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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
26 June 2014 (26.06.2014)

WIPO | PCT

(10) International Publication Number
WO 2014/100467 A1

(51) International Patent Classification:
C12N 15/82 (2006.01)

(21) International Application Number:
PCT/US2013/076672

(22) International Filing Date:
19 December 2013 (19.12.2013)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/739,499 19 December 2012 (19.12.2012) US
13/830,012 14 March 2013 (14.03.2013) US

(71) Applicants: **BOSTON MEDICAL CENTER CORPORATION** [US/US]; One Boston Medical Center Place, Boston, MA 02118 (US). **UNIVERSITY OF NORTH TEXAS** [US/US]; 3940 North Elm Street Suite A101, Denton, TX 76207-7102 (US).

(72) Inventors: **PURI, Vishwajeet**; 14 Washington Lane, Hopkinton, MA 01748 (US). **CHAPMAN, Kent**; 3613 Montecito Road, Denton, TX 76205 (US). **JAMES, Christopher**; 3232 Buckthorn Lane, Argyle, TX 76226 (US). **CAI, Yingqi**; 1408 Teasley Lane, Apt 2926, Denton, TX 76205 (US).

(74) Agents: **LLOYD, Jeff** et al.; Saliwanchik, Lloyd & Eisenschenk, PO Box 142950, Gainesville, FL 32614-2950 (US).

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

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

Published:

— *with international search report (Art. 21(3))*

— *with sequence listing part of description (Rule 5.2(a))*

(54) Title: METHODS FOR ELEVATING FAT/OIL CONTENT IN PLANTS

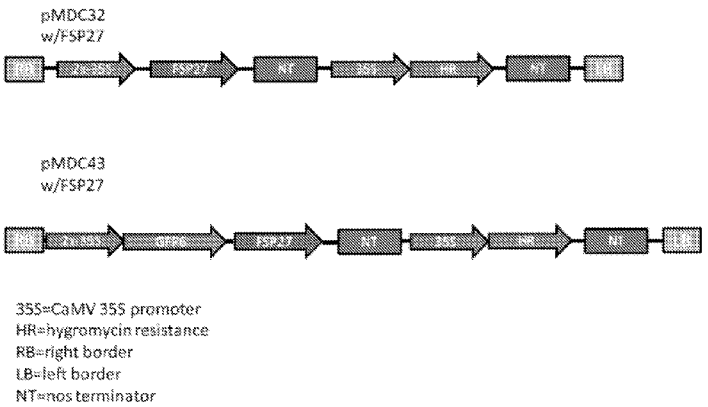


FIG. 1

(57) Abstract: In some embodiments, the present invention provides a method of elevating lipid content in vegetative (non-seed) plant or algal cells, plant tissues, or whole plants by genetically modifying the plant or algae to express a protein or polypeptide associated with lipid metabolism (such as fat-specific protein 27) of animal origin or plant origin. Also provided are genetically-modified plant or algal cells, plant tissues, or whole plants with elevated cellular lipid content, expressing a protein or polypeptide associated with lipid metabolism (such as fat-specific protein 27) of animal (e.g. human) origin or plant origin.

WO 2014/100467 A1

DESCRIPTION

METHODS FOR ELEVATING FAT/OIL CONTENT IN PLANTS

CROSS-REFERENCE TO RELATED APPLICATION

The present application claims the benefit of U.S. Provisional Application Serial No. 61/739,499, filed December 19, 2012, and U.S. Non-provisional Application Serial No. 13/830,012, filed March 14, 2013, both of which are hereby incorporated by reference in their entirety, including any figures, tables, or drawings.

BACKGROUND OF THE INVENTION

Plants are a primary source of human and/or animal food, excellent feedstock for fuels, and useful for production of desirable chemicals. Plants synthesize and store lipids, primarily, in cytosolic lipid droplets. In plants, seeds are the primary site of oil synthesis and storage; vegetable oils (such as triacylglycerol) are used as a form of energy during seed germination. Vegetable oils can be synthesized in non-seed (such as leaf) tissues; however, their abundance is low and the stored lipids are presumed to be metabolized rapidly, perhaps for the recycling of fatty acids for energy or the synthesis of membrane lipids.

Plants that can accumulate oils in non-seed tissues are commercially attractive. The biomass of non-seed parts (such as leaves, stems) of plants is generally far greater than the amount accounted for by seeds. Thus, the transformation of non-seed tissues into oil-producing machinery can significantly increase the energy-production capacity. Currently, the regulation and transient accumulation of stored oils in non-seed tissues are not well understood, and the production of oils in non-seed plant tissues for industrial applications remains challenging. Cellular lipid droplets are dynamic organelles that regulate triglyceride storage in mammalian cells. Lipid droplets are composed of a core of neutral lipids surrounded by a phospholipid monolayer and associated proteins. Various proteins associated with lipid metabolism, including fat specific protein 27 (FSP27), perilipins, (Bernardinelli-Seip congenital lipodystrophy type 2 protein), FIT1 (fat storage-inducing transmembrane protein 1), and FIT2 (fat storage-inducing transmembrane protein 2) have been well characterized for their ability to regulate fat metabolism in mammalian species.

BRIEF SUMMARY OF THE INVENTION

In some embodiments, the present invention provides a method of elevating oil content in algae, plants, or plant parts by genetically modifying the plant to express a protein or polypeptide associated with lipid metabolism (such as fat-specific protein 27) of animal or plant origin. In one specific embodiment, the present invention provides a method of elevating oil content in vegetative (non-seed) plant tissues or algae.

In some embodiments, the present invention also provides genetically-modified algal cells, plant cells, tissues, or whole plants with elevated cellular oil content, wherein the algal cell, plant cell, tissue, or whole plant expresses a protein or polypeptide associated with lipid metabolism (such as fat-specific protein 27) of exogenous origin, for example, of exogenous animal origin or exogenous plant origin. In certain embodiments, the proteins or polypeptides associated with lipid metabolism useful according to the present invention are of mammalian origin. In some embodiments, the present invention provides a method for obtaining a plant cell or algal cell with elevated lipid content, wherein the method comprises:

genetically modifying the plant cell or algal cell to express an exogenous protein or polypeptide associated with lipid metabolism, thereby obtaining a genetically-modified plant cell or algal cell with elevated lipid content;

wherein the protein or polypeptide associated with lipid metabolism induces adipogenesis, enhances the accumulation of cellular lipid droplets, and/or reduces lipase activity; and

wherein the expression of the protein or polypeptide associated with lipid metabolism increases lipid content of the genetically-modified plant cell or algal cell as compared to a wild-type (native) plant cell or algal cell of the same type.

In some embodiments, the present invention provides a method for obtaining a plant cell or algal cell with elevated lipid content, wherein the method comprises:

transforming the plant cell or algal cell with a vector comprising a nucleic acid sequence encoding an exogenous protein or polypeptide associated with lipid metabolism, wherein the nucleic acid is operably linked to a promoter and/or a regulatory sequence;

wherein the protein or polypeptide associated with lipid metabolism induces adipogenesis, enhances the accumulation of cellular lipid droplets, and/or reduces lipase activity;

wherein the transformed plant cell or algal cell expresses the protein or polypeptide associated with lipid metabolism; and

wherein the expression of the protein or polypeptide associated with lipid metabolism increases lipid content of the transformed plant cell or algal cell as compared to a wild-type (native) plant cell or algal cell of the same type.

In certain embodiments, the genetically-modified plant cell is contained in a plant tissue, plant part, or whole plant.

In some embodiments, the genetically-modified plant cell or algal cell comprises, in its genome or in its plastome, a nucleic acid molecule encoding a protein or polypeptide associated with lipid metabolism.

In some embodiments, the protein or polypeptide associated with lipid metabolism is not of plant origin. In certain embodiments, the protein or polypeptide associated with lipid metabolism is of animal origin, such as of insect, vertebrate, fish, bird, amphibian, or mammalian (*e.g.*, mouse, human) origin. In some embodiments, the protein or polypeptide associated with lipid metabolism is of plant origin.

In some embodiments, a T-DNA binary vector system is used for plant transformation. In one embodiment, plant transformation is performed using the floral dip method.

In certain embodiments, to elevate cellular lipid content and/or to induce lipid droplet production, the plant cell or the algal cell can be genetically engineered to expresses one or more proteins or polypeptides associated with lipid metabolism including, but not limited to, fat specific protein 27 (FSP27); perilipins including PLIN1 (perilipin 1) and PLIN2 (also called autosomal dominant retinitis pigmentosa (ADRP)); SEIPIN (Bernardinelli-Seip congenital lipodystrophy type 2 protein); FIT1 (fat storage-inducing transmembrane protein 1), and FIT2 (fat storage-inducing transmembrane protein 2); acyl-CoA:diacylglycerol acyltransferase 1 (DGAT-1) and phospholipid:diacylglycerol acyltransferase 1 (PDAT-1); cell death activator (Cidea); leafy cotyledon 2 (LEC2); and WRINKLED1 (WRI1).

In certain embodiments, to elevate cellular lipid content and/or to induce lipid droplet production, the plant cell or the algal cell can be genetically engineered to expresses one or more proteins or polypeptides associated with lipid metabolism including, but not limited to FSP27, PLIN1, PLIN2, SEIPIN, FIT1, FIT2, and LEC2.

In certain specific embodiments, the transgenic plants or algae express a combination of proteins or polypeptides associated with lipid metabolism, wherein the protein or polypeptide associated with lipid metabolism is selected from: DGAT-1 and FSP27; DGAT-1, *cgi58* (mutation), and FSP27; DGAT-1, PDAT-1, and FSP27; DGAT-1, PDAT-1, *cgi58*

(mutation), FSP27; FSP27, PLIN2, and *cgi58* (mutation); DGAT-1, FSP27, PLIN2, and *cgi58* (mutation); and DGAT-1, PDAT-1, FSP27, PLIN2, and *cgi58* (mutation). In a further embodiment of the invention, the transgenic plants or algae express any combination of proteins or polypeptides associated with lipid metabolism selected from: DGAT-1, FSP27, cgi58 (mutation), PDAT-1, PLIN2, FIT1, FIT2, SEIPIN, LEC2, and WRI1. In certain other embodiments, various proteins or polypeptides associated with lipid metabolism expressed in a transgenic plant or algae are of different origin. For example, in an embodiment of the invention, a plant or algal cell expresses human FSP27 and SEIPIN.

In another embodiment, the present invention provides a method for obtaining an algae or bacterial cell with elevated lipid content, wherein the method comprises:

transforming an algae or bacterial cell with a vector comprising a nucleic acid sequence encoding an exogenous protein or polypeptide associated with lipid metabolism, wherein the nucleic acid is operably linked to a promoter and/or a regulatory sequence;

wherein the protein or polypeptide associated with lipid metabolism induces adipogenesis, enhances the accumulation of cellular lipid droplets, and/or reduces lipase activity;

wherein the transformed algae or bacterial cell expresses the protein or polypeptide associated with lipid metabolism; and

wherein the expression of the protein or polypeptide associated with lipid metabolism increases lipid content of the transformed algae or bacterial cell as compared to a wild-type (native) algae or bacterial cell of the same type.

In certain embodiments, the algal cell can be genetically engineered to expresses any combinations of proteins associated with lipid metabolism and peptides including, but not limited to, FSP27; perilipins including PLIN1 and PLIN2; SEIPIN; FIT1 and FIT2 ; DGAT-1; PDAT-1; Cidea; LEC2; and WRI1.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1A is a diagram that illustrates embodiments of the transfer DNA (T-DNA) region of the binary vector for transformation of *A. thaliana* with the mouse fat specific protein 27 (FSP27) cDNA. The FSP27 open reading frame was inserted downstream from the 2X 35S promoter, either in-frame with green fluorescent protein (GFP) (pMDC43) or without (pMDC32). Binary vectors are known in the art, as described in Curtis and Grossniklaus (Plant Physiology, October 2003, Vol. 133, pp. 462–469), which is herein

incorporated by reference in its entirety. Plasmid vectors were transformed into *Agrobacterium tumefaciens* LBA4404 and clones were selected and verified by PCR. Arabidopsis plants were transformed by the floral dip method of Bent and Clough (Plant J. 1998 Dec;16(6):735-43.). Both wild-type plants (*A. thaliana*, ecotype Columbia), and plants with a T-DNA insertional mutation in the At4g24160 locus were used for transformations. The T-DNA knockout is in an exon of the Arabidopsis homolog of the human CGI-58 gene, and in Arabidopsis plants with this mutation there is an increase in cytosolic lipid droplets in leaves (James et al., Proc. Natl. Acad. Sci. U S A. 2010 Oct 12;107(41):17833-8).

Figure 1B are Confocal Laser Scanning Microscopy images of leaves of approximately 30-d-old Arabidopsis seedlings stained with the neutral lipid-specific stain, Nile blue. Red autofluorescence is from chlorophyll and shows the location of chloroplasts distributed around the perimeter of leaf mesophyll cells. Lipid droplets (blue) are distributed throughout the cytosol of the cells and are more abundant in transgenic seedlings expressing mouse FSP27 than in non-transformed cells (WT). Bar is 20 microns.

Figure 2 shows representative Confocal Laser Scanning Microscopy images of leaves of approximately 30-day-old *A. thaliana* seedlings stained with Nile blue - a neutral lipid-specific stain. Red autofluorescence emitted from chlorophylls shows the location of chloroplasts distributed around the perimeter of leaf mesophyll cells. Lipid droplets (blue) are distributed throughout the cytosol of the cells and are more abundant in transgenic seedlings expressing mouse FSP27 than in non-transformed cells (WT). Bar is 20 microns.

Figure 3 shows representative Confocal Laser Scanning Microscopy images of leaves of approximately 30-day-old *A. thaliana* seedlings stained with BODIPY 493/503 - a neutral lipid-specific stain. Red autofluorescence emitted from chlorophylls shows the location of chloroplasts distributed around the perimeter of leaf mesophyll cells. Lipid droplets (yellow-green with BODIPY staining) are distributed throughout the cytosol of the cells and are more abundant in transgenic seedlings expressing mouse FSP27 than in non-transformed cells (*cgi58*). Bar is 20 microns.

Figure 4 shows representative Confocal Laser Scanning Microscopy images of leaves of approximately 30-day-old *A. thaliana* seedlings stained with Nile blue - a neutral lipid-specific stain. Red autofluorescence emitted from chlorophylls shows the location of chloroplasts distributed around the perimeter of leaf mesophyll cells. GFP fluorescence (green) marks the location of the mouse FSP27- GFP fusion protein. Lipid droplets (blue) are distributed throughout the cytosol of the cells and are more abundant in the *cgi58* mutant

background than in the wild-type background. More lipid droplets are formed in leaves of transformed plants than in untransformed leaves (see also Fig. 2). Scale bars represent 20 microns.

Figure 5 shows the content of total fatty acids extracted from 15-day-old *A. thaliana* seedlings sown on solidified nutrient medium. The total fatty acid content is shown on a fresh weight basis. Transgenic plants (mouse FSP27-GFP in the *cgi58* mutant background) in the T1 generation are selected using hygromycin medium. Despite the inclusion of heterozygotes in the analysis, the FSP27-transformed plants exhibit a measureable increase in total lipid content. Also, it is postulated that the transfer of FSP27 stabilizes the variable *cgi58* phenotype (reduced standard deviation in the FSP27 expressing plants). Values are the means and standard deviation of three replicates.

Figure 6 shows the content of total fatty acids extracted from 15-day-old *A. thaliana* seedlings sown on solidified nutrient medium. The total fatty acid content is shown on a dry weight basis. Transgenic plants (expressing mouse FSP27-GFP or mouse autosomal dominant retinitis pigmentosa (ADRP)) in the T1 generation are selected on hygromycin medium. All FSP27-GFP or ADSP transgenic plants have a higher average lipid content in the T1 generation than that of the non-transformed plants, and one line (*cgi58*-43fsp27line1) has a statistically higher lipid content ($P < 0.05$) than that of non-transformed plants. Values are the means and standard deviations of five replicates.

Figure 7A-C show confocal fluorescence micrographs of leaves in Arabidopsis plants expressing ADRP (lower left; A-C) or FSP27 (lower right; A-C) in the *cgi58* knockout background. Red autofluorescence is marking chloroplasts; green fluorescence is from the neutral-lipid-specific stain– BODIPY 493/503, showing the accumulation of lipid droplets in leaves. The upper left is wild-type; upper left is the *cgi58* knockout background alone.

Figure 8 shows that amino acids 120-220 of FSP27 are associated with lipid accumulation. Amino acids 120-220 of FSP27 and the full length FSP27 are expressed in human adipocytes using lentivirus. X-axis shows total triglycerides in adipocytes. Note that the human adipocytes already have huge amount of triglycerides, and the expression of FSP27 (full length) and FSP27 (120-220) significant increase triglyceride contents in adipocytes by almost 40%. *, $p<0.05$, t-test.

Figure 9 shows sequence similarity between mouse and zebra fish FSP27 protein. NP_848460.1: CIDE-3 Mus musculus (mouse); NP_001038512.1: CIDE-3 Danio rerio (zebra fish).

Figure 10 shows motif locations of various SEIPIN homologs from *H. sapiens*, *S. cereviciae*, and *A. thaliana*.

Figure 11 shows sequence alingnment of various SEIPIN homologs from *H. sapiens*, *S. cereviciae*, and *A. thaliana*.

Figure 12 shows developmental and tissue-specific expression profiles of Arabidopsis SEIPIN genes identified by semi-quantitative reverse transcriptase (RT)-PCR analysis of Arabidopsis SEIPIN isoforms. Constitutively-expressed elongation factor (EF)1-alpha is included for comparison. SEIPIN2 and SEIPIN3 appear to be more constitutively expressed and may function in a partially redundant manner. Whereas, SEIPIN1 seems only to be expressed in seeds and seedlings.

Figure 13 shows lipid droplet staining in wild type and genetically modified yeast. Green fluorescence is from the neutral-lipid-specific stain–BODIPY 493/503, showing the accumulation of lipid droplets. The top left panel shows lipid droplets in wild type yeast, top middle panel shows lipid droplets in *ylr404wΔ*, which is a yeast having a deletion of yeast SEIPIN protein. The top right panel shows lipid droplets in *ylr404wΔ*, expressing yeast SEIPIN. The bottom left panel shows lipid droplets in *ylr404wΔ*, expressing yeast *A. thaliana* SEIPIN1, the bottom middle panel shows lipid droplets in *ylr404wΔ*, expressing *A. thaliana* SEIPIN2, and the bottom right panel shows lipid droplets in *ylr404wΔ*, expressing *A. thaliana* SEIPIN3. Expression of *A. thaliana* SEIPIN1, 2, or 3 restores lipid droplet accumulation in *ylr404wΔ*.

Figure 14 shows quantification of lipid droplets in terms of the number of lipid droplets per cell in wild type and genetically modified yeast. Number of lipid droplets is significantly reduced in *ylr404wΔ* compared to wild type yeast. Expression of *A. thaliana* SEIPIN1, 2, or 3 restores lipid droplet accumulation in *ylr404wΔ* to certain extent with *A. thaliana* SEIPIN3 having the maximum effect in terms of the number of lipid droplets per yeast cell.

Figure 15 shows lipid droplet staining in wild type and genetically modified yeast. Green fluorescence is from the neutral-lipid-specific stain–BODIPY 493/503, showing the accumulation of lipid droplets. The size of lipid droplets is significant increased in *ylr404wΔ* compared to wild type yeast. Expression of *A. thaliana* SEIPIN1, 2, or 3 did not restore the number of lipid droplets in *ylr404wΔ* to those observed in wild type yeast. Expression of *A. thaliana* SEIPINs also increased the size of lipid droplets in *ylr404wΔ* compared to wild type

yeast, with *A. thaliana* SEIPIN1 producing the biggest lipid droplets amongst the mutants tested.

