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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
15 May 2003 (15.05.2003)

PCT

(10) International Publication Number
WO 03/040297 A2

(51) International Patent Classification[7]: C12N

(21) International Application Number: PCT/EP02/13206

(22) International Filing Date:
5 November 2002 (05.11.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
PI 2001 5098 6 November 2001 (06.11.2001) MY

(71) Applicants *(for all designated States except US)*: **CENTRE DE COOPERATION INTERNATIONALE EN RECHERCHE AGRONOMIQUE POUR LE DEVELOPPEMENT (CIRAD)** [FR/FR]; 42, rue Scheffer, F-75116 Paris (FR). **MALAYSIAN PALM OIL BOARD** [MY/MY]; 6, Persiaran Institusi, Bandar Baru Bangi, Selangor, 43000 Kajang (MY).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **CHEAH, Suan, Choo** [MY/MY]; Persiaran Institusi, Bandar Baru Bangi, Selangor, 43000 Kajang (MY). **TREGEAR, James, W.** [GB/FR]; Cirad, CP/IRD Oil Palm Biotechnology Laboratory, Centre IRD de Montpellier, BP 5045, 911 Avenue Agropolis, F-34032 Montpellier (FR).

(74) Agents: **VIALLE-PRESLES, Marie José** et al.; Cabinet Ores, 6, avenue de Messine, F-75008 Paris (FR).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 03/040297 A2

(54) Title: NEW PROTEIN OF THE DEFENSIN FAMILY

(57) Abstract: The invention relates to new protein of the defensin family, its isolated polypeptide and an isolated polynucleotide encoding it. The invention further relates to the uses of the polynucleotide and or polypeptide, including its use to detect the mantled abnormality of oil palm.

NEW PROTEIN OF THE DEFENSIN FAMILY

The present invention relates to a new protein of the defensin family, which is useful in particular as a molecular marker of somaclonal variation events associated with the *mantled* homeotic flowering abnormality of oil palm.

Oil palm (*Elaeis guineensis* Jacq.) is a diploid monocotyledonous temporally dioecious species producing alternate cycles of male and female flowers.

In vitro micropropagation based on somatic embryogenesis has been used to carry out the multiplication of elite oil palm genotypes as described by RIVAL *et al.*, Plan Tiss. Cult. Biotechnol., 3, 74-83, 1996; however, the large scale use of this approach is hampered by the appearance of a somaclonal variant, known as *mantled* as described by CORLEY *et al.*, The Planter, 62, 233-240, 1986, which affects approximately 5-10% of palms cloned in this way as illustrated again by RIVAL *et al.*, Plan Tiss. Cult. Biotechnol., 3, 74-83, 1996). The *mantled* phenotype involves a feminisation of both male and female flowers: in the former, stamens develop as carpelloid structures, whilst in the latter, the staminodes (vestigial stamens) develop as pseudocarpel structures. In both cases, the petals appear to develop as sepalloid structures. In severe cases the flowers are sterile, although lesser affected female flowers may be fertilised to give *mantled* fruit.

In order to allow a more widespread use of *in vitro* micropropagation for oil palm, it is necessary to improve its reliability, by establishing an early detection method for the *mantled* abnormality. To this end, a detailed study has been carried out in order to identify molecular markers of the *mantled* phenotype.

Previous work revealed that the *mantled* abnormality is epigenetic in nature. Firstly, it was observed that reversion to a normal floral phenotype may occur in the field as described by DURAND-GASSELIN T. et al., Oléagineux 45, 1-11, 1990. Secondly, although the *mantled* abnormality is strongly transmitted through tissue culture, only a weak non-Mendelian transmission occurs via seeds as described by RAO V. and DONOUGH C.R., Elaeis, 2 199-207, 1990. Thirdly,

ploidy, RAPD and AFLP studies have shown that the abnormal phenotype is unlikely to have resulted from any gross changes in genomic structure as described by RIVAL *et al.*, Plant Cell. Rep., 16, 884-887, 1997; RIVAL et al., Plant Breeding, 117, 73-76, 1998.

The present invention is based on the cloning and characterisation of a new oil palm gene, hereinafter referred as *EGAD1*. The *EGAD1* gene codes for a 77 amino acid polypeptide displaying strong similarities with plant defensins.

Northern studies of *EGAD1* gene expression show that in the intact oil palm plant, *EGAD1* transcripts accumulate mostly in inflorescence tissues, both normal and *mantled*, and show a peak of abundance during the early stages of development of the male inflorescence. On the other hand, in tissue culture, *EGAD1* transcripts are observed to accumulate throughout the regeneration procedure used for oil palm, a higher level of expression being observed at the nodular callus stage in cultures carrying the *mantled* abnormality than in those lacking it.

Thus, the expression of *EGAD1* provides an early marker of the *mantled* phenotype.

The present invention relates to an isolated polypeptide of the defensin family, wherein said polypeptide, hereinafter referred as EGAD1, has at least 70% sequence identity, preferably at least 80% sequence identity, and more preferably at least 90% sequence identity, with the polypeptide of SEQ ID NO: 2 in the enclosed sequence listing. The percent sequence identity is based on the entire region SEQ ID NO: 2 and is determined according to the CLUSTAL W program as illustrated by THOMPSON *et al.*, Nucl. Acids. Res., 22, 4673-4680, 1994, using the default settings.

