or its complementary strand under stringent conditions; and
 - (iii) nucleic acid sequences which would hybridise to the DNA sequence of Figure 1 (SEQ ID: 1), or its complementary strand, but for the degeneracy of the genetic code.
2. A DNA sequence as claimed in claim 1 which encodes an enzyme having membrane-bound acyltransferase activity.
3. A DNA sequence as claimed in claim 2, in which the acyltransferase activity is 2-acyltransferase activity.
4. An isolated protein which is the expression product of a DNA sequence as claimed in claim 3.
5. An antibody capable of specifically binding to the protein as claimed in claim 4.
6. A transformed plant having one or more insoluble 2-acyltransferase enzymes having a substrate specificity which differs from the native enzyme of the plant.
7. A plant as claimed in claim 6, which is transgenic for an insoluble acyltransferase enzyme from another species.
8. A plant as claimed in claim 6 or claim 7, which is rape, maize, sunflower or soya.
9. A plant as claimed in claim 7, wherein the said other species is a species of the genus *Linnanthes* (such as *L. alba*, or preferably, *L. douglassi*) or *Crambe*.



This is the replacement sheet

Date: 24 SEP 1998

10. A plant as claimed in any one of claims 6 to 9, wherein the 2-acyltransferase has a higher specificity for erucic acid than the native enzyme of the plant.
11. A plant as claimed in any one of claims 6 to 10, wherein the said native enzyme is at least partially rendered inoperative or removed, for example by a ribozyme or by antisense nucleic acid.
12. A method of generating oil, the method including cultivating a plant as claimed in any one of claims 6 to 11 and harvesting oil produced by the plant or a part (particularly seeds) thereof.
13. A microbial host including a DNA sequence as claimed in any one of claims 1 to 3.
14. A protein which is substantially homologous to a protein as claimed in claim 4.
15. A fragment of a DNA sequence as claimed in any one of claims 1 to 3, including at least 15 nucleotides.
16. DNA encoding RNA which is antisense to RNA encoded by DNA as claimed in any one of claims 1 to 3.
17. DNA encoding a ribozyme specific to RNA encoded by DNA as claimed in any one of claims 1 to 3.
18. Isolated or recombinant DNA containing a promoter which naturally drives expression of a gene to produce a protein as claimed in claim 4 or claim 14.



This is the replacement sheet

Date:

24 SEP 1998

ABSTRACTDNA ENCODING 2-ACYLTRANSFERASES

5 Plants, particularly transgenic plants, may be produced
having a 2-acyltransferase enzyme or other insoluble
acyltransferase enzyme with an altered substrate
specificity compared to the native enzyme. For example,
oil seed rape (*Brassica napus*) may contain a 2-
10 acyltransferase transgene derived from *Limnanthes*
douglassi in order to increase the erucic acid content of
the oil. The cDNA sequence of maize (*Zea mays*) 2-
acyltransferase is disclosed and is useful for cloning
acyltransferase genes and/or cDNAs from other organisms,
15 including *L. douglassi*.

This is the replacement sheet

Date: 2 4 SEP 1998