Figure 16 shows quantification of lipid droplets in terms of the size of lipid droplets in wild type and genetically modified yeast. The size of lipid droplets is significant increased in *ylr404wΔ* compared to wild type yeast. Expression of *A. thaliana* SEIPIN1, 2, or 3 did not restore the size of lipid droplets in *ylr404wΔ* to those observed in wild type yeast. Expression of *A. thaliana* SEIPINs also increased the size of lipid droplets in *ylr404wΔ* compared to wild type yeast with *A. thaliana* SEIPIN1 producing the biggest lipid droplets amongst the mutants tested.

Figure 17 further illustrates changes in the size of the lipid droplets in wild type and genetically modified yeast.

Figure 18 shows localization of *A. thaliana* SEIPIN1 to lipid droplets when expressed in yeast. The top left panel indicates Nile Red staining of lipid droplets and the top right column shows green fluorescence indicating localization of *A. thaliana* SEIPIN1-GFP. The bottom left panel shows endoplasmic reticulum with blue fluorescence coming from cyano fluorescence protein (CFP) fused to HDEL, which is a C-terminal tetrapeptide found in yeast and plants allowing the sorting of the proteins in the lumen of the endoplasmic reticulum. The bottom right panel shows the merged figure of the other three panels indicating that *A. thaliana* SEIPIN1-GFP colocalises with lipid droplets in yeast.

Figure 19 shows localization of *A. thaliana* SEIPIN2 to lipid droplets when expressed in yeast. The top left panel indicates Nile Red staining of lipid droplets and the top right column shows green fluorescence indicating localization of *A. thaliana* SEIPIN2-GFP. The bottom left panel shows endoplasmic reticulum with blue fluorescence coming from CFP fused to HDEL. The bottom right panel shows the merged figure of the other three panels indicating that *A. thaliana* SEIPIN2-GFP colocalises with lipid droplets yeast.

Figure 20 shows localization of *A. thaliana* SEIPIN3 to lipid droplets when expressed in yeast. The top left panel indicates Nile Red staining of lipid droplets and the top right column shows green fluorescence indicating localization of *A. thaliana* SEIPIN3-GFP. The bottom left panel shows endoplasmic reticulum with blue fluorescence coming from CFP fused to HDEL. The bottom right panel shows the merged figure of the other three panels indicating that *A. thaliana* SEIPIN3-GFP colocalises with lipid droplets yeast.

Figure 21 shows quantification of lipid droplets in terms of the amount of triacylglyceride (TAG) amount in lipid droplets in the wild type and genetically modified

yeast. The amount of TAG in lipid droplets is significant decreased in *ylr404wΔ* compared to wild type yeast. Expression of yeast SEIPIN and *A. thaliana* SEIPIN1, 2, or 3 restored the amount of TAG in the lipid droplets in *ylr404wΔ* to those observed in wild type yeast. (* represents $p = 0.02$.)

Figures 22 and 23 show quantification of different types of TAG in lipid droplets in the wild type and genetically modified yeast. (* represents $p = 0.05$.)

Figure 24 provides a summary of the morphologies of lipid droplets in in the wild type and genetically modified yeast. The phrase "Not numbers" indicates that *A. thaliana* SEIPIN does not restore the number of lipid droplets in *ylr404wΔ* to those found in the wild type yeast. The phrase "Not size" indicates that *A. thaliana* SEIPIN does not restore the size of lipid droplets in *ylr404wΔ* to those found in the wild type yeast. The phrase "↑ numbers" indicates that *A. thaliana* SEIPIN increases the number of lipid droplets in *ylr404wΔ* when expressed therein; and the phrase "↑ size" indicates that *A. thaliana* SEIPIN increases the size of lipid droplets in *ylr404wΔ* when expressed therein.

Figure 25 shows schematic representation of transient expression of exogenous genes in *N. benthamiana*.

Figure 26 shows RT-PCR confirming the expression of exogenous genes in *N. benthamiana*.

Figure 27 shows lipid droplet and chloroplast staining of various *N. benthamiana* lines expressing exogenous genes. Red autofluorescence is marking chloroplasts; green fluorescence is from the neutral-lipid-specific stain–BODIPY 493/503, showing the accumulation of lipid droplets in leaves.

Figure 28 shows average number of lipid droplets in various *N. benthamiana* lines expressing exogenous genes.

I: Mock.

II: 35S:P19.

III: 35S:P19+35S:AtSEIPIN1.

IV: 35S:P19+35S:AtSEIPIN2.

V: 35S:P19+35S:AtSEIPIN3.

VI: 35S:P19+35S:AtSEIPIN1+35S:AtSEIPIN2.

VII: 35S:P19+35S:AtSEIPIN1+35S:AtSEIPIN3.

VIII: 35S:P19+35S:AtSEIPIN2+35S:AtSEIPIN3.

IX: 35S:P19+35S:AtSEIPIN1+35S:AtSEIPIN2+35S:AtSEIPIN3.

X: 35S:P19+35S:AtLEC2, XI: 35S:P19+35S:AtLEC2+35S:AtSEIPIN1.

XII: 35S:P19+35S:AtLEC2+35S:AtSEIPIN2.

IX: 35S:P19+35S:AtLEC2+35S:AtSEIPIN3.

XIV: 35S:P19+35S:AtLEC2+35S:AtSEIPIN1+35S:AtSEIPIN2+35S:AtSEIPIN3.

(# $0.005 < p < 0.05$, * $p < 0.005$.)

Figure 29 shows average number of lipid droplets of various sizes in various *N. benthamiana* lines expressing exogenous genes.

I: Mock.

II: 35S:P19.

III: 35S:P19+35S:AtSEIPIN1.

IV: 35S:P19+35S:AtSEIPIN2.

V: 35S:P19+35S:AtSEIPIN3.

VI: 35S:P19+35S:AtSEIPIN1+35S:AtSEIPIN2.

VII: 35S:P19+35S:AtSEIPIN1+35S:AtSEIPIN3.

VIII: 35S:P19+35S:AtSEIPIN2+35S:AtSEIPIN3.

IX: 35S:P19+35S:AtSEIPIN1+35S:AtSEIPIN2+35S:AtSEIPIN3.

X: 35S:P19+35S:AtLEC2, XI: 35S:P19+35S:AtLEC2+35S:AtSEIPIN1.

XII: 35S:P19+35S:AtLEC2+35S:AtSEIPIN2.

IX: 35S:P19+35S:AtLEC2+35S:AtSEIPIN3.

XIV: 35S:P19+35S:AtLEC2+35S:AtSEIPIN1+35S:AtSEIPIN2+35S:AtSEIPIN3.

(# $0.005 < p < 0.05$, * $p < 0.005$.)

Figure 30 shows lipid droplet and chloroplast staining of various *N. benthamiana* lines expressing exogenous genes.

Figure 31 shows transient expression of mouse FIT2 in *N. benthamiana* leaf tissue. Top left panel shows leaves transfected with empty vector, bottom left panel shows leaves transfected with 35S-P19, and large panel on the right shows leaves transfected with P19 and mouse FIT2. The presence of green fluorescence in P19 and mouse FIT2 transfected leaves indicates accumulation of lipid droplets in these leaves.

Figure 32 shows transient expression of *A. thaliana* LEC2 in *N. benthamiana* leaf tissue. Red autofluorescence is marking chloroplasts; green fluorescence is from the neutral-lipid-specific stain–BODIPY 493/503, showing the accumulation of lipid droplets in leaves. Top left panel shows leaves transfected with empty vector, bottom left panel shows leaves transfected with 35S-P19, and large panel on the right shows leaves transfected with P19 and

A. thaliana LEC2. The presence of green fluorescence in P19 and *A. thaliana* LEC2 transfected leaves indicates accumulation of lipid droplets in these leaves.

Figure 33 shows transient expression of GFP-mouse FIT2 in *N. benthamiana* leaf tissue. Top left panel shows green fluorescence originating from GFP-mouse FIT2 marking the ER. Top middle panel shows lipid droplets stained in yellow with Nile Red stain. Top right panel shows overlap of green endoplasmic reticulum fluorescence and yellow lipid droplet staining. Bottom left panel shows overlap of green endoplasmic reticulum fluorescence and yellow lipid droplet staining, further showing red autofluorescence marking chloroplasts. Bottom right panel shows a portion of the bottom left panel magnified to more clearly indicate the colocalization of endoplasmic reticulum and lipid droplets. These figures suggest that GFP-mouse FIT2 colocalize with lipid droplets in *N. benthamiana* leaves.

Figure 34 shows that stable expression of FIT2 increased lipid droplets accumulation in *A. thaliana* leaves. The top left panel shows Nile Red staining of wild type *A. thaliana* leaves and the top right panel shows a portion of the top left panel magnified to more clearly display Nile Red staining. The bottom left panel shows Nile Red staining of *A. thaliana* leaves in which GFP-FIT2 is overexpressed and the bottom right panel shows a portion of the bottom left panel magnified to more clearly display Nile Red staining. Increased Nile Red staining of *A. thaliana* leaves in which GFP-FIT2 is overexpressed indicates that FIT2 causes the accumulation of lipid droplets.

Figure 35 shows expression of GFP-mouse FIT2 in *A. thaliana.* Top left panel shows green fluorescence originating from GFP-mouse FIT2 indicating the ER. Top middle panel shows lipid droplets stained in yellow with Nile Red stain. Top right panel shows overlap of green endoplasmic reticulum fluorescence and yellow lipid droplet staining. Bottom left panel shows overlap of green endoplasmic reticulum fluorescence and yellow lipid droplet staining further showing red autofluorescence marking chloroplasts. These figures suggest that GFP-mouse FIT2 colocalizes with lipid droplets in *A. thaliana* leaves.

Figure 36 shows the oil contents of *A. thaliana* seeds sown on solidified nutrient medium. The total fatty acid content is shown on percent basis. Transgenic plants expressing mouse FSP27 or mouse autosomal dominant ADRP in the T2 or T3 generation are grown. Cgi-58 32 FSP 27, T2 lines 1-4 and cgi-58 32 FSP27, T3 lines 1-4 transgenic plants have a significantly higher average lipid content than that of the non-transformed plants. Values are the means and standard deviations.

BRIEF DESCRIPTION OF THE SEQUENCES

SEQ ID NO:1 is the amino acid sequence of a human fat specific protein 27 (FSP27) (GenBank Accession Q96AQ7).

SEQ ID NO:2 is the amino acid sequence of a mouse fat specific protein 27 (FSP27) (GenBank Accession NP_848460).

SEQ ID NO:3 is the amino acid sequence of a human PLN1 (perilipin 1) (GenBank Accession NP_002657).

SEQ ID NO:4 is the amino acid sequence of a mouse PLN1 (perilipin 1) (GenBank Accession Q96AQ7).

SEQ ID NO:5 is the amino acid sequence of a human PLIN2 (also called autosomal dominant retinitis pigmentosa (ADRP)) (GenBank Accession NP_001106942).

SEQ ID NO:6 is the amino acid sequence of a mouse PLIN2 (also called autosomal dominant retinitis pigmentosa (ADRP)) (GenBank Accession NP_031434).

SEQ ID NO:7 is the amino acid sequence of a human SEIPIN (Bernardinelli-Seip congenital lipodystrophy type 2 protein) (GenBank Accession Q96G97).

SEQ ID NO:8 is the amino acid sequence of a mouse SEIPIN (Bernardinelli-Seip congenital lipodystrophy type 2 protein) (GenBank Accession AAH43023).

SEQ ID NO:9 is the amino acid sequence of a human FIT1 (fat storage-inducing transmembrane protein 1) (GenBank Accession A5D6W6).

SEQ ID NO:10 is the amino acid sequence of a mouse FIT1 (fat storage-inducing transmembrane protein 1) (GenBank Accession NP_081084).

SEQ ID NO:11 is the amino acid sequence of a human FIT2 (fat storage-inducing transmembrane protein 2) (GenBank Accession Q8N6M3).

SEQ ID NO:12 is the amino acid sequence of a mouse FIT2 (fat storage-inducing transmembrane protein 2) (GenBank Accession NP_775573).

SEQ ID NO:13 is the mRNA sequence of the At4g24160 gene (GenBank Accession BT029749).

SEQ ID NO:14 is the amino acid sequence of the full length polypeptide encoded at the At4g24160 locus (GenBank Accession ABM06019).

SEQ ID NO:15 is the amino acid sequence of a diacylglycerol acyltransferase 1 [*Jatropha curcas*] (GenBank Accession ACA49853).SEQ ID NO:16 is the amino acid sequence of a phospholipid:diacylglycerol acyltransferase 1 [*Jatropha curcas*] (GenBank Accession AED91921).

SEQ ID NO:17 is the amino acid sequence of a phospholipid:diacylglycerol acyltransferase 1 [*Laccaria bicolor*] (GenBank Accession EDR11533).

SEQ ID NO:18 is the amino acid sequence of a phospholipid:diacylglycerol acyltransferase 1 [*Scheffersomyces stipitis*] (GenBank Accession ABN67418).

SEQ ID NO:19 is the amino acid sequence of an adipose triglyceride lipase [*Homo sapiens*] (GenBank Accession AAW81962).

SEQ ID NO:20 is the amino acid sequence of an adipose triglyceride lipase [*Mus musculus*] (GenBank Accession AAW81963).

SEQ ID NO:21 is the amino acid sequence of a cell death activator [*Homo sapiens*] (GenBank Accession AAQ65241).

SEQ ID NO:22 is the amino acid sequence of a cell death activator [*Mus musculus*] (GenBank Accession NP_031728).

SEQ ID NO:23 is the amino acid sequence of a WRINKLED1 [*A. thaliana*] (GenBank Accession AAP80382).

SEQ ID NO:24 is the amino acid sequence of a cell death activator CIDE-3 [*Danio rerio*](GenBank Accession NP_001038512).

SEQ ID NO:25 is the amino acid sequence of human lysophosphatidic acid acyltransferase alpha (LPAAT) (GenBank Accession NP_116130).

SEQ ID NO:26 is the amino acid sequence of mouse lysophosphatidic acid acyltransferase alpha isoform 1 (GenBank Accession NP_001156851).

SEQ ID NO:27 is the amino acid sequence of mouse Glycerol-3-phosphate acyltransferase 1, mitochondrial (GenBank Accession NP_032175).

SEQ ID NO:28 is the amino acid sequence of wild boar (*Sus scrofa*) Glycerol-3-phosphate acyltransferase 1, partial (GenBank Accession AAP74372).

SEQ ID NO:29 is the amino acid sequence of mouse Complement factor D (adipsin) (GenBank Accession AAI38780).

SEQ ID NO:30 is the amino acid sequence of wild boar (*Sus scrofa*) Complement factor D (adipsin), partial (GenBank Accession AAQ63882).

SEQ ID NO:31 is the amino acid sequence of mouse phosphatidate phosphatase PLIN1 isoform a (GenBank Accession NP_001123884).

SEQ ID NO:32 is the amino acid sequence of mouse phosphatidate phosphatase PLIN2 isoform 1 (GenBank Accession NP_001158357).

SEQ ID NO:33 is the amino acid sequence of *A. thaliana* SEIPIN1 (GenBank Accession AED92296).

SEQ ID NO:34 is the amino acid sequence of *A. thaliana* SEIPIN2 (GenBank Accession AEE31126).

SEQ ID NO:35 is the amino acid sequence of *A. thaliana* SEIPIN3 (GenBank Accession AEC08966).

SEQ ID NO:36 is the amino acid sequence of *A. thaliana* LEC2 (GenBank Accession ABE65660).

SEQ ID NO:37 is the amino acid sequence of tomato bushy stunt virus P19 protein (GenBank Accession AEC08966).

DETAILED DISCLOSURE OF THE INVENTION

In some embodiments, the present invention relates the use of proteins associated with lipid metabolism originated from animals or plants to elevate the lipid content in vegetative tissues (such as leaves) of plants. In certain embodiments, the proteins or polypeptides associated with lipid metabolism useful according to the present invention are of mammalian origin.

As lipid has more than twice the energy content of carbohydrate or protein, the present invention can be used to increase energy content in crop biomass, useful for production of biofuel, renewable chemical feedstocks, animal feed, and nutritional products. The term “lipid,” as used throughout, encompasses oils (such as triglyceride), and in some embodiments “lipid” is oil.

For the purpose of this invention, the term “protein or polypeptide associated with lipid metabolism” refers to a protein or polypeptide which is a “lipid droplet-associated protein or polypeptide,” “endoplasmic reticulum (ER) associated protein or polypeptide that localizes to domains of ER that form lipid droplets,” “lipid droplet forming protein or polypeptide,” or "lipid forming protein or polypeptide.” In some embodiments, a protein associated with lipid metabolism, designated as fat storage protein 27 (FSP27), is expressed in leaves of transgenic *Arabidoposis thaliana* plants.

Neutral lipid-specific fluorescent staining of cystolic lipid droplets reveals a marked increase in the number and size of lipid droplets in the mesophyll cells of the levels of transgenic plants, when compared with non-transformed plants of the same type. The expression of a fluorescent-tagged mouse FSP27 protein in transgenic plants shows the

FSP27 protein associated with the lipid droplets in plant cells, similar to that of mouse adipocytes. When the FSP27 protein is expressed in the Arabidopsis *cgi58* mutant background, lipid droplet formation and lipid content in leaves are further augmented, when compared to transgenic Arabidopsis plants that only express FSP27 or Arabidopsis *cgi58* mutant.

In some embodiments, the present invention provides a method of elevating lipid content in a plant or plant part by genetically modifying the plant to express a protein or polypeptide associated with lipid metabolism (such as fat-specific protein 27) of animal origin in the plant or plant part. In one specific embodiment, the present invention provides a method of elevating lipid content in vegetative (non-seed) plant tissues.

In some embodiments, the present invention also provides genetically-modified algal cells, plant cells, tissues, or whole plants with elevated cellular lipid content, wherein the algal cells, plant cells, tissues or whole plants express a protein or polypeptide associated with lipid metabolism (such as fat-specific protein 27) of animal origin or plant origin.

Genetically-Modified Plants with Elevated Lipid Content and/or Lipid Droplet Production

In some embodiments, the present invention provides a method for obtaining a plant cell or an algal cell with elevated lipid content, wherein the method comprises:

genetically modifying the plant cell or the algal cell to express an exogenous protein or polypeptide associated with lipid metabolism, thereby obtaining a genetically-modified plant cell with elevated lipid content;

wherein the protein or polypeptide associated with lipid metabolism induces adipogenesis, enhances the accumulation of cellular lipid droplets, and/or reduces lipase activity; and

wherein the expression of the protein or polypeptide associated with lipid metabolism increases lipid content of the genetically-modified plant cell or algal cell, when compared to a wild-type (native) plant cell or algal cell of the same type.

In some embodiments, the present invention provides a method for obtaining a plant cell or an algal cell with elevated lipid content, wherein the method comprises:

transforming the plant cell or the algal cell with a vector comprising a nucleic acid sequence encoding an exogenous protein or polypeptide associated with lipid metabolism,

yielding a transformed cell wherein the nucleic acid is operably linked to a promoter and/or a regulatory sequence;

wherein the protein or polypeptide associated with lipid metabolism induces adipogenesis, enhances the accumulation of cellular lipid droplets, and/or reduces lipase activity;

wherein the transformed plant cell or algal cell expresses the protein or polypeptide associated with lipid metabolism; and

wherein the expression of the protein or polypeptide associated with lipid metabolism increases lipid content of the transformed plant cell or algal cell as compared to a wild-type (native) plant cell or algal cell of the same type.

In certain embodiments, the genetically-modified plant cell is contained in an algal cell, a plant tissue, plant part, or whole plant.

In some embodiments, the genetically-modified plant cell comprises, in its genome, a nucleic acid molecule encoding a protein or polypeptide associated with lipid metabolism.

In some embodiments, the protein or polypeptide associated with lipid metabolism is not of plant origin. In certain embodiments, the protein or polypeptide associated with lipid metabolism is of animal origin, such as of insect, vertebrate, amphibian, or mammalian (*e.g.*, mouse, human) origin. In another embodiment, the protein or polypeptide associated with lipid metabolism is of plant origin.