The invention also comprises fragments of an EGAD1 polypeptide of the invention, comprising at least 5, preferably 10 contiguous amino-acids of said polypeptide.

The term "polypeptide" herein refers to a polymer of amino acid residues, resulting from the translation of a nucleic acid sequence, and possibly having post-translational

modifications such as glycosylation, lipid attachment, hydroxylation, etc...

The present invention also relates to an isolated polynucleotide selected among:

a) a polynucleotide encoding an EGAD1 polypeptide; this includes for instance *EGAD1* genes, *EGAD1* cDNA, or *EGAD1* mRNA. A particular example is the cDNA of SEQ ID NO: 1;

b) a fragment of a polynucleotide a) at least 10, preferably at least 15, and more preferably at least 20 contiguous nucleotides thereof ;

c) a polynucleotide complementary to a polynucleotide a) or b) above.

The invention also comprises polynucleotides hybridising selectively under stringent conditions with any polynucleotide of the invention. This includes in particular nucleic acid probes for detecting a nucleic acid sequence encoding an EGAD1 polypeptide, as well as nucleic acid primers for amplifying a nucleic acid sequence encoding an EGAD1 polypeptide, or a portion thereof.

The term "stringent conditions" refers to conditions under which a polynucleotide will hybridize to its target sequence to a detectably greater degree than to other sequences (e.g., at least 2-fold, preferably at least 5-fold over background). Stringent conditions can easily be defined by one of skill in the art for a polynucleotide of given length and base composition.

According to a preferred embodiment, a nucleic acid probe of the invention consists essentially of a polynucleotide of at least 10, preferably at least 15, and more preferably at least 20 nucleotides hybridising selectively under stringent conditions with a nucleic acid of SEQ ID NO: 1. Said probe may further comprise modifications, such as the addition of an appropriate label allowing its detection.

According to a preferred embodiment, a set of PCR primers of the invention consists of a fragment of at least 10, preferably at least 15, and more preferably at least 20 contiguous nucleotides of the polynucleotide of SEQ ID NO: 1,

and a fragment of at least 10, preferably at least 15, and more preferably at least 20 contiguous nucleotides of the complementary of SEQ ID NO: 1. Appropriate primers can easily be designed by one of skill in the art from the sequence SEQ ID NO: 1.

The invention also encompasses any polynucleotide obtained from an oil palm genomic or cDNA library by PCR amplification with a set of primers of the invention.

The invention also comprises:

- a recombinant expression cassette comprising a polynucleotide of the invention encoding an EGAD1 polypeptide operably linked to a promoter;
- a recombinant nucleic acid vector having an insert consisting of a polynucleotide or of a recombinant expression cassette of the invention;
- a host cell transfected with a polynucleotide or a recombinant expression cassette of the invention;
- a transgenic plant comprising a recombinant expression cassette of the present invention.

The invention also provides a process for producing an EGAD1 polypeptide of the invention, wherein said process comprises culturing a host cell transfected with a recombinant expression cassette of the invention, and recovering the EGAD1 polypeptide from said culture. Said polypeptide can be obtained from the cell lysate or from the culture medium using standard protein isolation techniques, in particular techniques suitable for isolation of defensins.

The invention also comprises polyclonal or monoclonal antibodies specifically directed against an EGAD1 polypeptide of the invention.

The present invention also relates to the use of a polypeptide or a polynucleotide of the invention, or of antibodies of the invention, for detecting the *mantled* phenotype in oil palm tissue, or in an oil palm tissue culture.

More specifically, the present invention provides a method for detecting the *mantled* abnormality in an oil palm tissue culture, wherein said method comprises detecting the

overexpression of the *EGAD1* gene encoding the EGAD1 polypeptide in said culture.

According to a preferred embodiment of the invention, the overexpression of the *EGAD1* gene is detected in an oil palm tissue culture at the callus stage.

The overexpression of the *EGAD1* gene is herein defined as a level of expression of said gene higher than the base-line expression level in a normal oil palm tissue culture.

Generally, the level of expression of the *EGAD1* gene in oil palm tissue culture bearing the *mantled* abnormality is at least 2 times higher, preferably at least 3 times higher, and more preferably at least 4 times higher, than its level of expression in a normal oil palm tissue culture.

The level of expression of the *EGAD1* gene can be determined by evaluating the amount of *EGAD1* transcripts, or the amount of EGAD1 polypeptide.

Techniques allowing the detection of an increased production of a mRNA or a protein are well known in the art.

An increased production of mRNA may for instance be detected using the technique of northern blot analysis. An increased production of EGAD1 polypeptide may for instance be detected by immunoassay using antibodies of the invention directed against said polypeptide.

The invention also encompasses reagents for practising the method of the invention. These reagents include in particular nucleic acid probes and sets of PCR primers as defined above.

The *EGAD1* gene and the EGAD1 polypeptide can also be used in the same way as previously described for other plant defensins (THEVISSEN *et al.*, J. Biol. Chem. 271, 15018-15025, 1996; GAO et al. Nature Biotech., 18, 1307-1310 (2000); US Patent 6,121,436; (TERRAS *et al.*, Plant Cell, 7, 573-588, 1995)), for the purpose of protecting plants against pathogens, in particular microorganisms such as fungi.

The use of the *EGAD1* gene and the EGAD1 polypeptide to protect plants against pathogens can be achieved by standard methods, for instance by transforming

the plant to be protected with a sequence encoding an EGAD1 polypeptide, placed under the transcriptional control of an appropriate promoter.

The present invention will be further illustrated by the additional description which follows, which refers to the isolation of the *EGAD1* cDNA and the use of *EGAD1* sequences for detecting the *mantled* phenotype. It should be understood however that these examples are given only by way of illustration of the invention and do not constitute in any way a limitation thereof.

EXAMPLES

General methods

Plant material

Oil palm tissue cultures were established and maintained as previously described by PANNETIER *et al.* (Oléagineux, 36, 119-122, 1981). Two different clonal lines, LAB146 and LAB147, were used for differential display analysis. The LAB146 culture was obtained by direct cloning of a seed-derived palm and may therefore be assumed to carry little or no *mantled* abormality (i.e. it should generate either a low percentage of *mantled* regenerants or only normal clonal progeny). The LAB147 culture was obtained by recloning a *mantled* tissue culture-derived palm and can therefore be assumed to produce 100% abnormal regenerants as illustrated by RIVAL *et al.*, (Plant Tiss. Cult. Biotechnol., 3, 74-83, 1996). For northern analysis, a range of cultures was used. Tissues were harvested corresponding to 3 developmental stages in the regeneration process: nodular callus, somatic embryos and 1 cm shoot apex-containing segments excised from leafy shoots. Inflorescence material was obtained from of seed-derived and tissue culture-regenerated oil palms grown in Côte d'Ivoire and Malaysia. Root and leaf material was harvested from seed-derived greenhouse plants. For DNA extraction, leaf material from adult palms of several different genotypes (see below) was used.

Extraction and analysis of RNA and DNA

All standard cloning procedures used in this study were carried out essentially as described by (SAMBROOK

et al., Molecular Cloning: A laboratory manual, Cold Spring Harbour Laboratory Press, Cold Spring Harbour, USA, 1989) unless otherwise indicated. Total RNA was extracted as described previously by CORRE *et al.*, (L. Plant Sci., 117, 139-149, 1996). RNA gel electrophoresis and northern transfer were carried out using a NorthernMax-Gly® kit (AMBION CORPORATION). Membranes were hybridised with either ^{32}P-radiolabelled RNA probes synthesised using a Strip-EZ® RNA probe kit (AMBION CORPORATION) or ^{32}P-radiolabelled DNA probes obtained using the random priming method (FEINBERG and VOGELSTEIN, Analytical Biochemistry, 132, 6-13, 1983). Genomic DNA was extracted as previously described by RIVAL *et al.*, (Plant Breeding, 117, 73-76, 1998) and analysed by Southern blotting and hybridisation using standard techniques.

Differential display analysis

Differential display analysis was carried out using the primers and PCR conditions described by MALHOTRA *et al.* (Nucl. Acids Res.,26, 854-856, 1998) with [α-^{33}P]dATP as the radiolabel. After denaturing polyacrylamide gel electrophoresis, differential bands were excised and re-amplified by conventional PCR using the same primers as were used for the differential display amplifications. Re-amplified DNAs were blunt end cloned into the *Eco*RV site of the pBluescript SK- phagemid® (STRATAGENE). Several clones were sequenced for each marker re-amplification reaction and when more than one cDNA sequence was found to be cloned for a given marker, each cDNA was tested separately.

Construction and screening of cDNA library

A cDNA library was constructed from 1 cm shoot apex-containing segments excised from leafy shoots of the LAB147 culture grown for 3 weeks on medium containing 10-5 M benzylaminopurine (BAP). Screening was carried out by plaque hybridisation at 60°C in a buffer containing 5 x Denhardt's solution, 6xSSC, 4 µg/ml sheared salmon sperm DNA and 0.5% SDS.

EXAMPLE 1: CHARACTERISATION BY DIFFERENTIAL DISPLAY OF A CDNA FRAGMENT OVEREXPRESSED IN *MANTLED*-BEARING TISSUE CULTURES

Differences in mRNA accumulation between shoot apex-containing segments of the LAB146 (normal seed palm-derived) and LAB147 (abnormal clone-derived) cultures were studied using differential display, enabling the identification of differential bands specific to normal or abnormal material which were subsequently excised, re-amplified and cloned.

One of the abnormal-specific cDNA fragments obtained, provisionally named m5B, was shown by Northern hybridisation to correspond to an mRNA of ca. 0.6 kb which accumulates at higher levels in LAB147 shoot apex-containing segments than in those obtained from the LAB146 culture, as shown in Figure 1.

Figure 1 shows a northern hybridisation analysis of transcript accumulation using a ^{32}P-radiolabelled RNA probe synthesised from the m5B differential display clone containing part of the *EGAD1* cDNA sequence. The m5B (*EGAD1*) probe was hybridised to total RNA extracted from oil palm tissue cultures and organs of intact plants (upper panel). As a reference, a control hybridisation to oil palm ^{32}P-radiolabelled 25S ribosomal DNA probe was also performed (lower panel). Two different tissue cultures were examined: LAB146, a normal seed palm-derived culture (labelled "SPDC") and LAB147, an abnormal (*mantled*) clone-derived culture (labelled "ACDC"). Other abbreviations are as follows:

- INFLO, spikelets of female inflorescences;
- N, normal inflorescence (25 cm in length);
- A, abnormal inflorescence (30 cm in length).