In some embodiments, a T-DNA binary vector system is used for plant transformation. A T-DNA binary vector system is a pair of plasmids consisting of a binary plasmid and a helper plasmid. In one embodiment, the T-DNA region located on the binary vector comprises a vector nucleic acid sequence encoding an exogenous protein or polypeptide associated with lipid metabolism.

T-DNA binary vector systems are routinely used in plant transformation. A variety of vectors and expression cassettes useful for performing plant transformation are described in Curtis and Grossniklaus (2003), which is herein incorporated by reference in its entirety. Non-limiting examples of vectors and expression cassettes useful in accordance with the present invention include pMDC32, pMDC7, pMDC30, pMDC45, pMDC44, pMDC43, pMDC83, pMDC84, pMDC85, pMDC139, pMDC140, pMDC141, pMDC107, pMDC111, pMDC110, pMDC162, pMDC163, pMDC164, pMDC99, pMDC100, and pMDC123.

In some embodiments, plant transformation is performed using the floral dip method, as describe in Bent and Clough (1998), which is herein incorporated by reference in its entirety.

In certain embodiments, to elevate cellular lipid content and/or to induce lipid droplet production, the plant cell can be genetically engineered to expresses one or more proteins or polypeptides associated with lipid metabolism including, but not limited to, fat specific protein 27 (FSP27); perilipins including PLIN1 (perilipin 1) and PLIN2 (also called autosomal dominant retinitis pigmentosa (ADRP)); SEIPIN (Bernardinelli-Seip congenital lipodystrophy type 2 protein); FIT1 (fat storage-inducing transmembrane protein 1), and FIT2 (fat storage-inducing transmembrane protein 2); acyl-CoA:diacylglycerol acyltransferase 1 (DGAT-1); phospholipid:diacylglycerol acyltransferase 1 (PDAT-1); cell death activator (Cidea); and WRINKLED1 (WRI1).

In certain specific embodiments, the plant cell or the algal cell can be genetically engineered to express one or more functional domains of the proteins associated with lipid metabolism, wherein the functional domain is involved lipid metabolism, including, but not limited to, the synthesis, protection, accumulation, storage, or breakdown of lipids.

In another embodiment, to elevate cellular lipid content and/or to induce lipid droplet production, the plant cell or the algal cell can be genetically engineered to over-express one or more proteins or polypeptides associated with lipid metabolism of plant origin.

A variety of proteins associated with lipid metabolism are known in the art; amino acid sequences of proteins associated with lipid metabolism, as well as cDNA sequences encoding proteins associated with lipid metabolism, are publically available, such as via the GenBank database.

Fat Specific Protein 27 (FSP27), a lipid droplet (LD) associated protein in adipocytes, regulates triglyceride (TG) storage. FSP27 plays a key role in LD morphology to accumulate TGs. FSP27 facilitates LD clustering and promotes their fusion to form enlarged droplets, resulting in triglyceride accumulation. Functional domains of FSP27 responsible for LD formation have been characterized (see Jambunathan *et al.*, 2011, which is hereby incorporated by reference in its entirety). Specifically, amino acids 173–220 of human FSP27 are necessary and sufficient for both the targeting of FSP27 to LDs and the initial clustering of the droplets. Amino acids 120–140 of human FSP27 are essential but not sufficient for LD enlargement, whereas amino acids 120–210 of human FSP27 are necessary and sufficient for both clustering and fusion of LDs to form enlarged droplets. In addition,

FSP27-mediated enlargement of LDs, but not their clustering, is associated with triglyceride accumulation. CIDEC (human ortholog of FSP27) results in the accumulation of multiple, small LD's in white adipocytes *in vivo*.

In certain embodiments, the plant cell or the algal cell can be genetically engineered to express one or more functional domains of FSP27, including, but not limited to, amino acids 173-220 of human FSP27, amino acids 120-140 of human FSP27, amino acids 120-210 of human FSP27, or any fragment having no fewer than 10 consecutive amino acids (such as, more than 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 consecutive amino acids) of the aforementioned functional domains.

In certain embodiments, the plant cell or the algal cell can be genetically engineered to express a FSP protein or peptide that corresponds to amino acids 120-220 of mouse FSP27 of SEQ ID NO:2 (GenBank Accession No. NP_848460), or any fragment thereof having no fewer than 10 consecutive amino acids (such as, more than 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 consecutive amino acids).

Members of the PAT family (also called the perilipin (PLIN) family), which regulate lipolysis, are a family of proteins associated with lipid metabolism that have been well characterized in the art. Perilipins function as a protective coating from the body's natural lipases, such as hormone-sensitive lipase, which break triglycerides into glycerol and free fatty acids for use in metabolism - a process called lipolysis.

Acyl-CoA:diacylglycerol acyltransferase 1 (DGAT-1) and phospholipid:diacylglycerol acyltransferase 1 (PDAT-1) proteins are essential for triacylglyceride (Oil) biosynthesis in plants and seeds. DGAT-1 is also responsible for triglyceride biosynthesis in mammals. See Zhang *et al.* (2009) Plant Cell 21, 3885-901, PMID: 20040537, which is hereby incorporated as reference in its entirety.

Mutations in *cgi58* (plant ortholog is also called *cgi58*) can be used to increase in plant oil contents. See James *et al.* (2010) PNAS 107, 17833-1838, PMID: 20876112, which is hereby incorporated as reference in its entirety.

Yeast gene SEIPIN (human ortholog is also called SEIPIN) can be used to increase the size of oil droplets in mammalian cells. See Szymanski *et al.* (2007) PNAS 104, 20890-5, PMID: 18093937, which is hereby incorporated as reference in its entirety.

FIT1 and FIT2 proteins, which belong to the FIT family (also have orthologues in yeast), play an important role in lipid droplet formation. Gross *et al.* (2011) PNAS 108, 19581-19586; PMID: 22106267, which is hereby incorporated as reference in its entirety.

Mammalian genes PLIN1 and PLIN2 play a role in protecting against breakdown of fat (called hydrolysis or lipolysis).

Cgi58 activate lipases (*e.g.*, adipose triglyceride lipase (ATGL)), which catalyze the breakdown of lipids.

Cell death activator (Cidea), a novel gene identified by the inventors, plays a role in triglyceride accumulation in humans.

In certain embodiments, the plant cell or the algal cell can be genetically engineered to expresses any combinations of proteins associated with lipid metabolism and peptides including, but not limited to, fat specific protein 27 (FSP27); perilipins including PLIN1 (perilipin 1) and PLIN2 (also called autosomal dominant retinitis pigmentosa (ADRP)); SEIPIN (Bernardinelli-Seip congenital lipodystrophy type 2 protein); FIT1 (fat storage-inducing transmembrane protein 1), and FIT2 (fat storage-inducing transmembrane protein 2); acyl-CoA:diacylglycerol acyltransferase 1 (DGAT-1); phospholipid:diacylglycerol acyltransferase 1 (PDAT-1); cell death activator (Cidea); and WRINKLED1 (WRI1).

In one embodiment, the plant cell can be genetically engineered to expresses one or more proteins associated with lipid metabolism in a *cgi58* (mutation) background, wherein the *cgi58* (mutation) background results in enhanced lipid/oil content in plants.

In certain specific embodiments, the transgenic plants or algae express a combination of nucleic acids expressing proteins associated with lipid metabolism selected from: DGAT-1 and FSP27; DGAT-1, *cgi58* (mutation), and FSP27; DGAT-1, PDAT-1, and FSP27; DGAT-1, PDAT-1, *cgi58* (mutation), FSP27; FSP27, PLIN2, and *cgi58* (mutation); DGAT-1, FSP27, PLIN2, and *cgi58* (mutation); and DGAT-1, PDAT-1, FSP27, PLIN2, and *cgi58* (mutation). In some embodiments, any protein or polypeptide associated with lipid metabolism of animal origin can be used in accordance with the present invention. In certain embodiments, suitable proteins or polypeptides associated with lipid metabolism can be originated from insects, fish, birds, vertebrates, amphibians, and mammalian species including, but not limited to apes, chimpanzees, orangutans, humans, monkeys, dogs, cats, horses, cattle, pigs, sheep, goats, chickens, mice, rats, guinea pigs, and hamsters.

In certain embodiments, the plant cell or the algal cell can be genetically engineered to expresses a protein or polypeptide associated with lipid metabolism comprising any of

SEQ ID NOs: 1-12 and 14-36, a homolog or variant thereof, or a functional fragment of a protein or polypeptide associated with lipid metabolism comprising any of SEQ ID NOs: 1-12, 14-36 or a homolog or variant thereof, wherein the functional variant and the functional fragment induces adipogenesis, enhances the accumulation of cellular lipid droplets, and/or reduces lipase activity.

In certain embodiments, a variant of a protein or polypeptide associated with lipid metabolism comprising a sequence of SEQ ID NOs:1-12, 14-36 comprises an amino acid sequence that may share about at least 70%, 75%, 80%, 85%, 90%, 95%, 98%, 99%, or greater sequence similarity at the respective amino acid sequence of SEQ ID NOs:1-12, 14-36.

The term “homolog," as used herein, refers to genes or proteins related to each other by descent from a common ancestral DNA (such as genes) or protein sequence. In certain embodiments, a homolog of a protein or polypeptide associated with lipid metabolism comprising a sequence of SEQ ID NOs:1-12, 14-36 comprises an amino acid sequence that may share about at least 70%, 75%, 80%, 85%, 90%, 95%, 98%, 99%, or greater sequence similarity at the respective amino acid sequence of SEQ ID NOs:1-12, 14-36.

The sequence identity will typically be greater than 75%, preferably greater than 80%, more preferably greater than 90%, and can be greater than 95%. The identity and/or similarity of a sequence can be 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, or 99% as compared to a sequence exemplified herein.

Unless otherwise specified, as used herein percent sequence identity and/or similarity of two sequences can be determined using the algorithm of Karlin and Altschul (1990), modified as in Karlin and Altschul (1993). Such an algorithm is incorporated into the NBLAST and XBLAST programs of Altschul *et al.* (1990). BLAST searches can be performed with the NBLAST program, score = 100, word length = 12, to obtain sequences with the desired percent sequence identity. To obtain gapped alignments for comparison purposes, Gapped BLAST can be used as described in Altschul *et al.* (1997). When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (NBLAST and XBLAST) can be used. *See* NCBI/NIH website.

Furthermore, as various proteins associated with lipid metabolism have been well characterized in the art, a skilled artisan can readily make modifications to native or naturally-occurring sequences without substantially affecting their function of regulating lipid metabolism. In certain embodiments, the present invention relates to use of proteins or

polypeptides associated with lipid metabolism comprising no more than 150, 140, 130, 120, 110, 100, 90, 80, 70, 60, 50, 40, 30, 25, 20, 15, 10, 9, 8, 7, 6, 5, 4, 3, 2, or 1 conservative modification(s) (*e.g.*, conservative substitutions, additions, deletions) to any of naturally-occurring sequences, such as SEQ ID NOs:1-12, 14-36.

In addition, the present invention relates to the use of functional fragments of naturally-occurring proteins or polypeptides associated with lipid metabolism. In certain embodiments, the functional fragments comprise at least 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 75, 80, 85, 90, 95, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 280, 300, 330, or 350 consecutive amino acids of any of SEQ ID NOs:1-12, 14-36.

In certain embodiments, plant species that can be genetically-modified in accordance with the current invention include, but are not limited to, monocots, dicots, crop plants (*i.e.*, any plant species grown for purposes of agriculture, food production for animals including humans), trees (*i.e.*, fruit trees, trees grown for wood production, trees grown for decoration, etc.), flowers of any kind (*i.e.*, plants grown for purposes of decoration, for example, following their harvest), and cacti. More specific examples of plants that can be genetically-modified to express one or more proteins or polypeptides associated with lipid metabolism include, but are not limited to, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta, Euphyllophytes, Spermatophyta, Magnoliophyta, Liliopsida, Commelinidae, Poales, Poaceae, Oryza, *Oryza sativa*, Zea, *Zea mays*, Hordeum, *Hordeum vulgare*, Triticum, *Triticum aestivum*, Eudicotyledons, Core eudicots, Asteridae, Euasterids, Rosidae, Eurosids II, Brassicales, Brassicaceae, Arabidopsis, Magnoliopsida, Solananae, Solanales, Solanaceae, Solanum, and Nicotiana. Thus, the embodiments of the invention have uses over a broad range of plants including, but not limited to, species from the genera *Anacardium, Arachis, Asparagus, Atropa, Avena, Brassica, Citrus, Citrullus, Capsicum, Carthamus, Cocos, Coffea, Cucumis, Cucurbita, Daucus, Elaeis, Fragaria, Glycine, Gossypium, Helianthus, Heterocallis, Hordeum, Hyoscyamus, Lactuca, Linum, Lolium, Lupinus, Lycopersicon, Malus, Manihot, Majorana, Medicago, Nicotiana, Olea, Oryza, Panieum, Panneserum, Persea, Phaseolus, Pistachia, Pisum, Pyrus, Prunus, Raphanus, Ricinus, Secale, Senecio, Sinapis, Solanum, Sorghum, Theobromus, Trigonella, Titicum, Vicia, Vitis, Vigna,* and *Zea*.

In certain embodiments, plant species that can be genetically-modified in accordance with the current invention include, but are not limited to, corn, sugarcane, sorghum, millet, rice, wheat, barley, soybean, olive, peanut, castor, oleaginous fruits such as palm and

avocado, Glycine sp., grape, canola, *Arabidopsis, Brassica sp.*, cotton, tobacco, bamboo, sugar beet, sunflower, willow, switchgrass (Panicum virgatum), giant reed (Arundo donax), reed canarygrass (Phalaris arundinacea), Miscanthus crossed with giganteus (Miscanthus X giganteus), Miscanthus sp., Sericea lespedeza (Lespedeza cuneata), ryegrass (Lolium multiflorum, lolium sp.), timothy, kochia (Kochia scoparia), forage soybeans, alfalfa, clover, turf grass, sunn hemp, kenaf, bahiagrass, bermudagrass, dallisgrass, pangolagrass, big bluestem, indiangrass, fescue (Festuca sp.) including tall fescue, Dactylis sp., Brachypodium distachyon, smooth bromegrass, orchardgrass, kentucky bluegrass, yellow nutsedge, pine, poplar (Populus sp.), and eucalyptus, among others.

In certain specific embodiments, plant species that can be genetically-modified in accordance with the current invention include, but are not limited to, sorghum; switchgrass (panicum); wheat (triticum); sugarcane (for expression in leaves and stems); camelina, canola (for expression in oil seeds); soybean; safflower; and jatropha (*e.g.*, for expression in seeds).

In certain embodiments, plant species that can be genetically-modified in accordance with the current invention include grasses such as the Poaceae (or Gramineae) family, the sedges (Cyperaceae), and the rushes (Juncaceae).

While *A. thaliana* is used in the present invention as an example of plant species to demonstration that plants transformed with proteins associated with lipid metabolism have elevated cellular lipid content and/or increased lipid droplet formation, those skilled in the art would readily obtain transgenic plants of other species with elevated cellular lipid content and/or increased lipid droplet formation, wherein transgenic plants express proteins associated with lipid metabolism.

Triacylglycerols (TG) can be synthesized in non-seed tissues; however, their abundance is low and these storage lipids are presumed to be metabolized rapidly, perhaps for the recycling of fatty acids for energy or the synthesis of membrane lipids.

In certain embodiments, the algal cells that can be genetically modified in accordance with the current invention include, but are not limited to, algae selected from Amphora, Anabaena, Anikstrodesmis, Botryococcus, Chaetoceros, Chlorella, Chlorococcum, Cyclotella, Cylindrotheca, Euglena, Hematococcus, Isochrysis, Monodus, Monoraphidium, Nannochloris, Nannochloropsis, Navicula, Nephrochloris, Nephroselmis, Nitzschia, Nodularia, Nostoc, Oochromonas, Oocystis, Oscillartoria, Parachlorella, Pavlova, Phaeodactylum, Pinguiococcus, Playtomonas, Pleurochrysis, Porphyra, Pseudoanabaena, Pyramimonas, Rhodomonas, Selenastrum, Scenedesmus, Sticococcus, Synechococcus,

Tetraselmis, Thalassiosira, and Trichodesmium. In certain embodiments, the algal cells are selected from *Botryococcus braunii, Chlorella spp., Dunaliella tertiolecta, Gracilaria spp., Pleurochrysis carterae* (also called CCMP647), *Sargassum spp., Ankistrodesmus spp., Botryococcus braunii, Chlorella protothecoides, Cyclotella DI- 35, Dunaliella tertiolecta, Hantzschia DI-160, Nannochloris spp., Nannochloropsis spp., Nitzschia TR-114, Phaeodactylum tricornutum, Scenedesmus TR-84, Stichococcus spp., Tetraselmis suecica, Thalassiosira pseudonana, Crypthecodinium cohnii, Neochloris oleoabundans,* and *Schiochytrium spp.*

In certain embodiments, the present invention provides a method of elevating lipid content and/or inducing lipid droplet accumulation in vegetative plant (non-seed) tissues or plant parts including, but not limited to, leaves, roots, stems, shoots, buds, tubers, fruits, and flowers. In another embodiment, the present invention provides elevated lipid content and/or induces lipid droplet accumulation in seeds.

In some embodiments, the present invention can be used to increase total fatty acid content of the plant cell or the algal cell. In certain embodiments, the present invention can be used to increase the level of fatty acids including leaf-specific fatty acids, including but not limited to, triacylglycerol, hydroxyl, epoxy, cyclic, acetylenic, saturated, polyunsaturated (such as omega-3, omega-6 fatty acids), and short-chain or long-chain fatty acids, which can be incorporated into neutral lipids that can be compartmentalized in lipid droplets, including TAGs, wax-esters, and steryl-esters.

In some embodiments, the method for obtaining a plant cell or an algal cell with elevated lipid content further comprises: downregulating, in the plant cell or the algal cell, the function of an At4924160 gene product.

Chanarin-Dorfman Syndrome is a neutral-lipid storage disorder (Lefevre et al., 2001; Bruno et al., 2008). CGI58, also known as ABHD5, associates with lipid droplets in human cells and participates in storage lipid hydrolysis. A mutation in this protein results in hyperaccumulation of lipid droplets in cells and the pathology associated with this syndrome. The CGI58 protein sequence includes a so-called "alpha/beta hydrolase fold" that is shared by members of the esterase/lipase/thioesterase family, suggesting that it might be a TAG lipase. Recent analyses of its functional properties have indicated that the mammalian polypeptide stimulates the activity of a lipase called ATGL (Adipose Triglyceride Lipase), which is the major lipase responsible for catalyzing the initial step of TAG breakdown in both adipose and non-lipid storing cell types (e.g. Lass et al., 2006; Yen & Farese, 2006; Schweiger et al.,

2006; Yamaguchi et al., 2007). Interestingly, CGI58 also possesses lysophosphatidic acid acyltransferase (LPAAT) activity in vitro, suggesting that, in addition to its role in stimulating lipase activity, it may play a role in recycling of fatty acids into membrane phospholipids (Ghosh et al., 2008).

At4g24160 has been identified as a putative homolog of human CGI58, in *A. thaliana*. The gene in Arabidopsis is apparently expressed as two alternative transcripts (two distinct cDNAs corresponding to the same gene have been identified) and the predicted protein products share domain architecture with other lipases/esterases and acyltransferases. Arabidopsis mutant lines lacking the function of the CGI58 homolog (i.e., At4g24160) contained vegetative (i.e. non-seed) tissues with metabolic machinery capable of synthesizing and storing oil as TAG, demonstrating that there are mechanisms in place to regulate this process in non-seed tissues.

The term "down-regulating," as used herein, refers to reducing the expression or function of a gene of interest. In certain embodiments, the reduction in expression or function of a gene of interest may be least a 25%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 100%, when compared to wild-type. The down-regulation of function may also be measured by assaying the enzymatic activity of a polypeptide that is regulated by a polypeptide encoded by the gene of interest.