In a preliminary characterisation of the tissue specificity of the m5B mRNA, it was found that the difference in accumulation of the 0.6 kb transcript between the LAB146 (normal) and LAB147 (abnormal) cultures was even more marked for calli than for shoot apex-containing segments. Little or no difference in signal intensity was observed for somatic embryos.

In whole plants, the 0.6 kb transcript was found to be present in both normal and abnormal female

inflorescences (25 cm and 30 cm respectively in length) harvested from clonal palms, similar levels being observed in each case.

No m5B-specific signal was observed for leaves or roots of intact seedling plants, although a longer exposure of the autoradiograph shown produced a faint signal for roots.

Overall, it can be concluded from the Northern hybridisation shown in Figure 1 that transcripts corresponding to the m5B cDNA accumulate principally in inflorescence tissues in the intact plant, as well as in tissue cultures in an apparently *mantled*-dependent fashion.

EXAMPLE 2: ISOLATION AND CHARACTERISATION OF THE FULL LENGTH *EGAD1* CDNA

The cDNA insert in the m5B clone was used as a probe to screen an oil palm cDNA library prepared from shoot apex-containing segments of leafy shoots grown on BAP-containing medium, allowing the purification and sequencing of 6 positive clones containing the previously determined m5B nucleotide sequence. A database search revealed that each of the cDNAs thus obtained codes for a cysteine-rich polypeptide sharing close similarities with the plant defensin (or γ-thionin) proteins (TERRAS *et al.*, Plant Cell, 7, 573-588, 1995). Although slight variations in size were observed between the cloned cDNAs, only one (m5B-15) displayed a divergent nucleotide sequence, the others all being base identical. Since the m5B-15 cDNA diverges from the other 5 cDNAs at only 3 nucleotide positions, we believe that it might represent a different allele of the same gene locus. This hypothesis is consistent with our genomic Southern data (see Example 3 below). The longest cloned cDNA obtained, m5B-7, was selected for use in subsequent experiments and renamed *EGAD1*. The nucleotide sequence of the 535 bp *EGAD1* cDNA and the deduced sequence of the polypeptide for which it codes are represented respectively as SEQ ID NO: 1 and SEQ ID NO: 2 and are also shown in Figure 2.

Figure 2 also shows the nucleotide sequence of the *EGAD1* (m5B-7) cDNA and the deduced sequence of the encoded polypeptide. The position of the putative signal

peptide sequence is indicated by underlining. The position of the putative translation termination codon is denoted by an asterisk.

EGAD1 codes for a 77 amino acid polypeptide of predicted size 5.3 kD. The *EGAD1* cDNA sequence shown in Figure 2 is assumed to contain a full length coding region, since it includes a putative translation initiation codon which is conserved amongst plant defensins and which is preceded in the same reading frame by a TAG stop codon 12 bp further upstream. The polypeptide encoded by the *EGAD1* gene contains a putative signal peptide sequence of 30 amino acids which is likely to be responsible for targeting the nascent polypeptide into the secretory pathway. In order to examine amino acid sequence conservation between the EGAD1 precursor and those of other defensin/γ-thionin proteins, a CLUSTAL alignment was carried out. The EGAD1 precursor sequence was aligned, using the CLUSTAL W program (THOMPSON *et al.*, Nucl. Acids Res., 22, 4673-4680, 1994) with those of its 3 closest published relatives, namely the products of the genes *PPT* from *Petunia inflata* (KARUNANANDAA *et al.*, Plant Mol. Biol., 26, 459-464, 1994), *J1-1* from *Capsicum annuum* (MEYER *et al.*, Plant Physiol., 112, 615-622, 1996) and *P322* from potato (STIEKEMA *et al.*, Plant Mol. Biol., 11, 255-269, 1988).

Figure 3 shows the alignment of the deduced EGAD1 protein sequence with those encoded by the genes *PPT* from *Petunia inflata* (SEQ ID NO: 3), j1-1 from *Capsicum annuum* (SEQ ID NO: 5) and P322 from potato (SEQ ID NO: 4). Asterisks indicate residues which are absolutely conserved in all 4 polypeptide sequences, whilst single and double dots respectively denote conservations within "weaker" or "stronger" groups as defined within the CLUSTAL W program. The position of the putative translation termination codon is indicated by an asterisk.

The closest identified relative of EGAD1 is PPT, the two respective sequences sharing 63.6% identical residues. It is interesting to note that the positions of the 8 cysteine residues are absolutely conserved between all 4 polypeptides, strongly suggesting that they all share the same secondary structure. Defensins are thought to play a

role in pathogen defence and in some cases have been shown to exert an antifungal action, probably brought about by electrostatic interaction with hyphal cell membranes (THEVISSEN *et al.*, J. Biol. Chem. 271, 15018-15025, 1996).

EXAMPLE 3: ESTIMATION OF THE NUMBER OF *EGAD1*-LIKE GENE LOCI PRESENT IN THE OIL PALM GENOME

Oil palm genomic DNA was extracted from leaves of 3 different adult oil palms, X, Y and E, of different genotype in each case (NB palms X and E were used to obtain the tissue cultures X1 and LAB146 respectively). DNAs were digested with 3 different restriction enzymes, Southern blotted and hybridised to the *EGAD1* cDNA.

Figure 4 shows a Southern hybridisation of a ^{32}P-radiolabelled *EGAD1* cDNA probe to oil palm genomic DNA digests from the three genotypically different oil palms, X, Y and E,. The restriction enzymes used and size marker migration positions are indicated.

These results suggest that the *EGAD1* gene occurs as a single copy per haploid genome. All 3 restriction enzymes used (none of which cuts the *EGAD1* cDNA) were found to produce only one hybridising fragment, except in the case of *Eco*RI when used with DNA of the Y genotype; in this case, a doublet was observed. One possible explanation for the latter result is that it might be caused by a difference in restriction pattern between allelic regions in or bordering the *EGAD1* gene locus in the Y genome.

EXAMPLE 4: INVESTIGATION OF THE *MANTLED*-DEPENDENT EXPRESSION OF *EGAD1* IN OTHER OIL PALM GENOTYPES AND CULTURES

Since the northern hybridisation data shown in Figure 1 revealed a strong differential accumulation of *EGAD1* transcripts at the callus stage of tissue culture regeneration, we investigated this phenomenon further. The N and A cultures used for the differential display and preliminary expression analysis were not of identical genotype; thus it was important to check that the difference in transcript levels seen was not simply attributable to differences in genetic background. We therefore investigated *EGAD1* transcript accumulation in genotypically identical

cultures differing only in their *mantled* status, namely the cultures X1 (normal), X2 (intermediate) and X3 (abnormal). These cultures were initiated from respectively a normal seed-derived palm, a normal regenerant palm previously cloned from the seed-derived palm and an abnormal regenerant palm again previously cloned from the same original palm. The X1, X2 and X3 cultures were all initiated at the same time, so as to eliminate any possible effects due to culture age. In terms of abnormality, the X2 culture can be considered as representing an intermediate situation in that it has been initiated from a normal palm, but consists of cells which have undergone an extra round of tissue culture compared with the X1 culture. In the same experiment, in order to check the reproducibility of *EGAD1* expression between different cultures obtained from a given starting material, we analysed callus lines N' (LAB148) and A' (LAB149), which were respectively initiated from the same palms as the LAB146 and LAB147 cultures. Figure 6 shows the northern hybridisation profiles obtained with each callus line using the *EGAD1* probe. The hybridisation profiles of the N' and A' samples are consistent with those shown in Figure 1 for the N and A lines, thus confirming the reproducibility of these results between cloning experiments and providing further evidence that the differential *EGAD1* gene expression patterns observed may be associated with the *mantled* abnormality. This hypothesis is further strengthened by the hybridisation profile observed for the X1, X2 and X3 cultures, which differ in their *mantled* status within a common genotypic background. As expected, the X3 (abnormal) culture produces a dramatically stronger signal than the X1 (normal) culture. It is moreover interesting to note that the X2 (intermediate) culture produces a signal greater than that of X1, but weaker than that of X3, presumably reflecting the extra cycle of tissue culture "history" which it carries compared with the X1 line.

The above data confirm that *EGAD1* gene expression is affected by epigenetic factors which prevail in tissue culture, resulting in increased transcript accumulation in abnormal clonal lines at the callus stage. The

reproducibility of *EGAD1* gene expression profiles between cultures established at different times from the same palm has been demonstrated. Most importantly, studies performed on the genotypically homogeneous X1/X2/X3 cultures shows that *EGAD1* gene expression provides a sensitive marker for the presence of the *mantled* abnormality at the callus stage.

CLAIMS

1) An isolated polypeptide of the defensin family, hereinafter referred as EGAD1, having at least 70% sequence identity with the polypeptide of SEQ ID NO: 2.

2) A fragment of an EGAD1 polypeptide of claim 1, comprising at least 5 contiguous amino-acids of said polypeptide.

3) An isolated polynucleotide selected among :

a) a polynucleotide encoding an EGAD1 polypeptide of claim 1;

b) a fragment of a polynucleotide a) comprising at least 10 contiguous nucleotides ;

c) a polynucleotide complementary of a polynucleotide a) or b) above;

d) a polynucleotide hybridising selectively under stringent conditions with a polynucleotide a) b) or c) above.

4) A polynucleotide of claim 3 selected among :

a) the cDNA of SEQ ID NO:1;

b) a polynucleotide comprising at least 10 contiguous nucleotides of SEQ ID NO: 1;

c) a polynucleotide complementary of a polynucleotide a) or b) above;

d) a polynucleotide hybridising selectively under stringent conditions with a polynucleotide a) b) or c) above.

5) An antibody directed against a polypeptide of any of claims 1 or 2.

6) The use of a polypeptide of any of claims 1 or 2 or of a polynucleotide of any of claims 3 or 4, or of an antibody of claim 5, for detecting the *mantled* phenotype in an oil palm tissue culture.

7) A method for detecting the *mantled* abnormality in an oil palm tissue culture, wherein said method comprises detecting the overexpression of the *EGAD1* gene, encoding an EGAD1 polypeptide of claim 1, in said culture.