In certain embodiments of the invention, down-regulation of the activity of a polypeptide encoded by a gene may be accomplished using antisense-mediated-, or dsRNA-mediated-, or other forms of RNA-mediated-interference (RNAi), as is well known in the art. Methods for identification of candidate nucleotide sequences for RNA-mediated gene suppression, and design of oligonucleotides and constructs to achieve RNA-mediated gene suppression, are well known (e.g. Reynolds et al., 2004; Lu and Mathews, 2008).

In one embodiment, the plant cell can be genetically engineered to expresses one or more proteins associated with lipid metabolism in a *cgi58* (mutation) background, wherein the CGI58 (mutation) background results in enhanced lipid content in plants. In one embodiment, the plant cell of the present invention has a *cgi58* (mutation) background described in US2010/0221400.

Methods for the genetic control of lipid accumulation in vegetative (non-seed) portions of plants by down-regulation of activity of At4g24160 or a homolog thereof are described in US2010/0221400, which is herein incorporated by reference in its entirety.

In certain embodiments, the present invention provides a transgenic plant cell or an algal cell with elevated lipid content, wherein the transgenic plant or algal cell expresses an exogenous protein or polypeptide associated with lipid metabolism, wherein the protein associated with lipid metabolism induces adipogenesis, enhances the accumulation of cellular lipid droplets, and/or reduces lipase activity; and wherein the expression of the protein or polypeptide associated with lipid metabolism increases lipid content of the genetically-modified plant or algal cell, when compared to a wild-type plant cell. In certain embodiments, the genetically-modified plant cell is contained in a plant tissue, plant part, or whole plant. In one embodiment, the genetically-modified plant or algal cell comprises, in its genome, a transgene encoding a protein or polypeptide associated with lipid metabolism that induces adipogenesis, enhances the accumulation of cellular lipid droplets, and/or reduces lipase activity.

As used herein, the term "genetically modified plant or plant parts" refers to a plant or a plant part, whether it is attached or detached from the whole plant. It also includes progeny of the genetically modified plant or plant parts that are produced through sexual or asexual reproduction. Similarly, "transformed plant cell" refers to the initial transformant as well as progeny cells of the initial transformant in which the heterologous genetic sequence is found.

"Progeny" includes the immediate and all subsequent generations of offspring traceable to a parent.

In some embodiments, the present invention provides a method for obtaining an algal or bacterial cell with elevated lipid content, wherein the method comprises:

genetically modifying an algal or bacterial cell to express an exogenous protein or polypeptide associated with lipid metabolism, thereby obtaining a genetically-modified algae or bacterial cell with elevated lipid content;

wherein the protein or polypeptide associated with lipid metabolism induces adipogenesis, enhances the accumulation of cellular lipid droplets, and/or reduces lipase activity; and

wherein the expression of the protein or polypeptide associated with lipid metabolism increases lipid content of the genetically-modified algal or bacterial cell, when compared to a wild-type (native) algal or bacterial cell of the same type.

In another embodiment, the present invention provides a method for obtaining an algal or bacterial cell with elevated lipid content, wherein the method comprises:

transforming an algal or bacterial cell with a vector comprising a nucleic acid sequence encoding an exogenous protein or polypeptide associated with lipid metabolism, wherein nucleic acid is operably linked to a promoter and/or a regulatory sequence;

wherein the protein associated with lipid metabolism induces adipogenesis, enhances the accumulation of cellular lipid droplets, and/or reduces lipase activity;

wherein the transformed algal or bacterial cell expresses the protein or polypeptide associated with lipid metabolism; and

wherein the expression of the protein or polypeptide associated with lipid metabolism increases lipid content of the transformed algal or bacterial cell as compared to a wild-type (native) algal or bacterial cell of the same type.

In certain embodiments, the algal cell can be genetically engineered to expresses any combinations of proteins associated with lipid metabolism and peptides including, but not limited to, fat specific protein 27 (FSP27); perilipins including PLIN1 (perilipin 1) and PLIN2 (also called autosomal dominant retinitis pigmentosa (ADRP)); SEIPIN (Bernardinelli-Seip congenital lipodystrophy type 2 protein); FIT1 (fat storage-inducing transmembrane protein 1), and FIT2 (fat storage-inducing transmembrane protein 2); acyl-CoA:diacylglycerol acyltransferase 1 (DGAT-1); phospholipid:diacylglycerol acyltransferase 1 (PDAT-1); cell death activator (Cidea); and WRINKLED1 (WRI1). In some embodiments, algae can be selected from the group consisting of *Amphora, Anabaena, Anikstrodesmis, Botryococcus, Chaetoceros, Chlorella, Chlorococcum, Cyclotella, Cylindrotheca, Euglena, Hematococcus, Isochrysis, Monodus, Monoraphidium, Nannochloris, Nannochloropsis, Navicula, Nephrochloris, Nephroselmis, Nitzschia, Nodularia, Nostoc, Oochromonas, Oocystis, Oscillartoria, Parachlorella, Pavlova, Phaeodactylum, Pinguiococcus, Playtomonas, Pleurochrysis, Porphyra, Pseudoanabaena, Pyramimonas, Rhodomonas, Selenastrum, Scenedesmus, Sticococcus, Synechococcus, Tetraselmis, Thalassiosira,* and *Trichodesmium*.

Nucleic Acid Constructs, Expression Cassettes, and Host Cells

The terms "polynucleotide" and "nucleic acid," used interchangeably herein, refer to a polymeric form of nucleotides of any length, either ribonucleotides or deoxynucleotides. Thus, this term includes, but is not limited to, single-, double-, or multi-stranded DNA or RNA, genomic DNA, cDNA, DNA-RNA hybrids, or a polymer comprising purine and

pyrimidine bases or other natural, chemically or biochemically modified, non-natural, or derivatized nucleotide bases.

As used herein, the terms "operon" and "single transcription unit" are used interchangeably to refer to two or more contiguous coding regions (nucleotide sequences that encode a gene product such as an RNA or a protein) that are coordinately regulated by one or more controlling elements (e.g., a promoter).

As used herein, the term "gene product" refers to RNA encoded by DNA (or *vice versa*) or protein that is encoded by an RNA or DNA, where a gene will typically comprise one or more nucleotide sequences that encode a protein, and may also include introns and other non-coding nucleotide sequences.

The terms "peptide," "polypeptide," and "protein" are used interchangeably herein, and refer to a polymeric form of amino acids of any length, which can include coded and non-coded amino acids, chemically or biochemically modified or derivatized amino acids, and polypeptides having modified peptide backbones.

The term "naturally-occurring" or “native” as used herein as applied to a nucleic acid, a cell, or an organism, refers to a nucleic acid, cell, or organism that is found in nature. For example, a polypeptide or polynucleotide sequence that is present in an organism (including viruses) that can be isolated from a source in nature and which has not been intentionally modified by a human in the laboratory is naturally occurring, and includes “wild-type” plants.

The term "heterologous nucleic acid," as used herein, refers to a nucleic acid wherein at least one of the following is true: (a) the nucleic acid is foreign ("exogenous") to (i.e., not naturally found in) a given host microorganism or host cell; (b) the nucleic acid comprises a nucleotide sequence that is naturally found in (e.g., is "endogenous to") a given host microorganism or host cell (e.g., the nucleic acid comprises a nucleotide sequence endogenous to the host microorganism or host cell); however, in the context of a heterologous nucleic acid, the same nucleotide sequence as found endogenously is produced in an unnatural (e.g., greater than expected or greater than naturally found) amount in the cell, or a nucleic acid comprising a nucleotide sequence that differs in sequence from the endogenous nucleotide sequence but encodes the same protein (having the same or substantially the same amino acid sequence) as found endogenously is produced in an unnatural (e.g., greater than expected or greater than naturally found) amount in the cell; (c) the nucleic acid comprises two or more nucleotide sequences that are not found in the same relationship to each other in nature, e.g., the nucleic acid is recombinant. An example of a

heterologous nucleic acid is a nucleotide sequence encoding a protein or polypeptide associated with lipid metabolism operably linked to a transcriptional control element (for example, a promoter) to which an endogenous (naturally-occurring) sequence coding for a protein or polypeptide associated with lipid metabolism is not normally operably linked. Another example of a heterologous nucleic acid is a high copy number plasmid comprising a nucleotide sequence encoding a protein or polypeptide associated with lipid metabolism. Another example of a heterologous nucleic acid is a nucleic acid encoding a protein or polypeptide associated with lipid metabolism, where a host cell that does not normally produce a protein or polypeptide associated with lipid metabolism is genetically modified with the nucleic acid encoding a protein or polypeptide associated with lipid metabolism; because protein associated with lipid metabolism-encoding nucleic acids are not naturally found in the host cell, the nucleic acid is heterologous to the genetically modified host cell.

"Recombinant," as used herein, means that a particular nucleic acid (DNA or RNA) is the product of various combinations of cloning, restriction, and/or ligation steps resulting in a construct having a structural coding or non-coding sequence distinguishable from endogenous nucleic acids found in natural systems. Generally, DNA sequences encoding the structural coding sequence can be assembled from cDNA fragments and short oligonucleotide linkers, or from a series of synthetic oligonucleotides, to provide a synthetic nucleic acid which is capable of being expressed from a recombinant transcriptional unit contained in a cell or in a cell-free transcription and translation system. Such sequences can be provided in the form of an open reading frame uninterrupted by internal non-translated sequences, or introns, which are typically present in eukaryotic genes. Genomic DNA comprising the relevant sequences can also be used in the formation of a recombinant gene or transcriptional unit. Sequences of non-translated DNA may be present 5' or 3' from the open reading frame, where such sequences do not interfere with manipulation or expression of the coding regions, and may indeed act to modulate production of a desired product by various mechanisms (see "DNA regulatory sequences", below).

Thus, for example, the term "recombinant" polynucleotide or nucleic acid refers to one which is not naturally occurring, for example, is made by the artificial combination of two otherwise separated segments of sequence through human intervention. This artificial combination is often accomplished by either chemical synthesis means, or by the artificial manipulation of isolated segments of nucleic acids, e.g., by genetic engineering techniques. Such is usually done to replace a codon with a redundant codon encoding the same or a

conservative amino acid, while typically introducing or removing a sequence recognition site. Alternatively, it is performed to join together nucleic acid segments of desired functions to generate a desired combination of functions. This artificial combination is often accomplished by either chemical synthesis means, or by the artificial manipulation of isolated segments of nucleic acids, e.g., by genetic engineering techniques.

By "construct" is meant a recombinant nucleic acid, generally recombinant DNA, which has been generated for the purpose of the expression of a specific nucleotide sequence(s), or is to be used in the construction of other recombinant nucleotide sequences.

As used herein, the term "exogenous nucleic acid" refers to a nucleic acid that is not normally or naturally found in and/or produced by a given bacterium, organism, or cell in nature. As used herein, the term "endogenous nucleic acid" refers to a nucleic acid that is normally found in and/or produced by a given bacterium, organism, or cell in nature. An "endogenous nucleic acid" is also referred to as a "native nucleic acid" or a nucleic acid that is "native" to a given bacterium, organism, or cell.

The terms "DNA regulatory sequences," "control elements," and "regulatory elements," used interchangeably herein, refer to transcriptional and translational control sequences, such as promoters, enhancers, polyadenylation signals, terminators, protein degradation signals, and the like, that provide for and/or regulate expression of a coding sequence and/or production of an encoded polypeptide in a host cell.

The terms "transformation" or "transformed" are used interchangeably herein with "genetic modification" or "genetically modified" and refer to a permanent or transient genetic change induced in a cell following introduction of new nucleic acid (i.e., DNA exogenous to the cell). Genetic change ("modification") can be accomplished either by incorporation of the new DNA into the genome of the host cell, or by transient or stable maintenance of the new DNA as an episomal element. Where the cell is a eukaryotic cell, a permanent genetic change is generally achieved by introduction of the DNA into the genome of the cell or into a plastome of the cell. In prokaryotic cells, permanent changes can be introduced into the chromosome or *via* extrachromosomal elements such as plasmids, plastids, and expression vectors, which may contain one or more selectable markers to aid in their maintenance in the recombinant host cell.

"Operably linked" refers to a juxtaposition wherein the components so described are in a relationship permitting them to function in their intended manner. For instance, a promoter is operably linked to a coding sequence if the promoter affects its transcription or

expression. As used herein, the terms "heterologous promoter" and "heterologous control regions" refer to promoters and other control regions that are not normally associated with a particular nucleic acid in nature. For example, a "transcriptional control region heterologous to a coding region" is a transcriptional control region that is not normally associated with the coding region in nature.

A "host cell," as used herein, denotes an *in vivo* or *in vitro* eukaryotic cell, a prokaryotic cell, or a cell from a multicellular organism (for example, a cell line) cultured as a unicellular entity, which eukaryotic or prokaryotic cells can be, or have been, used as recipients for a nucleic acid (for example, an expression vector that comprises a nucleotide sequence encoding one or more gene products such as proteins or polypeptides associated with lipid metabolism), and include the progeny of the original cell which has been genetically modified by the nucleic acid. It is understood that the progeny of a single cell may not necessarily be completely identical in morphology or in genomic or total DNA complement as the original parent, due to natural, accidental, or deliberate mutation. A "recombinant host cell" (also referred to as a "genetically modified host cell") is a host cell into which has been introduced a heterologous nucleic acid, e.g., an expression vector. For example, a subject prokaryotic host cell is a genetically modified prokaryotic host cell (for example, a bacterium), by virtue of introduction into a suitable prokaryotic host cell a heterologous nucleic acid, e.g., an exogenous nucleic acid that is foreign to (not normally found in nature in) the prokaryotic host cell, or a recombinant nucleic acid that is not normally found in the prokaryotic host cell; and a subject eukaryotic host cell is a genetically modified eukaryotic host cell, by virtue of introduction into a suitable eukaryotic host cell a heterologous nucleic acid, for example, an exogenous nucleic acid that is foreign to the eukaryotic host cell, or a recombinant nucleic acid that is not normally found in the eukaryotic host cell.

As used herein the term "isolated" is meant to describe a polynucleotide, a polypeptide, or a cell that is in an environment different from that in which the polynucleotide, the polypeptide, or the cell naturally occurs. An isolated genetically modified host cell may be present in a mixed population of genetically modified host cells.

Expression cassettes may be prepared comprising a transcription initiation or transcriptional control region(s) (for example, a promoter), the coding region for the protein of interest, and a transcriptional termination region. Transcriptional control regions include those that provide for over-expression of the protein of interest in the genetically modified

host cell; those that provide for inducible expression, such that when an inducing agent is added to the culture medium, transcription of the coding region of the protein of interest is induced or increased to a higher level than prior to induction.

An expression cassette may contain at least one polynucleotide of interest to be co-transformed into the organism. Such an expression cassette is preferably provided with a plurality of restriction sites for insertion of the sequences of the invention to be under the transcriptional regulation of the regulatory regions. The expression cassette may additionally contain selectable marker genes.

The cassette may include 5' and 3' regulatory sequences operably linked to a polynucleotide of interest. By "operably linked" is intended, for example, a functional linkage between a promoter and a second sequence, wherein the promoter sequence initiates and mediates transcription of the DNA sequence corresponding to the second sequence. Generally, operably linked means that the nucleic acid sequences being linked are contiguous and, where necessary to join two protein coding regions, contiguous and in the same reading frame. When a polynucleotide comprises a plurality of coding regions that are operably linked such that they are under the control of a single promoter, the polynucleotide may be referred to as an "operon".

The expression cassette will optionally include in the 5'-3' direction of transcription, a transcriptional and translational initiation region, a polynucleotide sequence of interest and a transcriptional and translational termination region functional in plants. The transcriptional initiation region, the promoter, is optional, but may be native or analogous, or foreign or heterologous, to the intended host. Additionally, the promoter may be the natural sequence or alternatively a synthetic sequence. By "foreign" is intended that the transcriptional initiation region is not found in the native organism into which the transcriptional initiation region is introduced. As used herein, a chimeric gene comprises a coding sequence operably linked to a transcriptional initiation region that is heterologous to the coding sequence.

The termination region may be native with the transcriptional initiation region, may be native with the operably linked DNA sequence of interest, or may be derived from another source. Convenient termination regions are available from the Ti-plasmid of *A. tumefaciens*, such as the octopine synthase and nopaline synthase termination regions. See also Guerineau *et al.* (1991) *Mol. Gen. Genet.* 262:141-144; Proudfoot (1991) *Cell* 64:671-674; Sanfacon *et al.* (1991) *Genes Dev.* 5:141-149; Mogen *et al.* (1990) *Plant Cell* 2:1261-1272; Munroe *et al.*

(1990) Gene 91:151-158; Ballas *et al.* (1989) *Nucleic Acids Res.* 17:7891-7903; and Joshi *et al.* (1987) *Nucleic Acid Res.* 15:9627-9639.

Where appropriate, the proteins or polynucleotides of interest may be optimized for expression in the transformed organism. That is, the genes can be synthesized using plant or algae genomic preferred codons (for genomic transformation) or plastid-preferred codons corresponding to the plastids of the plant or algae of interest (for plastidic transformation). Methods are available in the art for synthesizing such codon optimized polynucleotides. See, for example, U. S. Patent Nos. 5,380,831 and 5,436,391, and Murray *et al.* (1989) *Nucleic Acids Res.* 17:477-498, herein incorporated by reference. Of course, the skilled artisan will appreciate that for the transplastomic purposes described herein, sequence optimization should be conducted with plastid codon usage frequency in mind, rather than the plant or algae genome codon usage exemplified in these references.

It is now well known in the art that when synthesizing a protein or polynucleotide of interest for improved expression in a host cell it is desirable to design the gene such that its frequency of codon usage approaches the frequency of codon usage of the host cell. It is also well known that plastome codon usage may vary from that of the host plant genome. For purposes of the subject invention, "frequency of preferred codon usage" is viewed in the context of whether the transformation is to be genomic or plastidic. For example, in the case of the latter, the phrase refers to the preference exhibited by a specific host cell plastid in usage of nucleotide codons to specify a given amino acid. To determine the frequency of usage of a particular codon in a gene, the number of occurrences of that codon in the gene is divided by the total number of occurrences of all codons specifying the same amino acid in the gene. Similarly, the frequency of preferred codon usage exhibited by a plastid can be calculated by averaging frequency of preferred codon usage in a number of genes expressed by the plastid. It usually is preferable that this analysis be limited to genes that are among those more highly expressed by the plastid or in the host cell's genome, as appropriate. Alternatively, the polynucleotide of interest may be synthesized to have a greater number of the host plastid's most preferred codon for each amino acid, or to reduce the number of codons that are rarely used by the host.

The expression cassettes may additionally contain 5' leader sequences in the expression cassette construct. Such leader sequences can act to enhance translation. Translation leaders are known in the art and include: picornavirus leaders, for example, EMCV leader (Encephalomyocarditis 5' noncoding region), Elroy-Stein *et al.* (1989) *PNAS*

USA 86:6126-6130; potyvirus leaders, for example, TEV leader (Tobacco Etch Virus), Allison *et al.* (1986); MDMV Leader (Maize Dwarf Mosaic Virus) *Virology* 154:9-20; and human immunoglobulin heavy-chain binding protein (BiP), Macejak *et al.* (1991) *Nature* 353:90-94; untranslated leader from the coat protein mRNA of alfalfa mosaic virus (AMV RNA 4), Jobling *et al.* (1987) *Nature* 325:622-625; tobacco mosaic virus leader (TMV), Gallie *et al.* (1989) in *Molecular Biology of RNA*, ed. Cech (Liss, New York), pp. 237-256; and maize chlorotic mottle virus leader (MCMV), Lommel *et al.* (1991) *Virology* 81:382-385. See also, Della-Cioppa *et al.* (1987) *Plant Physiol.* 84:965-968. Other methods known to enhance translation can also be utilized, for example, introns, and the like.