8) A method of claim 7, wherein the overexpression of the *EGAD1* gene is detected in an oil palm tissue culture at the callus stage.

9) A method of any of claims 7 or 8, wherein the overexpression of the *EGAD1* gene is detected by evaluating the amount of *EGAD1* transcripts.

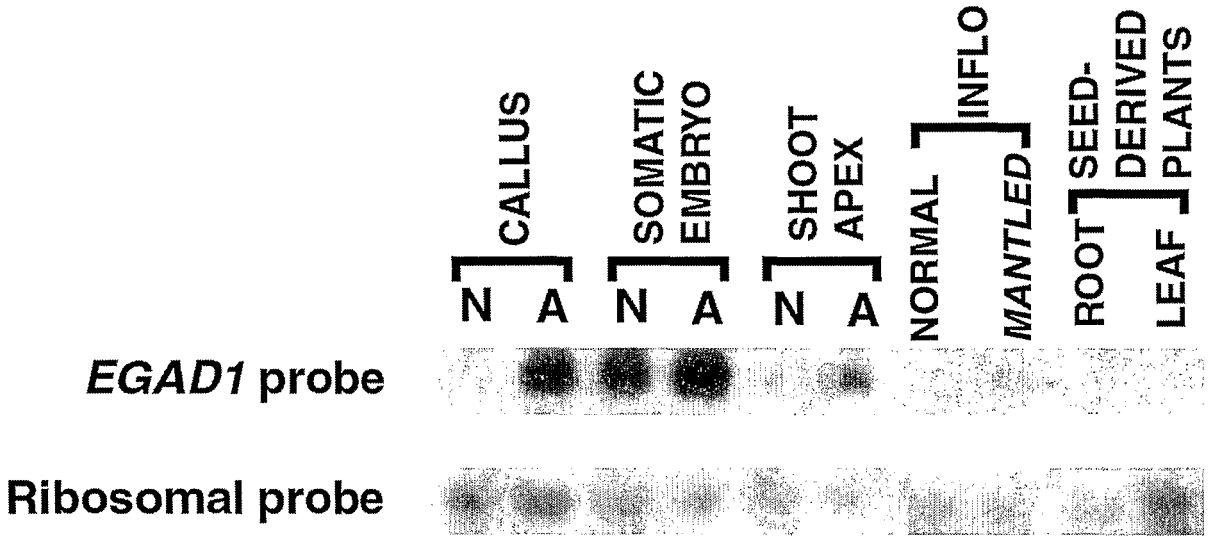
CALLUS
SOMATIC EMBRYO
SHOOT APEX
NORMAL
INFLO
MANTLED
ROOT
SEED-DERIVED
LEAF
PLANTS
N A N A N A
EGAD1 probe
Ribosomal probe

FIG. 1

```
1/1                                                    31/11
TTT TTG CTC TGA GCG TGT GTT AGC TAG TGT GTT GCG ATG GAG CAC TCT CGG CGA ATG CTT
                                                       M   E   H   S   R   R   M   L

61/21                                                  91/31
CCA GCT ATC CTC CTG CTC TTG TTC CTT CTC ATG CCC TCT GAG ATG GGG ACG AAG GTG GCG
P   A   I   L   L   L   L   F   L   L   M   P   S   E   M   G   T   K   V   A

121/41                                                 151/51
GAG GCA AGG ACC TGC GAG TCT CAA AGC CAC AAG TTC CAG GGC ACG TGC TTG AGA GAA AGC
E   A   R   T   C   E   S   Q   S   H   K   F   Q   G   T   C   L   R   E   S

181/61                                                 211/71
AAC TGT GCA AAC GTG TGC CAG ACC GAG GGC TTC CAA GGA GGA GTT TGC CGG GGA GTC CGG
N   C   A   N   V   C   Q   T   E   G   F   Q   G   G   V   C   R   G   V   R

241/81                                                 271/91
CGC CGA TGC TTT TGC ACC AGG CTT TGC TAA TGA TCT ACA CTT CGC ACA TAG GAT GGT GAG
R   R   C   F   C   T   R   L   C   *

301/101                                                331/111
GGT TAT GTG GTT GCC CGT AGC TTT CAT GCC TCC CAG AAT AAA ATA AGC CTA GTT TTA GGA

361/121                                                391/131
TGT GTT GCT CTG TTC TAT CCT TTG TGG TAG TCA AGT CCT TAT GGC GTG TTA ACT GTG TGT

421/141                                                451/151
TTG AAC AAG TCT AAT TGT GCT TAT GAA TGA TGC TGT TTC GCT TAG TTG AGC TAG CTT GTA

481/161                                                511/171
TTT TGC TGT TGT CTA ATG TGA GTA TTT GAG TAT ATT AAA AAA AAA AAA AAA AAA AA
```