In preparing an expression cassette, the various proteins or polynucleotide may be manipulated, so as to provide for the polynucleotide sequences in the proper orientation and, as appropriate, in the proper reading frame. Toward this end, adapters or linkers may be employed to join the polynucleotide fragments or other manipulations may be involved to provide for convenient restriction sites, removal of superfluous nucleotides, removal of restriction sites, or the like. For this purpose, *in vitro* mutagenesis, primer repair, restriction, annealing, resubstitutions, e.g., transitions and transversions, may be involved.

Tissue-specific promoters are well known in the art and can be used to localize expression of the heterologous coding sequence in desired plant parts.

In addition, expressed gene products may be localized to specific organelles in the target cell by ligating DNA or RNA coded for peptide leader sequences to the polynucleotide of interest. Such leader sequences can be obtained from several genes of either plant or other sources. These genes encode cytoplasmically-synthesized proteins directed to, for example, mitochondria (the F1-ATPase beta subunit from yeast or tobacco, cytochrome c1 from yeast), chloroplasts (cytochrome oxidase subunit Va from yeast, small subunit of rubisco from pea), endoplasmic reticulum lumen (protein disulfide isomerase), vacuole (carboxypeptidase Y and proteinase A from yeast, phytohemagglutinin from French bean), peroxisomes (D-aminoacid oxidase, uricase) and lysosomes (hydrolases).

A nucleic acid is "hybridizable" to another nucleic acid, such as a cDNA, genomic DNA, or RNA, when a single stranded form of the nucleic acid can anneal to the other nucleic acid under the appropriate conditions of temperature and solution ionic strength.

Hybridization and washing conditions are well known and exemplified in Sambrook, J., Fritsch, E. F. and Maniatis, T. Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor (1989), particularly Chapter 11

and Table 11.1 therein; and Sambrook, J. and Russell, W., Molecular Cloning: A Laboratory Manual, Third Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor (2001).

As used herein, "stringent" conditions for hybridization refers to conditions wherein hybridization is carried out overnight at 20-25 °C below the melting temperature (Tm) of the DNA hybrid in 6x SSPE, 5x Denhardt's solution, 0.1% SDS, 0.1 mg/ml denatured DNA. The melting temperature, Tm, is described by the following formula (Beltz *et al.*, 1983):

Tm=81.5 C+16.6 Log[Na+]+0.41(%G+C)-0.61(% formamide)-600/length of duplex in base pairs.

Washes are typically carried out as follows:

(1) Twice at room temperature for 15 minutes in 1x SSPE, 0.1% SDS (low stringency wash).

(2) Once at Tm-20 °C for 15 minutes in 0.2x SSPE, 0.1% SDS.

Hybridization requires that the two nucleic acids contain complementary sequences, although depending on the stringency of the hybridization, mismatches between bases are possible. The appropriate stringency for hybridizing nucleic acids depends on the length of the nucleic acids and the degree of complementation, variables well known in the art. The greater the degree of similarity or homology between two nucleotide sequences, the greater the value of the melting temperature (Tm) for hybrids of nucleic acids having those sequences. The relative stability (corresponding to higher Tm) of nucleic acid hybridizations decreases in the following order: RNA:RNA, DNA:RNA, DNA:DNA. For hybrids of greater than 100 nucleotides in length, equations for calculating Tm have been derived (see Sambrook et al., supra, 9.50 9.51). For hybridizations with shorter nucleic acids, i.e., oligonucleotides, the position of mismatches becomes more important, and the length of the oligonucleotide determines its specificity (see Sambrook et al., supra, 11.7 11.8). Typically, the length for a hybridizable nucleic acid is at least about 10 nucleotides. Illustrative minimum lengths for a hybridizable nucleic acid are: at least about 15 nucleotides; at least about 20 nucleotides; and at least about 30 nucleotides. Furthermore, the skilled artisan will recognize that the temperature and wash solution salt concentration may be adjusted as necessary according to factors such as length of the probe.

The term "conservative amino acid substitution" refers to the interchangeability in proteins of amino acid residues having similar side chains. For example, a group of amino acids having aliphatic side chains consists of glycine, alanine, valine, leucine, and isoleucine; a group of amino acids having aliphatic-hydroxyl side chains consists of serine and threonine;

a group of amino acids having amide-containing side chains consists of asparagine and glutamine; a group of amino acids having aromatic side chains consists of phenylalanine, tyrosine, and tryptophan; a group of amino acids having basic side chains consists of lysine, arginine, and histidine; and a group of amino acids having sulfur-containing side chains consists of cysteine and methionine. Exemplary conservative amino acids substitution groups are: valine-leucine-isoleucine, phenylalanine-tyrosine, lysine-arginine, alanine-valine, and asparagine-glutamine. A protein or polypeptide associated with lipid metabolism containing conserved amino acid substitutions as compared to a protein or polypeptide associated with lipid metabolism exemplified herein would fall within the scope of "variants" of proteins or polypeptides associated with lipid metabolism.

"Synthetic nucleic acids" can be assembled from oligonucleotide building blocks that are chemically synthesized using procedures known to those skilled in the art. These building blocks are ligated and annealed to form gene segments which are then enzymatically assembled to construct the entire gene. "Chemically synthesized," as related to a sequence of DNA, means that the component nucleotides were assembled *in vitro*. Manual chemical synthesis of DNA may be accomplished using well-established procedures, or automated chemical synthesis can be performed using one of a number of commercially available machines. The nucleotide sequence of the nucleic acids can be modified for optimal expression based on optimization of nucleotide sequence to reflect the codon bias of the host cell. The skilled artisan appreciates the likelihood of successful expression if codon usage is biased towards those codons favored by the host. Determination of preferred codons can be based on a survey of genes derived from the host cell where sequence information is available. Fragments of full-length proteins can be produced by techniques well known in the art, such as by creating synthetic nucleic acids encoding the desired portions; or by use of Bal 31 exonuclease to generate fragments of a longer nucleic acid.

A polynucleotide or polypeptide has a certain percent "sequence identity" to another polynucleotide or polypeptide, meaning that, when aligned, that percentage of bases or amino acids are the same, and in the same relative position, when comparing the two sequences. Sequence similarity can be determined in a number of different manners. To determine sequence identity, sequences can be aligned using the methods and computer programs, including BLAST, available over the world wide web at ncbi.nlm.nih.gov/BLAST. See, e.g., Altschul et al. (1990), J. Mol. Biol. 215:403-410. Another alignment algorithm is FASTA, available in the Genetics Computing Group (GCG) package, from Madison, Wis., USA, a

wholly owned subsidiary of Oxford Molecular Group, Inc. Other techniques for alignment are described in Methods in Enzymology, vol. 266: Computer Methods for Macromolecular Sequence Analysis (1996), ed. Doolittle, Academic Press, Inc., a division of Harcourt Brace & Co., San Diego, Calif., USA. Of particular interest are alignment programs that permit gaps in the sequence. The Smith-Waterman is one type of algorithm that permits gaps in sequence alignments. See Meth. Mol. Biol. 70: 173-187 (1997). Also, the GAP program using the Needleman and Wunsch alignment method can be utilized to align sequences. See J. Mol. Biol. 48: 443-453 (1970).

As used herein, the term "variant" refers either to a naturally occurring genetic mutant of protein associated with lipid metabolism or a recombinantly prepared variation of protein associated with lipid metabolism, each of which contains one or more mutations in its DNA.

The term "variant" may also refer to either a naturally occurring variation of a given peptide or a recombinantly prepared variation of a given peptide or protein in which one or more amino acid residues have been modified by amino acid substitution, addition, or deletion. In certain embodiments, the variants include less than 75, less than 70, less than 60, less than 65, less than 60, less than 55, less than 50, less than 45, less than 40, less than 35, less than 30, less than 25, less than 20, less than 15, less than 10, less than 5, less than 4, less than 3, or less than 2 amino acid substitutions, rearrangements, insertions, and/or deletions relative to a naturally-occurring or native protein or polypeptide associated with lipid metabolism.

In some embodiments, the transformation vector further comprises a nucleic acid that confers resistance to a selection agent selected from bar, pat, ALS, HPH, HYG, EPSP, and Hm1.

Selectable marker genes include genes encoding antibiotic resistance, such as those encoding neomycin phosphotransferase II (NEO) and hygromycin phosphotransferase (HPT) as well as genes conferring resist insensitive to the herbicide or for an enzyme that degrades or detoxifies the herbicide in the plant before it can act. (See DeBlock et al. (1987) EMBO J, 6:2513-2518; DeBlock et al. (1989) Plant Physiol., 91:691-704; Fromm et al. (1990) 8:833-839. For example, resistance to glyphosate or sulfonylurea herbicides has been obtained by using genes coding for the mutant target enzymes, 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) and acetolactate synthase (ALS). Resistance to glufosinate ammonium, bromoxynil, and 2,4-dichlorophenoxyacetate (2,4-D) have been obtained by using bacterial

genes encoding phosphinothricin acetyltransferase, a nitrilase, or a 2,4-dichlorophenoxyacetate monooxygenase, which detoxify the respective herbicides.

For purposes of the present invention, selectable marker genes include, but are not limited to genes encoding: neomycin phosphotransferase II (Fraley et a. (1986) CRC Critical Reviews in Plant Science, 4:1-25); cyanamide hydratase (Maier-Greiner et al. (1991) Proc. Natl. Acad. Sci. USA, 88:4250-4264); aspartate kinase; dihydrodipicolinate synthase (Perl et al. (1993) Bio/Technology, 11:715-718); tryptophan decarboxylase (Goddijn et al. (1993) Plant Mol. Bio., 22:907-912); dihydrodipicolinate synthase and desensitized aspartade kinase (Perl et al. (1993) Bio/Technology, 11:715-718); bar gene (Toki et al. (1992) Plant Physiol., 100:1503-1507 and Meagher et al. (1996) and Crop Sci, 36:1367); tryptophane decarboxylase (Goddijn et al. (1993) Plant Mol. Biol., 22:907-912); neomycin phosphotransferase (NEO) (Southern et al. (1982) J. Mol. Appl. Gen., 1:327; hygromycin phosphotransferase (HPT or HYG) (Shimizu et al. (1986) Mol. Cell Biol., 6:1074); dihydrofolate reductase (DHFR) (Kwok et al. (1986) PNAS USA 4552); phosphinothricin acetyltransferase (DeBlock et al. (1987) EMBO J., 6:2513); 2,2-dichloropropionic acid dehalogenase (Buchanan-Wollatron et al. (1989) J. Cell. Biochem. 13D:330); acetohydroxyacid synthase (Anderson et al U.S. Pat. No. 4,761,373; Haughn et al. (1988) Mol. Gen. Genet. 221:266); 5-enolpyruvyl-shikimate-phosphate synthase (aroA) (Comai et al. (1985) Nature 317:741); haloarylnitrilase (Stalker et al., published PCT applon WO87/04181); acetyl-coenzyme A carboxylase (Parker et al. (1990) Plant Physiol. 92:1220); dihydropteroate synthase (sul I) (Guerineau et al. (1990) Plant Mol. Biol. 15:127); 32 kD photosystem II polypeptide (psbA) (Hirschberg et al. (1983) Science, 222:1346); etc.

Also included are genes encoding resistance to: chloramphenicol (Herrera-Estrella et al. (1983) EMBO J., 2:987-992); methotrexate (Herrera-Estrella et al. (1983) Nature, 303:209-213; Meijer et al. (1991) Plant Mol Bio., 16:807-820 (1991); hygromycin (Waldron et al. (1985) Plant Mol. Biol., 5:103-108; Zhijian et al. (1995) Plant Science, 108:219-227 and Meijer et al. (1991) Plant Mol. Bio. 16:807-820); streptomycin (Jones et al. (1987) Mol. Gen. Genet., 210:86-91); spectinomycin (Bretagne-Sagnard et al. (1996) Transgenic Res., 5:131-137); bleomycin (Hille et al. (1986) Plant Mol. Biol., 7:171-176); sulfonamide (Guerineau et al. (1990) Plant Mol. Bio., 15:127-136); bromoxynil (Stalker et al. (1988) Science, 242:419-423); 2,4-D (Streber et al. (1989) Bio/Technology, 7:811-816); glyphosate (Shaw et al. (1986) Science, 233:478-481); phosphinothricin (DeBlock et al. (1987) EMBO

J., 6:2513-2518); spectinomycin (Bretagne-Sagnard and Chupeau (1996) Transgenic Research 5:131-137).

The bar gene confers herbicide resistance to glufosinate-type herbicides, such as phosphinothricin (PPT) or bialaphos, and the like. As noted above, other selectable markers that could be used in the vector constructs include, but are not limited to, the pat gene, also for bialaphos and phosphinothricin resistance, the ALS gene for imidazolinone resistance, the HPH or HYG gene for hygromycin resistance, the EPSP synthase gene for glyphosate resistance, the Hm1 gene for resistance to the Hc-toxin, and other selective agents used routinely and known to one of ordinary skill in the art.

Screening Methods for Obtaining Plants with Elevated Lipid Content

In some embodiments, the invention provides methods for screening for a functional protein or polypeptide associated with lipid metabolism for elevating lipid content and/or inducing lipid droplet accumulation in a plant, bacterial, or algal cell, wherein the method comprises:

obtaining a test plant, bacterial, or algal cell genetically-modified to express a candidate exogenous protein or polypeptide associated with lipid metabolism; and

growing the genetically-modified test cell and selecting the genetically-modified test cell having elevated lipid content and/or increased lipid droplet level when compared to a native (wild-type) cell of the same type.

Embodiments of this invention also pertain to methods for screening for a functional protein or polypeptide associated with lipid metabolism for elevating lipid content and/or inducing lipid droplet accumulation in a plant, bacterial, or algal cell, wherein the method comprises:

transforming a test plant, bacterial, or algal cell with a vector nucleic acid sequence encoding a candidate exogenous protein or polypeptide associated with lipid metabolism, wherein the nucleic acid is operably linked to a promoter and/or a regulatory sequence; and

growing the genetically-modified test cell and selecting the genetically-modified test cell having elevated lipid content and/or increased lipid droplet level when compared to a native (wild-type) cell of the same type.

In certain embodiments of the screening method, the transformed or genetically-modified test cell is a plant cell. In certain embodiments, the plant test cell is in a plant tissue, plant part, or whole plant.

In certain embodiments of the screening method, vegetative plant (non-seed) cells, tissues or plant parts including, but not limited to, leaves, roots, stems, shoots, buds, tubers, fruits, and flowers, are genetically-modified or transformed. In another embodiment of the screening method, a plant seed cell or tissue is genetically-modified or transformed.

In some embodiments, a method may employ marker-assisted breeding to identify plants, including cultivars or breeding lines, displaying a trait of interest, such as elevated levels of neutral lipids in vegetative portions of plant biomass.

When an exogenous nucleic acid comprising a nucleotide sequence that encodes a protein or polypeptide associated with lipid metabolism is introduced into the host cell, lipid content of the test cell is elevated. In certain embodiments, a candidate protein or polypeptide associated with lipid metabolism is selected if there is an elevation of the lipid content of the cell of at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, or at least about 90%, or more, as compared to a non-genetically-modified host.

In some embodiments, for example, where the exogenous nucleic acid is a plurality of exogenous nucleic acids (such as, for example, a cDNA library, a genomic library, or a population of nucleic acids, each encoding a protein or polypeptide associated with lipid metabolism with a different amino acid sequence, etc.), the exogenous nucleic acids are introduced into a plurality of host cells, forming a plurality of test cells. In certain embodiments, the test cells are in some embodiments grown in normal culture conditions.

Methods of isolating the exogenous nucleic acid from a test cell are well known in the art. Suitable methods include, but are not limited to, any of a number of alkaline lysis methods that are standard in the art.

In some embodiments, a subject screening method will further comprise further characterizing a candidate gene product. In these embodiments, the exogenous nucleic acid comprising nucleotide sequence(s) encoding protein or polypeptide associated with lipid metabolism are isolated from a test cell; the gene product(s) are expressed in a cell and/or in an *in vitro* cell-free transcription/translation system. In some embodiments, the exogenous nucleic acid is subjected to nucleotide sequence analysis, and the amino acid sequence of the gene product deduced from the nucleotide sequence. In some embodiments, the amino acid sequence of the gene product is compared with other amino acid sequences in a public database of amino acid sequences, to determine whether any significant amino acid sequence identity to an amino acid sequence of a known protein exists. In addition, the gene product(s)

are expressed in a cell and/or in an *in vitro* cell-free transcription/translation system; and the effect of the gene product(s) on a metabolic pathway intermediate or other metabolite is analyzed.

Exogenous nucleic acids that are suitable for introducing into a host cell, to produce a test cell, include, but are not limited to, naturally-occurring nucleic acids isolated from a cell; naturally-occurring nucleic acids that have been modified (for example, by mutation) before or subsequent to isolation from a cell; synthetic nucleic acids, e.g., nucleic acids synthesized in a laboratory using standard methods of chemical synthesis of nucleic acids, or generated by recombinant methods; synthetic or naturally-occurring nucleic acids that have been amplified *in vitro*, either within a cell or in a cell-free system; and the like.

Exogenous nucleic acids that are suitable for introducing into a host cell include, but are not limited to, genomic DNA; RNA; a complementary DNA (cDNA) copy of mRNA isolated from a cell; recombinant DNA; and DNA synthesized *in vitro*, e.g., using standard cell-free *in vitro* methods for DNA synthesis. In some embodiments, exogenous nucleic acids are a cDNA library made from cells, either prokaryotic cells or eukaryotic cells. In some embodiments, exogenous nucleic acids are a genomic DNA library made from cells, either prokaryotic cells or eukaryotic cells.

Nucleic acids will in some embodiments be mutated before being introduced into a host cell. Methods of mutating a nucleic acid are well known in the art and include well-established chemical mutation methods, radiation-induced mutagenesis, and methods of mutating a nucleic acid during synthesis. Chemical methods of mutating DNA include exposure of DNA to a chemical mutagen, e.g., ethyl methanesulfonate (EMS), methyl methanesulfonate (MMS), N-nitrosourea (ENU), N-methyl-N-nitro-N'-nitrosoguanidine, 4-nitroquinoline N-oxide, diethylsulfate, benzopyrene, cyclophosphamide, bleomycin, triethylmelamine, acrylamide monomer, nitrogen mustard, vincristine, diepoxyalkanes (for example, diepoxybutane), ICR-170, formaldehyde, procarbazine hydrochloride, ethylene oxide, dimethylnitrosamine, 7,12 dimethylbenz(a)anthracene, chlorambucil, hexamethylphosphoramide, bisulfan, and the like. Radiation mutation-inducing agents include ultraviolet radiation, .gamma.-irradiation, X-rays, and fast neutron bombardment. Mutations can also be introduced into a nucleic acid using, e.g., trimethylpsoralen with ultraviolet light. Random or targeted insertion of a mobile DNA element, e.g., a transposable element, is another suitable method for generating mutations. Mutations can be introduced into a nucleic acid during amplification in a cell-free *in vitro* system, e.g., using a polymerase

chain reaction (PCR) technique such as error-prone PCR. Mutations can be introduced into a nucleic acid *in vitro* using DNA shuffling techniques (e.g., exon shuffling, domain swapping, and the like). Mutations can also be introduced into a nucleic acid as a result of a deficiency in a DNA repair enzyme in a cell, e.g., the presence in a cell of a mutant gene encoding a mutant DNA repair enzyme is expected to generate a high frequency of mutations (i.e., about 1 mutation/100 genes-1 mutation/10,000 genes) in the genome of the cell. Examples of genes encoding DNA repair enzymes include but are not limited to Mut H, Mut S, Mut L, and Mut U, and the homologs thereof in other species (e.g., MSH 1 6, PMS 1 2, MLH 1, GTBP, ERCC-1, and the like). Methods of mutating nucleic acids are well known in the art, and any known method is suitable for use. See, e.g., Stemple (2004) Nature 5:1-7; Chiang et al. (1993) PCR Methods Appl 2(3): 210-217; Stemmer (1994) Proc. Natl. Acad. Sci. USA 91:10747-10751; and U.S. Pat. Nos. 6,033,861, and 6,773,900.

Isolation of Homologs

Isolation of additional homologs from other plant species may be accomplished by laboratory procedures well known and commonly used in the art. Standard techniques are used for identification, cloning, isolation, amplification, and purification of nucleic acid sequences and polypeptides. These techniques and various others are generally performed as described for instance in Sambrook et al., 1989. Genome walking techniques may be performed according to manufacturer's specifications (CLONTECH Laboratories, Inc., Palo Alto, Calif.).

One such technique for isolation of homologs is the use of oligonucleotide probes based on sequences disclosed in this specification to identify the desired gene in a cDNA or genomic DNA library. To construct genomic libraries, large segments of genomic DNA are generated by digestion with restriction endonucleases and then ligating the resultant segments with vector DNA to form concatemers that can be packaged into an appropriate vector. To prepare a cDNA library, mRNA is isolated from the desired organ, such as seed tissue, and a cDNA library is prepared from the mRNA.

A cDNA or genomic DNA library can be screened using a probe based upon the sequence of a cloned naturally-occurring protein or polypeptide sequence. Probes may be used to hybridize with genomic DNA or cDNA sequences to isolate homologous genes in the same or different plant species. Usefully employed such probes include, without limitation, 5' UTRs which, may function as promoters. Alternatively, antibodies raised against a

polypeptide, or homolog thereof, can be used to screen an mRNA expression library to isolate sequences of interest. Homologs may also be identified in silico, for instance by similarity-based database searches as described below.

Nucleic acid sequences can be screened for the presence of a protein encoding sequence that is homologous to genes of other organisms with known protein encoding sequence using any of a variety of search algorithms. Such search algorithms can be homology-based or predictive-based. Similarity-based searches (e.g., GAP2, BLASTX supplemented by NAP and TBLASTX) can detect conserved sequences during comparison of DNA sequences or hypothetically translated protein sequences to public and/or proprietary DNA and protein databases.

Existence of a gene is inferred if significant sequence similarity extends over the majority of the target gene. Since such methods may overlook genes unique to the source organism, for which homologous nucleic acid molecules have not yet been identified in databases, gene prediction programs may also be used. Gene prediction programs generally use "signals" in the sequences, such as splice sites or "content" statistics, such as codon bias, to predict gene structures (Stormo, 2000).

Alternatively, the nucleic acids of interest can be amplified from nucleic acid samples using amplification techniques. For example, polymerase chain reaction technology can be used to amplify the sequences of a gene of interest or the homolog gene directly from genomic DNA, from cDNA, from genomic libraries, and cDNA libraries. PCR and other in vitro amplification methods may also be useful, for example, in cloning nucleic acids sequences that code for proteins to be expressed, to make nucleic acids to use as probes for detecting the presence of desired mRNA in samples, for nucleic acid sequencing, or for other purposes.

Appropriate primers and probes for identifying homolog sequences from plant tissues are generated from comparisons of the sequences provided herein. For a general overview of PCR, see, Innis, et al., eds., 1990.

PCR or other primers may be used under standard PCR conditions, preferably using nucleic acid sequences as identified in EST libraries or other GenBank accessions as a template. The PCR products generated by any of the reactions can then be used to identify nucleic acids useful in the context of the present invention by their ability to hybridize to known homolog genes found in GenBank and other databases.

Plant Transformation

To use isolated sequences in the above techniques, recombinant DNA vectors suitable for transformation of plant cells are prepared. Techniques for transforming a wide variety of higher plant species are well known and described in the technical and scientific literature. See, for example, Weising et al., 1988; and Sambrook et al., 1989. Methods of plant cell culture are well known in the art. A DNA sequence coding for the desired polypeptide, for example a cDNA sequence encoding a full length protein, will preferably be combined with transcriptional and translational initiation regulatory sequences that will direct the transcription of the sequence from the gene in the intended tissues of the transformed plant.

Vectors used for plant transformation may include, for example, plasmids, cosmids, yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), plant artificial chromosomes (PACs), or any suitable cloning system. It is contemplated the utilization of cloning systems with large insert capacities will allow introduction of large DNA sequences comprising more than one selected gene. Introduction of such sequences may be facilitated by use of BACs or YACs, or even PACs. For example the use of BACs for Agrobacterium-mediated transformation was disclosed by Hamilton et al., 1999.

Particularly useful for transformation are expression cassettes that have been isolated from such vectors. DNA segments used for transforming plant cells will, of course, generally comprise the cDNA, gene or genes that one desires to introduce into and have expressed in the host cells. These DNA segments can further include structures such as promoter, enhancers, 3' untranslated regions (such as polyadenylation sites), polylinkers, or even regulatory genes as desired. The DNA segment or gene chosen for cellular introduction may encode a protein that will be expressed in the resultant recombinant cells resulting in a screenable or selectable trait and/or will impart an improved phenotype to the resulting transgenic plant. However, this may not always be the case, and the present invention also encompasses transgenic plants incorporating non-expressed transgenes.

A number of promoters that are active in plant cells have been described in the literature, and are preferred elements included in the context of the present invention. Such promoters would include but are not limited to those isolated from the following genes: nopaline synthase (NOS; Ebert et al., 1987) and octopine synthase (OCS): cauliflower mosaic virus (CaMV) 19S (Lawton et al. 1987) and 35S (Odell et al., 1985), as well as the enhanced CaMV 35S promoter (e35S; described by Kay et al., 1987); figwort mosaic virus (FMV) 35S; the small subunit of ribulose bisphosphate carboxylase (ssRUBISCO, a very abundant plant

polypeptide); napin (Kridl et al., 1991); Adh (Walker et al., 1987); sucrose synthase (Yang et al., 1990); tubulin; actin (Wang et al., 1992); cab (Sullivan et al., 1989); PEPCase (Hudspeth et al., 1989); 7S-alpha'-conglycinin (Beachy et al., 1985); R gene complex promoters (Chandler et al. 1989); tomato E8; patatin; ubiquitin; mannopine synthase (mas); soybean seed protein glycinin (Gly); soybean vegetative storage protein (vsp); waxy; Brittle; Shrunken 2; Branching enzymes I and II; starch synthases; debranching enzymes; oleosins; glutelins; globulin 1; BETL1; and Arabidopsis banyuls promoter. The rice actin 1 promoter, the AGL11 promoter, the BETL1 promoter, and the e35S promoter may find use in the practice of the present invention. All of these promoters have been used to create various types of DNA constructs that have been expressed in plants (see, for example, Rogers et al., WO 84/02913).

Promoter hybrids can also be constructed to enhance transcriptional activity (Hoffman, U.S. Pat. No. 5,106,739, herein incorporated by reference), or to combine desired transcriptional activity, inducibility, and tissue or developmental specificity. Promoters that function in plants include but are not limited to promoters that are classified as, among others, inducible, viral, synthetic, constitutive, tissue-specific, developmentally-regulated, chemically or environmentally inducible, or senescence-related, for instance as described (Odell et al., 1985). Promoters that are tissue specific, tissue-enhanced, or developmentally regulated are also known in the art and envisioned to have utility in the practice of this present invention. For instance, a tissue specific promoter, such as the ST-LS1 promoter (e.g. Stockhaus et al., 1989), that is functional in plant vegetative tissues such as leaves, stems, and/or roots, may be of use. Such a promoter may also be expressed to at least some degree in seed or embryo tissues. In certain embodiments, the promoter to be utilized may be expressed preferentially in green parts of a plant such as leaves or stems. A senescence-related promoter (e.g. from SAG12) may also be utilized.

The promoters used in the present invention may be modified to affect their control characteristic. Promoters can be derived by means of ligation with operator regions, random or controlled mutagenesis, or other means well known in the art. Furthermore the promoter regions can be altered to contain multiple enhancer sequences to assist in elevating gene expression. Examples of such enhancer sequences have been reported (Kay et al., 1987).

Where an enhancer is used in conjunction with a promoter for the expression of a selected protein, it is believed that it will be preferred to place the enhancer between the promoter and the start codon of the selected coding region. However, one could also use a

different arrangement of the enhancer relative to other sequences and still realize the beneficial properties conferred by the enhancer. For example, the enhancer could be placed 5' of the promoter region, within the promoter region, within the coding sequence, or 3' of the coding region. The placement and choice of sequences used as enhancers is known to those of skill in the art in light of the present disclosure. Transformation constructs prepared in accordance with the current invention will typically include a 3' untranslated region (3' UTR), and typically contains a polyadenylation sequence. One type of 3' UTR that may be used is a 3' UTR from the nopaline synthase gene of Agrobacterium tumefaciens (NOS 3'-end; Bevan et al., 1983). Other 3' UTR sequences can be used and are commonly known to those of skill in the art.

A number of selectable marker genes are known in the art and can be used in the present invention (Wilmink and Dons, 1993). By employing a selectable or screenable marker gene in addition to the gene of interest, one can provide or enhance the ability to identify transformants. Useful selectable marker genes for use in the present invention would include genes that confer resistance to compounds such as antibiotics like kanamycin and herbicides like glyphosate or dicamba. Other selectable markers known in the art may also be used and would fall within the scope of the present invention.

DNA constructs of the present invention may be introduced into the genome of the desired plant host by a variety of techniques that are well known in the art. For example, the DNA construct may be introduced directly into the genomic DNA of the plant cell using techniques such as electroporation and microinjection of plant cell protoplasts, or the DNA constructs can be introduced directly to plant tissue using DNA particle bombardment.

Microinjection techniques are known in the art and well described in the scientific and patent literature. The introduction of DNA constructs using polyethylene glycol precipitation is described in Paszkowski et al., 1984. Electroporation techniques are described in Fromm et al., 1985. Ballistic transformation techniques are described in Klein et al., 1987.

Alternatively, the DNA constructs may be combined with suitable T-DNA flanking regions and introduced into a conventional Agrobacterium tumefaciens host vector. The virulence functions of the Agrobacterium tumefaciens host direct the insertion of the construct and adjacent marker into the plant cell DNA when the cell is infected by the bacteria. Agrobacterium tumefaciens-mediated transformation techniques, including disarming and use of binary vectors, are well described in the scientific literature. See, for example, Horsch, 1984; and Fraley, 1983.

After transformation by any of the above transformation techniques, the transformed plant cells or tissues may be grown in an appropriate medium to promote cell proliferation and regeneration. Plant regeneration from cultured protoplasts is described in Evans et al., 1983; and Binding, Regeneration of Plants, Plant Protoplasts, pp. 21 73, CRC Press, Boca Raton, 1985. For gene gun transformation of wheat and maize, see, U.S. Pat. Nos. 6,153,812 and 6,160,208. See also, Christou, 1996. See, also, U.S. Pat. Nos. 5,416,011; 5,463,174; and 5,959,179 for Agrobacterium-mediated transformation of soy; U.S. Pat. Nos. 5,591,616 and 5,731,179 for Agrobacterium-mediated transformation of monocots such as maize; and U.S. Pat. No. 6,037,527 for Agrobacterium-mediated transformation of cotton. Other Rhizobiaceae may be used for plant cell transformation as well (e.g. Broothaerts et al., 2007).

To generate a subject genetically modified host cell according to the subject invention, one or more nucleic acids comprising nucleotide sequences encoding one or more proteins or polypeptides associated with lipid metabolism can be introduced stably or transiently into a parent host cell, using established techniques, including, but not limited to, electroporation, calcium phosphate precipitation, DEAE-dextran mediated transfection, liposome-mediated transfection, particle bombardment, Agrobacterium-mediated transformation, and the like. For stable transformation, a nucleic acid will generally further include a selectable marker, for example, any of several well-known selectable markers such as neomycin resistance, ampicillin resistance, tetracycline resistance, chloramphenicol resistance, kanamycin resistance, and the like.

Where a parent host cell has been genetically modified to produce two or more proteins or polypeptides associated with lipid metabolism, nucleotide sequences encoding the two or more proteins or polypeptides associated with lipid metabolism will in some embodiments each be contained on separate expression vectors. Where the host cell is genetically modified to express one or more proteins or polypeptides associated with lipid metabolism, nucleotide sequences encoding the one or more proteins or polypeptides associated with lipid metabolism will in some embodiments be contained in a single expression vector. Where nucleotide sequences encoding the one or more proteins or polypeptides associated with lipid metabolism are contained in a single expression vector, in some embodiments, the nucleotide sequences will be operably linked to a common control element (for example, a promoter), such that the common control element controls expression of all of the nucleotide sequences on the single expression vector.

Where nucleotide sequences encoding proteins or polypeptides associated with lipid metabolism are contained in a single expression vector, in some embodiments, the nucleotide sequences will be operably linked to different control elements (for example, a promoter), such that, the different control elements control expression of each of the nucleotide sequences separately on a single expression vector.

In many embodiments, the exogenous nucleic acid is inserted into an expression vector. Expression vectors that are suitable for use in prokaryotic and eukaryotic host cells are known in the art, and any suitable expression vector can be used. Suitable expression vectors are as described above.

As noted above, an exogenous nucleic acid will in some embodiments be isolated from a cell or an organism in its natural environment. In some embodiments, the nucleic acid of the cell or organism will be mutated before nucleic acid is isolated from the cell or organism. In other embodiments, the exogenous nucleic acid is synthesized in a cell-free system *in vitro*.

In some embodiments, the exogenous nucleic acid is a synthetic nucleic acid. In some embodiments, a synthetic nucleic acid comprises a nucleotide sequence encoding a variant protein or polypeptide associated with lipid metabolism, for example, a variant protein or polypeptide associated with lipid metabolism that differs in amino acid sequence by one or more amino acids from a naturally-occurring protein or polypeptide associated with lipid metabolism. In some embodiments, a variant protein or polypeptide associated with lipid metabolism differs in amino acid sequence by from about 10 amino acids to about 15 amino acids, from about 15 amino acids to about 20 amino acids, from about 20 amino acids to about 25 amino acids, from about 25 amino acids to about 30 amino acids, from about 30 amino acids to about 35 amino acids, from about 35 amino acids to about 40 amino acids, from about 40 amino acids to about 50 amino acids, or from about 50 amino acids to about 60 amino acids, compared to the amino acid sequence of a naturally-occurring parent protein or polypeptide associated with lipid metabolism.

In some embodiments, a nucleic acid comprising a nucleotide sequence encoding a naturally-occurring protein or polypeptide associated with lipid metabolism is mutated, using any of a variety of well-established methods, giving rise to a nucleic acid comprising a nucleotide sequence encoding a variant protein or polypeptide associated with lipid metabolism.

Suitable mutagenesis methods include, but are not limited to, chemical mutation methods, radiation-induced mutagenesis, and methods of mutating a nucleic acid during synthesis, as described above. Thus, for example, a nucleic acid comprising a nucleotide sequence encoding a naturally-occurring protein or polypeptide associated with lipid metabolism is exposed to a chemical mutagen, as described above, or subjected to radiation mutation, or subjected to an error-prone PCR, and the mutagenized nucleic acid introduced into a genetically modified host cell(s) as described above. Methods for random mutagenesis using a "mutator" strain of bacteria are also well known in the art and can be used to generate a variant. See, e.g., Greener et al., "An Efficient Random Mutagenesis Technique Using an *E. coli* Mutator Strain", Methods in Molecular Biology, 57:375-385 (1995). Saturation mutagenesis techniques employing a polymerase chain reaction (PCR) are also well known and can be used. See, e.g., U.S. Pat. No. 6,171,820.

An embodiment of the invention provides a host cell comprising a vector according to the invention. Other embodiments include plant plastid transformation vectors or nuclear transformation vectors containing nucleotide sequences encoding proteins or polypeptides associated with lipid metabolism, such as containing the full-length protein or polypeptide associated with lipid metabolism, or variants or fragments thereof, for the expression of the protein or polypeptide associated with lipid metabolism with elevated lipid content in the plant cell. These plant vectors may contain other sequences for the generation of chimeric proteins or polypeptides associated with lipid metabolism which may contain mutations, deletions, or insertions of nucleic acid sequences.

According to embodiments of the present invention, a wide variety of plants and plant cell systems can be engineered for the desired physiological and agronomic characteristics described herein using the nucleic acid constructs of the present invention by various transformation methods known in the art, including Agrobacterium-mediated transformation (Horsch et al., Science 227: 1227-1231, 1985) or plastid transformation (Staub and Maliga, Plant J. 6: 547-553, 1994; Hahn and Kuehnle, 2003, cited herein above).

In preferred embodiments, target plants and plant cells for engineering include, but are not limited to, those monocotyledonous and dicotyledonous plants, such as crops, including grain crops (for example, wheat, maize, rice, millet, barley), tobacco, fruit crops (for example, tomato, strawberry, orange, grapefruit, banana), forage crops (for example, alfalfa), root vegetable crops (for example, carrot, potato, sugar beets, yam), leafy vegetable crops (for example, lettuce, spinach); flowering plants (for example, petunia, rose,

chrysanthemum), conifers and pine trees (for example, pine, fir, spruce); oil crops (for example, sunflower, rape seed); and plants used for experimental purposes (for example, *Arabidopsis*).

According to other embodiments of the present invention, desired plants may be obtained by engineering one or more of the vectors expressing proteins or polypeptides associated with lipid metabolism as described herein into a variety of plant cell types, including but not limited to, protoplasts, tissue culture cells, tissue and organ explants, pollens, embryos, as well as whole plants. In an embodiment of the present invention, the engineered plant material is selected or screened for transformants (those that have incorporated or integrated the introduced gene construct(s)) following the approaches and methods described below. An isolated transformant may then be regenerated into a plant and progeny thereof (including the immediate and subsequent generations) *via* sexual or asexual reproduction or growth. Alternatively, the engineered plant material may be regenerated into a plant before subjecting the derived plant to selection or screening for the marker gene traits. Procedures for regenerating plants from plant cells, tissues or organs, either before or after selecting or screening for marker gene(s), are well known to those skilled in the art.

According to another embodiment of the present invention, tissue-specific promoters may be used to target the expression of proteins or polypeptides associated with lipid metabolism in fruits, roots or leaves so that an edible plant part is provided low-temperature tolerance. Examples of tissue-specific promoters include those encoding *rbsC* (Coruzzi et al., EMBO J. 3:1671-1697, 1984) for leaf-specific expression and SAHH or SHMT (Sivanandan et al., Biochimica et Biophysica Acta 1731:202-208, 2005) for root-specific expression. Another exemplary root-specific promoter is taught by Ekramoddoullah et al., US Patent No. 7,285,656 B2. Also, the Cauliflower Mosaic Virus (CaMV) *35S* promoter has been reported to have root-specific and leaf-specific modules in its promoter region (Benfey et al., EMBO J. 8:2195-2202, 1989). Other tissue-specific promoters are well known and widely available to those of ordinary skill in the art. Further, a wide variety of constitutive or inducible promoters are also well known and widely available to those of ordinary skill in the art.

Proplastid and chloroplast genetic engineering have been shown to varying degrees of homoplasmy for several major agronomic crops including potato, rice, maize, soybean, grape, sweet potato, and tobacco including starting from non-green tissues. Non-lethal selection on antibiotics is used to proliferate cells containing plastids with antibiotic resistance genes. Plastid transformation methods use two plastid-DNA flanking sequences that recombine with

plastid sequences to insert chimeric DNA into the spacer regions between functional genes of the plastome, as is established in the field (see Bock and Hagemann, Prog. Bot. 61:76-90, 2000, and Guda *et al.*, Plant Cell Reports 19:257-262, 2000, and references therein).