FIG. 2

```
‡
PPT             MGRSIRLFATFFLIAMLFLSTEMGPMTSAEARTCESQSHRFHGTCVRESNCASVCQTEGF
P322            ----MRFFATFFLLAMLVVATKMGPMRIAEARHCESLSHRFKGPCTRDSNCASVCETERF
j1-1            MAGFSKVIATIFLMMMLVFATGM----VAEARTCESQSHRFKGLCFSKSNCGSVCHTEGF
EGAD1           MEHSRRMLPAILLLLFLLMPSEMG-TKVAEARTCESQSHKFQGTCLRESNCANVCQTEGF
                     :.:.:::*: :*...: *     **** *** **:*:* *  .***..**.** *

PPT             IGGNCRAFRRRCFCTRNC
P322            SGGNCHGFRRRCFCTKPC
j1-1            NGGHCRGFRRRCFCTRHC
EGAD1           QGGVCRGVRRRCFCTRLC
                 ** *:..*******: *
```

FIG. 3

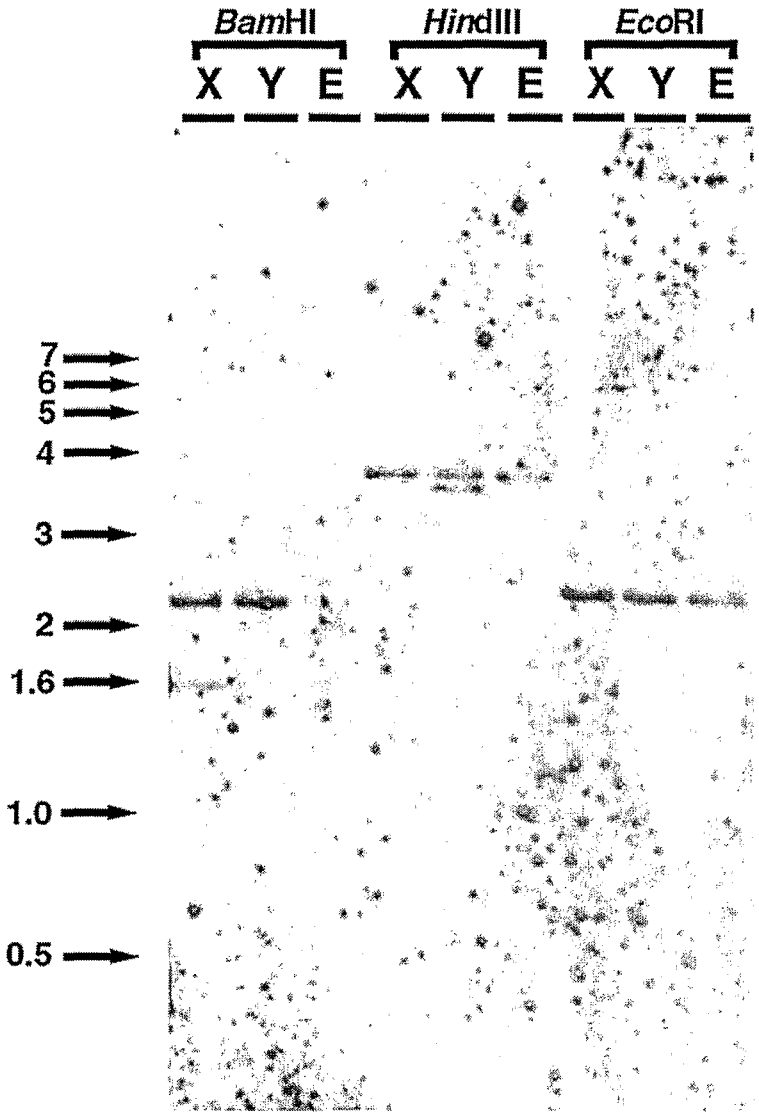

FIG. 4

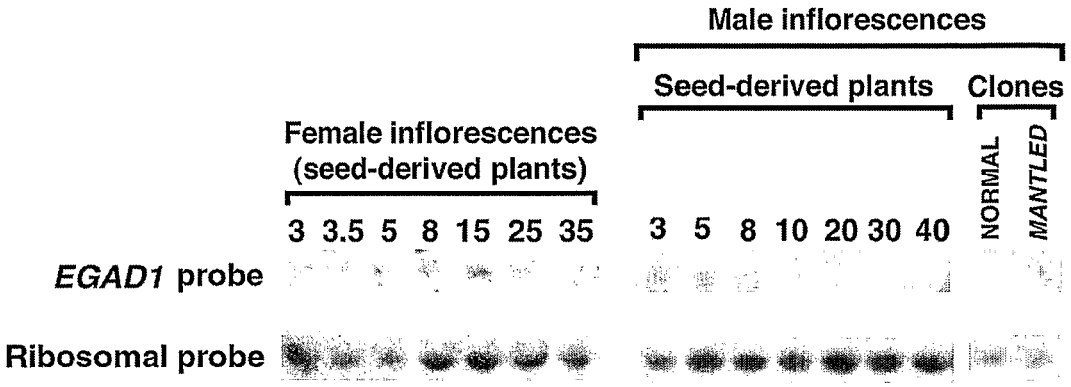
Male inflorescences
Seed-derived plants
Clones
Female inflorescences
(seed-derived plants)
NORMAL
MANTLED
3 3.5 5 8 15 25 35
3 5 8 10 20 30 40
EGAD1 probe
Ribosomal probe

FIG. 5

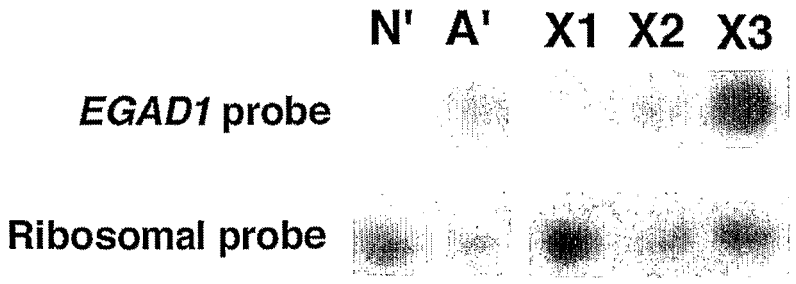
N' A' X1 X2 X3
EGAD1 probe
Ribosomal probe

FIG. 6

SEQUENCE LISTING

<110> CENTRE DE COOPERATION INTERNATIONALE EN RECHERCHE AGRONOMIQUE POUR LE DEVELOPPEMENT (CIRAD)

MALAYSIAN PALM OIL BOARD

CHEAH, Suan Choo

TREGEAR, James W

<120> NEW PROTEIN OF THE DEFENSIN FAMILY

<130> MJPbv1367/5

<150> PI 2001 5098

<151> 2001-11-06

<160> 5

<170> PatentIn version 3.1

<210> 1

<211> 536

<212> DNA

<213> Elaeis guineensis

<400> 1
tttttgctct gagcgtgtgt tagctagtgt gttgcgatgg agcactctcg gcgaatgctt 60

ccagctatcc tcctgctctt gttccttctc atgccctctg agatggggac gaaggtggcg 120

gaggcaagga cctgcgagtc tcaaagccac aagttccagg gcacgtgctt gagagaaagc 180

aactgtgcaa acgtgtgcca gaccgagggc ttccaaggag gagtttgccg gggagtccgg 240

cgccgatgct tttgcaccag gctttgctaa tgatctacac ttcgcacata ggatggtgag 300

ggttatgtgg ttgcccgtag ctttcatgcc tcccagaata aaataagcct agttttagga 360

tgtgttgctc tgttctatcc tttgtggtag tcaagtcctt atggcgtgtt aactgtgtgt 420

ttgaacaagt ctaattgtgc ttatgaatga tgctgtttcg cttagttgag ctagcttgta 480

ttttgctgtt gtctaatgtg agtatttgag tatattaaaa aaaaaaaaaa aaaaaa 536

<210> 2

<211> 77

<212> PRT

<213> Elaeis guineensis

<400> 2

Met Glu His Ser Arg Arg Met Leu Pro Ala Ile Leu Leu Leu Leu Phe
1 5 10 15

Leu Leu Met Pro Ser Glu Met Gly Thr Lys Val Ala Glu Ala Arg Thr
20 25 30

Cys Glu Ser Gln Ser His Lys Phe Gln Gly Thr Cys Leu Arg Glu Ser
35 40 45

Asn Cys Ala Asn Val Cys Gln Thr Glu Gly Phe Gln Gly Gly Val Cys
50 55 60

Arg Gly Val Arg Arg Arg Cys Phe Cys Thr Arg Leu Cys
65 70 75

<210> 3

<211> 78

<212> PRT

<213> Petunia inflata

<400> 3

Met Gly Arg Ser Ile Arg Leu Phe Ala Thr Phe Phe Leu Ile Ala Met
1 5 10 15

Leu Phe Leu Ser Thr Glu Met Gly Pro Met Thr Ser Ala Glu Ala Arg
20 25 30

Thr Cys Glu Ser Gln Ser His Arg Phe His Gly Thr Cys Val Arg Glu
35 40 45

Ser Asn Cys Ala Ser Val Cys Gln Thr Glu Gly Phe Ile Gly Gly Asn
50 55 60

Cys Arg Ala Phe Arg Arg Arg Cys Phe Cys Thr Arg Asn Cys
65 70 75

<210> 4

<211> 74

<212> PRT

<213> Solanum tuberosum

<400> 4

Met Arg Phe Phe Ala Thr Phe Phe Leu Leu Ala Met Leu Val Val Ala
1 5 10 15

Thr Lys Met Gly Pro Met Arg Ile Ala Glu Ala Arg His Cys Glu Ser
20 25 30

Leu Ser His Arg Phe Lys Gly Pro Cys Thr Arg Asp Ser Asn Cys Ala
35 40 45

Ser Val Cys Glu Thr Glu Arg Phe Ser Gly Gly Asn Cys His Gly Phe
50 55 60

Arg Arg Arg Cys Phe Cys Thr Lys Pro Cys
65 70

<210> 5

<211> 74

<212> PRT

<213> Capsicum annuum

<400> 5

```
Met Ala Gly Phe Ser Lys Val Ile Ala Thr Ile Phe Leu Met Met Met
1               5                  10                  15

Leu Val Phe Ala Thr Gly Met Val Ala Glu Ala Arg Thr Cys Glu Ser
            20                  25                  30

Gln Ser His Arg Phe Lys Gly Leu Cys Phe Ser Lys Ser Asn Cys Gly
        35                  40                  45

Ser Val Cys His Thr Glu Gly Phe Asn Gly Gly His Cys Arg Gly Phe
    50                  55                  60

Arg Arg Arg Cys Phe Cys Thr Arg His Cys
65                  70
```