Antibiotics such as spectinomycin, streptomycin, and kanamycin can shut down gene expression in chloroplasts by ribosome inactivation. These antibiotics bleach leaves and form white callus when tissue is put onto regeneration medium in their presence. The bacterial genes *aad*A and *neo* encode the enzymes aminoglycoside-3N-adenyltransferase and neomycin phosphotransferase, which inactivate these antibiotics, and can be used for positive selection of plastids engineered to express these genes. Polynucleotides of interest can be linked to the selectable genes and thus can be enriched by selection during the sorting out of engineered and non-engineered plastids. Consequently, cells with plastids engineered to contain genes for these enzymes (and linkages thereto) can overcome the effects of inhibitors in the plant cell culture medium and can proliferate, while cells lacking engineered plastids cannot proliferate. Similarly, plastids engineered with polynucleotides encoding enzymes from the mevalonate pathway to produce IPP from acetyl CoA in the presence of inhibitors of the non-mevalonate pathway can overcome otherwise inhibitory culture conditions. By utilizing the polynucleotides disclosed herein in accord with this invention, an inhibitor targeting the non-mevalonate pathway and its components can be used for selection purposes of transplastomic plants produced through currently available methods, or any future methods which become known for production of transplastomic plants, to contain and express said polynucleotides and any linked coding sequences of interest.

This selection process of the subject invention is unique in that it is the first selectable trait that acts by pathway complementation to overcome inhibitors. This is distinguished from the state of the art of selection by other antibiotics to which resistance is conferred by inactivation of the antibiotic itself, *e.g.* compound inactivation as for the aminoglyoside 3'-adenyltransferase gene or neo gene. This method avoids the occurrence of resistant escapes due to random insertion of the resistance gene into the nuclear genome or by spontaneous mutation of the ribosomal target of the antibiotic, as is known to occur in the state of the art. Moreover, this method requires the presence of an entire functioning mevalonate pathway in plastids. For example, if one of the enzyme activities of the mevalonate pathway is not present in the plastid, resistance will not be conferred.

A transformed plant cell, callus, tissue, or plant may be identified and isolated by selecting or screening the engineered plant material for traits encoded by the marker genes present on the transforming DNA. For instance, selection may be performed by growing the engineered plant material on media containing inhibitory amount of the antibiotic or herbicide to which the transforming gene construct confers resistance. Further, transformed plants and plant cells may also be identified by screening for the activities of any visible marker genes (*e.g.*, the β-glucuronidase, luciferase, B or C1 genes) that may be present on the vector of the present invention. Such selection and screening methodologies are well known to those skilled in the art. Alternatively or in addition, screening may be for improved low-temperature tolerance as taught herein, for example, by observing a reduction in growth-inhibition.

Physical and biochemical methods may also be used to identify plant or plant cell transformants containing the gene constructs of the present invention. These methods include but are not limited to: 1) Southern analysis or PCR amplification for detecting and determining the structure of the recombinant DNA insert; 2) Northern blot, S1 RNase protection, primer-extension or reverse transcriptase-PCR amplification for detecting and examining RNA transcripts of the gene constructs; 3) enzymatic assays for detecting enzyme activity, where such gene products are encoded by the gene construct; 4) protein gel electrophoresis (PAGE), Western blot techniques, immunoprecipitation, or enzyme-linked immunoassays, where the gene construct products are proteins. Additional techniques, such as *in situ* hybridization, enzyme staining, and immunostaining, also may be used to detect the presence or expression of the recombinant construct in specific plant organs and tissues. The methods for doing all these assays are well known to those skilled in the art. In a specific embodiment, the selectable marker gene *nptII*, which specifies kanamycin-resistance, is used in nuclear transformation.

Following transformation, a plant may be regenerated, *e.g.*, from single cells, callus tissue, or leaf discs, as is standard in the art. Almost any plant can be entirely regenerated from cells, tissues, and organs of the plant. Available techniques are reviewed in Vasil *et al.* (1984) in *Cell Culture and Somatic Cell Genetics of Plants, Vols. I, II, and III, Laboratory Procedures and Their Applications* (Academic press); and Weissbach *et al.* (1989) *Methods for Plant Mol. Biol.*

The transformed plants may then be grown, and either pollinated with the same transformed strain or different strains, and the resulting hybrid having expression of the

desired phenotypic characteristic identified. Two or more generations may be grown to ensure that expression of the desired phenotypic characteristic is stably maintained and inherited, and then seeds harvested to ensure expression of the desired phenotypic characteristic has been achieved.

The particular choice of a transformation technology will be determined by its efficiency to transform certain target species, as well as the experience and preference of the person practicing the invention with a particular methodology of choice. It will be apparent to the skilled person that the particular choice of a transformation system to introduce nucleic acid into plant plastids is not essential to or a limitation of the invention, nor is the choice of technique for plant regeneration.

Applications

In certain embodiments, the present invention can be used to:

a) provide higher efficiency and cost effective energy production;
b) increase production of lipids which are beneficial for human health, *e.g.*, omega-unsaturated fat in olives, canola, corns, peanuts, sunflower seeds, etc;
c) generate plants for protein therapy. Some proteins play a positive regulatory role in improving the metabolic health in humans suffering from insulin resistance, type 2 diabetes, cardiovascular diseases etc.;
d) produce genetically-modified plants with elevated lipid content for feeding animals including livestock such as cows to produce milk with high level of lipid droplets;
e) produce genetically-modified algal cells with elevated lipid content for production of biofuels, and feed; and
f) produce genetically-modified bacterial cells expressing proteins associated with lipid metabolism for cleaning oil spillage.

Increase the production of oils which are beneficial for human health. Our biochemical analysis shows that FSP27 expression in plants increase omega-6 and omega-3 unsaturated fatty acids.

Expressing fish homologs of FSP27 in combination with other nucleic acid molecules encoding proteins involved in the synthesis of long-chain polyunsaturated fatty acids in plants can be used to increase oil contents in plants, thereby producing plants with high omega-unsaturated fatty acid contents. In one embodiment, the transgenic plants of the present invention can serve as an inexpensive and safe source of dietary fatty acids.

Transgenic plants with high fat contents can be used to feed milk-producing cows, thereby increasing fat contents in dairy products.

The present invention can be used to increase oil contents in oil-producing plants including, but not limited to, olive, canola, sunflower, soybean, castor, and oleaginous fruits such as palm and avocado. The present invention can also be used to increase unsaturated oil contents in plants, to improve the quality and quantity of oil in plants, and to increase oil content in seeds.

The seeds of the transgenic plants with high lipid contents can be used to produce biodegradable plastic (also called as "bioplastic").

The proteins or polypeptides associated with lipid metabolism (such as FSP27) can be expressed in algae to increase biofuel production.

Common uses for oils comprising neutral lipids include the preparation of food for human consumption, feed for non-human animal consumption and industrial uses such as for preparation of biofuels.

As used herein, "industrial use" or "industrial usage" refers to non-food and non-feed uses for products prepared from plant parts prepared according to the present invention. As used herein, "biofuel" refers to a fuel combusted to provide power, heat, or energy, e.g. for an internal combustion engine, comprising at least 1%, 5%, 10%, 20% or more, by weight, of an oil, or product thereof, produced from a plant of the present invention, or by a method of the present invention.

Also included in this invention are plants, plant cell cultures, and plant parts thereof, oil obtained from the vegetative tissues of such plants and cells and progeny thereof, animal feed derived from the processing of such tissues, the use of the foregoing oil in food, animal feed, biofuels, cooking oil or industrial applications, and products made from the hydrogenation, fractionation, interesterification or hydrolysis of such oil.

MATERIALS AND METHODS

Expression of *A. thaliana SEIPIN* Genes in Yeast Cells

The coding regions of three *Arabidopsis SEIPIN* genes, designated *AtSEIPIN1*, *AtSEIPIN2*, and *AtSEIPIN3*, were isolated from wild type *Arabidopsis* (Columbia-0 [Col-0]) by using reverse transcriptase (RT)-PCR. RNA was purified by RNeasy Plant Mini kit (Qiagen) and treated by DNase (Promega) to avoid any DNA contamination. About 100 ng total RNA from each sample was used for RT-PCR. The RT-PCR was performed by using

SuperScript® One-Step RT-PCR System (Invitrogen). The RT-PCR program was set up as follows, reverse transcription at 42°C for 15 min, pre-denaturation at 95°C for 5 min, 35 amplification cycles (94°C for 30 sec, 50°C for 30 sec, 72°C 90 sec), and post-extension step at 72°C for 7 min. The Genbank accession numbers of AtSEIPIN1, AtSEIPIN2, and AtSEIPIN3 proteins are AED92296, AEE31126, and AEC08966, respectively. Wild type yeast strain (BY4742), SEIPIN-deletion yeast mutant (*ylr404wΔ*), and yeast expression plasmids (pRS315-PGK, pRS315-*ylr404w* and pRS316-CFP-HDEL) were obtained. The coding regions of *AtSEIPIN1*, *AtSEIPIN2* and *AtSEIPIN3* genes were inserted into yeast expression vector pRS315-PGK using restriction enzymes *Bam*HI and *Pst*I (Promega). Then, the recombined yeast expression plasmids (pRS315-*AtSEIPIN*1, pRS315-*AtSEIPIN2* and pRS315-*AtSEIPIN3*) containing *Arabidopsis SEIPIN* cDNAs were transformed into SEIPIN-deletion yeast mutant (*ylr404wΔ*) with Frozen-EZ Yeast Transformation II Kit™ (Zymo Research). The transformed yeast cells were selected by synthetic complete (SC)-Leu medium and then further confirmed by colony PCR.

Transient Expression of *A. thaliana Seipins* and Mouse *Fit2* in *N. benthamiana* by infiltration

Arabidopsis SEIPIN coding regions were cloned (as described above) and inserted into plant expression vector pMDC32 respectively to construct plant expression plasmids (pMDC32-*AtSEIPIN1*, *AtSEIPIN2* and *AtSEIPIN3*). The mouse FIT2 gene coding region was obtained and subcloned into pMDC32 vector to be expressed in plants. The recombined plant expression plasmids were transformed into *Agrobacterium tumefaciens* (GV3101) by electroporation. *Agrobacteria* containing appropriate cDNAs were mixed and diluted with infiltration buffer to make the final infiltration mixtures, which were used to infiltrate *N. benthamiana* leaf tissue. The recipe of infiltration buffer, *N. benthamiana* and *Agrobacterium* growth conditions, and infiltration procedures were described by Petrie *et al.*, 2010. Tomato bushy stunt virus protein P19 (Genbank accession number: AAB02538) plant expression plasmid pORE04-P19 was obtained and was included in all infiltration mixtures to enhance the gene expression in *N. benthamiana* leaf tissue. *A. thaliana* LEAFY COTYLEDON2 (AtLEC2) in pORE04 was also included in appropriate infiltration mixtures to enhance the synthesis of triacylglycerol (TAG) and further to simulate "seed metabolism" in *N. benthamiana* leaf tissue. The expression of different genes in *N. benthamiana* leaf tissue was tested at the transcriptional level by using RT-PCR. RNA was purified from *N.*

benthamiana leaf tissue by RNeasy Plant Mini kit (Qiagen), and treated by DNase (Promega) to avoid any DNA contamination. RT-PCR was performed by using One-Step Ex Tag RT-PCR kit (Takara). The reverse transcription step was incubation at 42°C for 15 min. The pre-denaturation step was at 95°C for 5 min. The post-extension step was at 72°C for 7 min. *EF1α* and *P19* were amplified by 28 cycles with 94°C for 30 sec, 55°C for 30 sec and 72°C for 1 min. *AtLEC2* and *AtSEIPIN1* were amplified by 35 cycles with 94°C for 30 sec, 50°C for 30 sec and 72°C for 1 min. *AtSEIPIN2* and *AtSEIPIN3* were amplified by 35 cycles with 94°C for 30 sec, 50°C for 30 sec and 72°C for 1.5 min. For samples infiltrated with less than two genes, infiltrated with three cDNAs, and infiltrated with more than three genes, 50 ng, 100 ng and 200 ng of total RNA were used for amplification, respectively.

Lipid Analysis and Colocalization

To visualize lipid droplets (LD) in yeast cells, yeast cells were grown in appropriate SC drop-out medium (with glucose or oleic acid) at 28°C to stationary phase (OD600 ~ 3.0), and lipid droplets were stained with 0.4 μg/ml Bodipy FL (Invitrogen, from 4 mg/ml stock in DMSO) in 50 mM PIPES buffer (pH = 7). To visualize lipid droplets in *N. benthamiana* leaf tissue, leaf discs were collected 5-7 days after infiltration, and lipid droplets were stained with 2 μg/ml Bodipy FL (from 4 mg/ml stock in DMSO) in 50 mM PIPES buffer (pH = 7). To colocalize *Arabodopsis* SEIPINs, ER and LDs in yeast cells, *Arabidopsis* SEIPINs were fused with GFP at both N and C terminus and inserted in yeast expression plasmid pRS315-PGK. Endoplasmic Reticulum (ER) was indicated by ER marker (pRS316-CFP-HDEL) co-expressed with GFP-fused Arabidopsis SEIPINs. LDs were stained with 0.4 μg/ml Nile Red (Sigma Aldrich, from 1 mg/ml stock in DMSO) in 50 mM PIPES buffer (pH = 7) to avoid overlapping of emission spectra with GFP and CFP. To colocalize mouse FIT2 and LDs in *N. benthamiana* leaf tissue, FIT2 was fused with GFP at N terminus and lipid droplets were stained with 2 μg/ml Nile Red (from 1 mg/ml stock in DMSO) in 50 mM PIPES buffer (pH = 7). Confocal images were acquired by Zeiss LSM10 confocal laser scanning microscope (funded by NSF- MRI grant #1126205). GFP and Bodipy FL was excited by 488 nm laser and the emission signal was collected in a spectra of 500-540 nm. CFP was excited by 405 nm laser and the fluorescent signal was collected from 450nm to 500 nm. Nile Red was excited by 488 nm laser and the emission was acquired from 520 nm to 560 nm. Chloroplast autofluorescence was collected in spectra of 640-720 nm. Both 2-D images and single

images in Z-stack series were saved as 512 × 512-pixel (for yeast) and 1024 ×1024-pixel (for *N. benthamiana*) images.

To profile the effects of AtSEIPINs on LD morphology in different organisms (yeast and tobacco), numbers and sizes of lipid droplets were quantified by using ImageJ. In yeast, 3 lines with more than 150 cells for each strain were used for number quantification, and 3 lines with 30 LDs for each strain were used for size quantification. For LD statistics in *N. benthamiana*, 9 confocal images from 3 individual infiltrations for each transient expression were used to quantify the number of LDs for different size categories.

Quantification of TAG content and composition in different yeast strains

Yeast cells were grown in appropriate SC drop-out medium (with glucose) until stationary phase (OD ~ 3.0) and about 50 OD600 units cells were used for lipid extraction. The cells were disrupted by glass beads and bead beater (BioSpec Mini-Beadbeater-16), and 5 μg TAG (tri-15:0) standard was added into each sample. Total lipid was extracted by using hot (70°C) isopropanol and chloroform in a ratio of 450 mg sample:2 ml isopropanol:1 ml chloroform at 4°C overnight. Then the total lipid was further purified by adding 1 ml chloroform and 2 ml 1M KCl, followed by washing with 2 ml 1 M KCl twice. The purified lipid was dried under N_2, and stored in 400 μl 1:1 chloroform/methanol at -20°C. The neutral lipid was separated from polar lipid by using solid phase extraction (SPE). The 6 ml silica column (Sigma Aldrich) was cleaned with 3 ml acetone, and then conditioned with 6 ml hexane. Each lipid sample was loaded onto one conditioned column. 5 mL of hexane/diethyl ether 4:1 and hexane/diethyl ether 1:1 were used to elute neutral lipids. Then, 3 mL methanol and 3 mL chloroform were loaded to the column to elute polar lipids. The neutral lipid and polar lipid samples were evaporated under nitrogen and re-dissolved in chloroform/methanol 1:1 for storage. To analyze TAG content and composition, 20 μL of neutral lipid for each sample was mixed with 5 μL 500 mM ammonium acetate and 230 μL chloroform/methanol 1:1, and injected into triple quadrupole mass spectrometer. The spectra were acquired using Xcalibur (v.2.0.7), and processed by Metabolite Imager (v.1.0) to quantify the total amount and composition of TAG.

Staining of lipid droplets with Nile red and BODIPY 493/503

Stock solutions contained BODIPY 493/503 dissolved in ethanol at a concentration of 1 mg/ml. This solution is stored in the dark at –20°C).

Nile red is dissolved in DMSO to give a stock solution of 50 μg/ml.

Paraformaldehyde is aspirated off after fixing the cells and the cells are rinsed with PBS. PBS + Nile red (at 1:2000 dilution) or PBS + BODIPY 493/503 (at 1:1000 dilution) is added to the cells and agitated for 15 minutes. The staining solution was aspired out and the cells were washed thrice with PBS. Cells were mounted to observe under the microscope.

EXAMPLES

Following are examples that illustrate procedures for practicing the invention. The examples should not be construed as limiting.

EXAMPLE 1 — INCREASE OF LIPID CONTENT AND INDUCTION OF LIPID DROPLET FORMATION IN PLANTS USING MAMMALIAN PROTEINS ASSOCIATED WITH LIPID METABOLISM

Plant transformation vectors are constructed and are propagated in *Eschericia coli* Top 10 cells. The vectors are sequenced for verification. Plasmid vectors are transformed into Agrobacterium, tunefaciens LBA4404, and the clones are selected and verified by PCR. Arabidopsis plants are transformed by the floral dip method as described in Bent and Clough, Plant J. 1998 Dec; 16(6):735-43, which is herein incorporated by reference in its entirety.

Both wild-type plants (*A. thaliana*, ecotype Columbia), and plants with a transfer DNA (T-DNA) insertion mutation in the At4g24160 locus are used for transformations. The T-DNA knockout is in an exon of the Arabidopsis homolog of the human CG1-58 gene. For Arabidopsis plants with CGI-58 mutation, there is an increase in cystosolic lipid droplets in leaves when compared to wild-type plants (James *et al.*, Proc Natl Acad Sci USA. 2010 Oct 12;107(41):17833-8).

Figure 1 is a diagram that illustrates the elements in the T-DNA regions of plant binary transformation vectors. Plants are allowed to set seed and the seed are screened on hygromycin medium for identification of transgenic plants.

Cystolic lipid droplets are normally low in abundance in leaves of wildtype plants and they can be visualized by neutral-lipid-specific fluorescent stains like Nile blue (Fig. 2) or Bodipy493/503 (Fig. 3). The loss of function mutant, cgi-58, results in more lipid droplets than in wildtype plants (James *et al.*, Proc Natl Acad Sci USA. 2010 Oct 12;107(41):17833-17838; see also Fig. 3. vs. Fig. 2). Expression of mouse FSP27 in either the wild-type or the cgi-58 background accentuates lipid droplet accumulation (Figs. 2-4).

Total fatty acid content is measured in seedlings as a crude estimate of changes in lipid content. Fatty acid methyl esters are quantified by gas chromatography-flame ionization detection (GC-FID) using heptadecanoic acid as an internal standard. Transgenic Tl seedlings are grown on hygromycin medium, and plants with five rosette leaves are combined for extraction. Total lipids are extracted and fatty acid methyl esters are prepared according to Chapman and Moore (Arch Biochcem Biophys. 1993 Feb 15;301(1):21-33), which is herein incorporated by reference in its entirety.

The results show that transformed lines expressing FSP27 in the T1 generation have higher total fatty acid content than that of corresponding non-transformed plants on a fresh weight (Fig. 5) and a dry weight (Fig. 6) basis. Transformed lines being homozygous for FSP27 will exhibit greater increase in total fat content. Also, there will be a greater increase in total fat content when neutral lipids are separated from polar membrane lipids, since changes in fat content will be in triacylglycerol levels only, but not to bulk changes in membrane lipids.

EXAMPLE 2 — GENERATION OF FSP27 AND PLIN2 EXPRESSING HOMOZYGOUS TRANSGENIC PLANTS WITH HIGH LIPID CONTENT

Seven homozygous lines of FSP27-expressing plants in the *cgi58* mutant background, as well as one homozygous line expressing PLIN2 (ADRP) are raised. The new plants are completely viable and healthy with higher lipid accumulation as shown by microscopic data (Fig. 7).

Seedlings are grown on solidified nutrient medium under selection. Seven Arabidopsis homozygous lines in T2 generation over-expressing the FSP27 in the *cgi58* knockout background are identified. Also, one Arabidopsis homozygous line in T2 generation overexpressing the ADRP in the *cgi58* knockout background is identified. Lines that are no longer segregating (homozygous) are selected for harvest and extraction. Figure 7 shows representative confocal images of leaves having preponderance of lipid droplets in both lines as well as the cgi-58 knockout background.

EXAMPLE 3 — IDENTIFICATION OF TRIGLYCERIDE-ACCUMULATORY DOMAIN OF FSP27

Using deletion-mutagenesis, the domain of amino acids 120-220 of the mouse FSP27 protein (SEQ ID NO: 2), which is associated with lipid accumulation in adipocytes, is

dissected. The domain 120-220 of mouse FSP27 is a core-portion of FSP27 protein. As shown in Fig.8, adipocytes expressing amino acids 120-220 of the mouse FSP27 protein accumulate lipids faster than adipocytes expressing the full length mouse FSP27 protein.

The present invention also provides genetically engineered plants expressing only the triglyceride-accumulating domain of FSP27 (such as amino acids 120-220 of mouse FSP27), in order to accumulate lipids / oils at a faster rate than the full length protein. For the plants that need to be harvested from time to time for biofuel production, expressing the triglyceride-accumulating domain can be useful for improving lipid / or production.

EXAMPLE 4 — EXPRESSION OF MAMMALIAN AND FISH ANALOGS OF FSP27/CIDEC/CIDE-3 IN PLANTS TO INCREASE LIPID CONTENTS

Homologs of mammalian proteins associated with lipid metabolism can be used to increase lipid / oil contents in transgenic plants. FSP27 plays a key role in triglyceride accumulation in mammals such as mouse and humans. As shown in Fig. 9, mammalian FSP27 and the zebra fish homolog of FSP27 protein share higher than 85% sequence similarity. In one embodiment, mammalian FSP27 and/or fish homologs of FSP27 can be used for expression in plants to generate transgenic plants with high oil and/or lipid contents.

EXAMPLE 5 — INCREASE OF LIPID CONTENT IN PLANTS BY EXPRESSING A COMBINATION OF PROTEINS ASSOCIATED WITH LIPID METABOLISM

In certain embodiments, to increase and maximize the efficiency of oil production in plants, transgenic plants are genetically modified to express a combination of proteins associated with lipid metabolism and peptides. Proteins or polypeptides associated lipid metabolism useful for improving plant lipid/oil content include, but are not limited to, proteins and peptides involved in lipid (such as triglyceride) metabolism, such as, for example, proteins involved in the synthesis, protection, accumulation, storage, and breakdown of lipid (such as triglyceride).

For instance, FSP27 expression in plants increase plant lipid / oil content, and FSP27 expressed in CGI58-mutants results in even greater increase in lipid / oil content.

In certain embodiments, the present invention provides transgenic plants expressing a combination of proteins associated with lipid metabolism including, but not limtied to, DGAT-1, PDAT-1, *cgi58* mutation, SEIPIN, FIT1, FIT2, PLIN1, PLIN2, FSP27/Cidec/cide-3, and Cidea.

In certain embodiments, the transgenic plants express a combination of nucleic acids expressing proteins associated with lipid metabolism selected from: DGAT-1 and FSP27; DGAT-1, *cgi58* (mutation), and FSP27; DGAT-1, PDAT-1, and FSP27; DGAT-1, PDAT-1, *cgi58* (mutation), FSP27; FSP27, PLIN2, and *cgi58* (mutation); DGAT-1, FSP27, PLIN2, and *cgi58* (mutation); and DGAT-1, PDAT-1, FSP27, PLIN2, and *cgi58* (mutation).

In one embodiment, a combination of "triglyceride accumulation" proteins is expressed in leaves of plants with globally up-regulated fatty acid biosynthesis. Plants with globally up-regulated fatty acid biosynthesis include, but are not limited to, plants with the WRINKLED1 transcription factor mis-expressed in leaves. The WRINKLED1 transcription is involved in the regulation of fatty acid biosynthesis. See Sanjaya *et al.*, 2011, Plant Biotechnology Journal (2011) 9, pp. 874–883), which is hereby incorporated as reference in its entirety.

EXAMPLE 6 – Homologues of Human Lipodystrophy genes in *A. thaliana*

Table 1 shows Homologues of Human Lipodystrophy genes in *A. thaliana*

Human gene	Protein function	Candidate Arabidopsis homolog(s)[a]
Agpat2	LPAT, synthesis of phosphatidic acid	*At1g80950; At1g51260; At3g57650; At3g18850; At1g75020; At4g30580*
Bscl2	SEIPIN, role in LD morphology	*At5g16460; At1g29760; At2g34380*
Akt2	Protein Kinase B	*At3g08730; At3g08720; At5g04510*[b]*; At310540*[b]
Zmpste24	Zinc metalloprotease; processing of lamin subunits	*At4g01320*
Cgi-58	Co-activator of ATGL, also has LPAT activity	*At4g24160*
Lipa	Lysosomal acid lipase; hydrolyzes cholesteryl esters and TAGs	*At5g14180; At2g15230*

[a]Best match by WU-BLAST against the Arabidopsis genome at TAIR [www.arabidopsis.org].

[b]Contains pleckstrin homology domains and has phosphoinositide-3-dependent kinase activity.

EXAMPLE 7 — INCREASED LIPID CONTENT IN PLANTS BY EXPRESSING PROTEINS ASSOCIATED WITH LIPID METABOLISM OR COMBINATIONS THEREOF

Proteins associated with lipid metabolism of animal origin, for example, mouse and human, or of plant origin, for example, *A. thaliana*, were transiently expressed in vegetative tissues of plants, for example, *N. benthamiana* (a close relative of tobacco and species of *Nicotiana* indigenous to Australia) and *A. thaliana*. Increased lipid accumulation in lipid droplets of plants transiently expressing exogenous proteins or polypeptide associated with lipid metabolism was observed indicating that overexpression of exogenous proteins associated with lipid metabolism in vegetative tissue of plants can be used to increase lipid production in these plants and such plants provide a valuable means of producing higher yields of biofuel.

Further, plants permanently expressing exogenous proteins or polypeptide associated with lipid metabolism, for example, having the exogenous proteins associated with lipid metabolism incorporated in the genomes of the plants to produce transgenic plants, can also be used to produce higher amounts of lipids in such plants. These plants can also provide valuable means of producing higher yields of biofuel.

Examples of techniques of expressing endogenous lipid droplets in vegetative tissues of plants and increased lipid accumulation in plants expressing exogenous proteins associated with lipid metabolism are provided in Figures 25 to 36.

Over-expression of SEIPINs in leaves enhances the capacity for neutral lipid storage, and provides additional strategies to engineer increased neutral lipid accumulation in plant cells, including even subcellular “packages” of different sizes. Transient overexpression of SEIPINs in tobacco leaves increases lipid droplet numbers and influences the size of LDs (S1, large; S2, medium; S3 small). The current invention provides that permanent overexpression of proteins associated with lipid metabolism, such as SEIPINs, can be used to produce higher amounts of oil in plants as compared to wild type plants of the same type.

EXAMPLE 8 — INCREASE OF LIPID CONTENT IN YEAST CELLS BY EXPRESSING PROTEINS ASSOCIATED WITH LIPID METABOLISM OR COMBINATIONS THEREOF

Wild type cells of *S. cerevisiae* produce lipid droplets (see, Figure 13, top left panel). A yeast SEIPIN (ScSEIPIN) plays an important role in the production of these lipid droplets in *S. cerevisiae* as shown by reduced accumulation of lipids in *S. cerevisiae* mutant (*ylr404wΔ*) lacking ScSEIPIN activity (see, Figure 13, top middle panel). The role of ScSEIPIN in lipid droplet production in yeast is further confirmed by restoration of lipid accumulation in *ylr404wΔ* expressing ScSEIPIN. Figure 13, bottom panels, further show that expression of exogenous SEIPINs, namely SEIPIN 1, 2, or 3 from *A. thaliana* also restores lipid accumulation in *ylr404wΔ*.

Further, expression of SEIPIN 1, 2, or 3 in *ylr404wΔ* produces lipid droplets of varying morphologies (see Figures 13-16 and 24). For example, overexpression of AtSEIPIN1 produces lipid droplets of larger size than the wild type, whereas overexpression of AtSEIPIN2 or 3, without affecting the size of the lipid droplets, increases the number of lipid droplets per yeast cell compared to *ylr404wΔ mutant*.

Furthermore, overexpression of AtSEIPIN 1, 2, or 3 in *ylr404wΔ* restores the amount of TAG accumulation comparable to that found in the wild type yeast cells (see, Figures 21-23).

These data show that the three *A. thaliana* SEIPIN homologues provide different developmental expression profiles. All localize to discrete domains of ER in heterologous system (yeast). AtSEIPINs 2 and 3 partially complement yeast mutants, indicating they function generally in a similar manner to yeast and human SEIPIN in the regulation of LD number and shape. AtSEIPIN1 generates supersize LDs in yeast (and plants).

EXAMPLE 9 — COLOCALIZATION OF SEIPINS AND LIPID DROPLETS IN YEAST

AtSEIPINs, when overexpressed in *ylr404wΔ,* localize to lipid droplets which further confirms the role of SEIPINs in lipid droplet accumulation in yeast (see, for example, Figures 17-20). AtSEIPIN-GFP and CFP-HDEL were overexpressed in a yeast cells. Conjugation with GFP allowed visualization of the location of AtSEIPINs in a cell by green fluorescence (see, Figures 18-20, top right panels), whereas expression of CFP-HDEL allowed visualization of endoplasmic reticulum as blue fluorescence in the yeast cell (see, Figures 18-

20, bottom left panels). Lipid droplets in these yeast cells is visualized by Nile Red staining (see, Figures 18-20, top left panels).

Overlapping the top left, bottom left, and top right columns in Figures 18-20 indicates that green fluorescence coming from AtSEIPIN GFP fusion proteins largely co-localized with the yellow staining of lipid droplets. Blue fluorescence of CFR-HDEL did not colocalize with either the lipid droplets or the AtSEIPIN GFP fusion proteins.

EXAMPLE 10 – EXPRESSION OF LIPID-DROPLETS ASSOCIATED PROTEINS IN ALGAE TO INCREASE ALGAL LIPID CONTENTS

Overexpression of various proteins associated with lipid metabolism from mammalian and plant origin, for example, FSP27, Cidea, PLIN1, PLIN2, SEIPIN, FIT1, and FIT2 in various cell types cause 3-10 fold increase in fat accumulation. Algae are widely used as an organism for production of biofuel. Accordingly, the current invention further provides algal cells expressing one or more of the proteins associated with lipid metabolism, either from animal or plant origin. These algal cells contain higher amounts of oil/fat.

Examples of various proteins or polypeptides associated with lipid metabolism that can be expressed in algae to produce increased oil in algae include, but are not limited to FSP27, Cidea, ADRP, PLIN1, FIT , /2, SEIPIN, SEIPIN 1, SEIPIN 2, SEIPIN 3, DGAT1, DGAT2, PDAT1, WRI1, and mutant CGI-58. Examples of algae that can be used according to the current invention to produce oil include, but are not limited to algae from *Chlamydomonas spp.*, *Botryococcus braunii*, *Chlorella spp.*, *Dunaliella tertiolecta*, *Gracilaria spp.*, *Pleurochrysis carterae* (also called CCMP647), *Sargassum spp.*, and *Eudorina elegans*.

Non-limiting examples of various fuel types that can be produced in algae expressing exogenous proteins associated with lipid metabolism include biodiesel, biobutanol, biogasoline, methane, ethanol, vegetable oil fuel, hydrocracking to traditional transport fuels, and jet fuel.

Thus, the algal cells of the current invention can be used to produce energy with higher efficiency and at a cost effective manner.

Algal cells of the current invention can also be used to increase production of oils which are beneficial for human health, e.g. omega-unsaturated fat in olives, canola oil, etc. For example, fatty acid analysis in FSP27 expressing plants show that besides increase in

overall oil content the content of omega-3 fatty acids, particularly linoleic (18:2) and alpha-linolenic (18:3) fatty acid, is increased in these plants.

Certain proteins associated with lipid metabolism play a positive regulatory role in improving the metabolic health in humans suffering from insulin resistance, type 2 diabetes, cardiovascular disease, etc. Generating algae expressing such proteins associated with lipid metabolism can have therapeutic use based on the positive role played by these proteins.

Various techniques discussed in references 11-14 can be used to genetically manipulate algae according to the current invention and are expressly incorporated by reference herein. Methods of genetically manipulating algae, in addition to those described in references 11-14, are well known to a person of ordinary skill in the art and such methods are within the purview of the current invention.

Non-limiting examples of vectors used for transformation in algae include pPmr3 plasmid, pmfg-GLuc (mfg refers to "my favorite gene"), pALM32, and pALM33.

All references, including publications, patent applications and patents, cited herein are hereby incorporated by reference to the same extent as if each reference was individually and specifically indicated to be incorporated by reference and was set forth in its entirety herein.

The terms "a" and "an" and "the" and similar referents as used in the context of describing the invention are to be construed to cover both the singular and the plural, unless otherwise indicated herein or clearly contradicted by context.

Recitation of ranges of values herein are merely intended to serve as a shorthand method of referring individually to each separate value falling within the range, unless otherwise indicated herein, and each separate value is incorporated into the specification as if it were individually recited herein. Unless otherwise stated, all exact values provided herein are representative of corresponding approximate values (*e.g.*, all exact exemplary values provided with respect to a particular factor or measurement can be considered to also provide a corresponding approximate measurement, modified by "about," where appropriate).

The use of any and all examples, or exemplary language (*e.g.*, "such as") provided herein, is intended merely to better illuminate the invention and does not pose a limitation on the scope of the invention unless otherwise indicated. No language in the specification should be construed as indicating any element is essential to the practice of the invention unless as much is explicitly stated.

The description herein of any aspect or embodiment of the invention using terms such as "comprising", "having", "including" or "containing" with reference to an element or

elements is intended to provide support for a similar aspect or embodiment of the invention that "consists of", "consists essentially of", or "substantially comprises" that particular element or elements, unless otherwise stated or clearly contradicted by context (*e.g.*, a composition described herein as comprising a particular element should be understood as also describing a composition consisting of that element, unless otherwise stated or clearly contradicted by context).

It should be understood that the examples and embodiments described herein are for illustrative purposes only and that various modifications or changes in light thereof will be suggested to persons skilled in the art and are to be included within the spirit and purview of this application.

REFERENCES

1. Curtis and Grossniklaus, A gateway cloning vector set for high-throughput functinal analysis of genes in planta, Plant Physiology, Vol. 133, p462–469 (2003).

2. Gross *et al.* (2011) PNAS 108, 19581-19586; PMID: 22106267.

3. Jambunathan *et al.*, FSP27 promotes lipid droplet clustering and then fusion to regulate triglyceride accumulation (2011).

4. James *et al.* (2010) PNAS 107, 17833-1838, PMID: 20876112

5. Sanjaya *et al.*, 2011, Plant Biotechnology Journal (2011) 9, pp. 874–883.

6. Szymanski *et al.* (2007) PNAS 104, 20890-5, PMID: 18093937.

7. Zhang *et al.* (2009) Plant Cell 21, 3885-901, PMID: 20040537.

8. U.S. Application Publication No. 2010/0221400.

9. Petrie, J.R., Shrestha, P., Liu, Q., Mansour, M.P., Wood, C.C., Zhou, X.R., Nichols, P.D., Green, A.G., and Singh, S.P. (2010). Rapid expression of transgenes driven by seed-specific constructs in leaf tissue: DHA production. Plant Methods 6, 8.

10. Szymanski, K.M., Binns, D., Bartz, R., Grishin, N.V., Li, W.P., Agarwal, A.K., Garg, A., Anderson, R.G., and Goodman, J.M. (2007). The lipodystrophy protein SEIPIN is found at endoplasmic reticulum lipid droplet junctions and is important for droplet morphology. Proc Natl Acad Sci U S A 104, 20890-20895.

11. Voinnet, O., Rivas, S., Mestre, P., and Baulcombe, D. (2003). An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. Plant J 33, 949-956.

12. Neupert J, Shao N, Lu Y, Bock R. (2012), Genetic transformation of the model green alga *Chlamydomonas reinhardtii*. Methods Mol Biol.; 847:35-47.

Lerche K, Hallmann A. (2013), Stable nuclear transformation of Eudorina elegans. BMC Biotechnol. 13:11.

13. Meslet-Cladière L, Vallon O. (2011), Novel shuttle markers for nuclear transformation of the green alga *Chlamydomonas reinhardtii*. Eukaryot Cell;10(12):1670-8.

14. U.S. Application Publication No. 2009/0176272.

CLAIMS

We claim:

1. A method for obtaining a plant cell or algal cell with elevated lipid content, wherein the method comprises:

genetically modifying a plant cell or algal cell to express an exogenous protein or polypeptide associated with lipid metabolism, thereby obtaining a genetically-modified plant cell or algal cell with elevated lipid content;

wherein the protein or polypeptide associated with lipid metabolism induces adipogenesis, enhances the accumulation of cellular lipid droplets, and/or reduces lipase activity; and

wherein the expression of the protein or polypeptide associated with lipid metabolism increases lipid content of the genetically-modified plant cell or algal cell as compared to a wild-type plant cell or algal cell of the same type.

2. A method according to claim 1, wherein the protein or polypeptide associated with lipid metabolism is selected from fat specific protein 27 (FSP27), PLIN1, PLIN2, SEIPIN, FIT1, FIT2, acyl-CoA:diacylglycerol acyltransferase 1 (DGAT-1), phospholipid:diacylglycerol acyltransferase 1 (PDAT-1), cell death activator (Cidea), leafy cotyledon 2 (LEC2), and WRINKLED1 (WRI1) protein or polypeptide.

3. A method according to claim 1, wherein the protein or polypeptide associated with lipid metabolism is of animal origin.

4. A method according to claim 1, wherein the protein or polypeptide associated with lipid metabolism is a fat specific protein 27 (FSP27) protein or polypeptide.

5. A method according to claim 1, further comprising modifying the plant cell or algal cell to express a combination of exogenous proteins or polypeptides associated with lipid metabolism, wherein at least one exogenous protein or polypeptide associated with lipid metabolism is selected from fat specific protein 27 (FSP27), PLIN1, PLIN2, SEIPIN, FIT1, FIT2, acyl-CoA:diacylglycerol acyltransferase 1 (DGAT-1), phospholipid:diacylglycerol

acyltransferase 1 (PDAT-1), cell death activator (Cidea), LEC2, and WRINKLED1 (WRI1) protein or polypeptide.

6. A method according to claim 1, wherein the cell is a plant cell and the method further comprises regenerating the plant cell into a plant.

7. A method according to claim 1, wherein the genetic modification of the plant cell or algal cell comprises transforming the plant cell or algal cell with a vector comprising a nucleic acid sequence encoding an exogenous protein or polypeptide associated with lipid metabolism, wherein the nucleic acid is operably linked to a promoter and/or a regulatory sequence.

8. A method according to claim 1, wherein the exogenous protein or polypeptide associated with lipid metabolism is selected from *Arabidopsis thaliana* SEIPIN1, SEIPIN2, SEIPIN3, or leafy cotyledon 2 (LEC2).

9. A method according to claim 8, wherein lipid droplet size is enhanced as compared to lipid droplet size of a wild-type cell of the same type.

10. A transgenic plant cell or algal cell having elevated lipid content as compared to a wild-type plant cell or algal cell of the same type, wherein the transgenic plant cell or algal cell expresses an exogenous protein or polypeptide associated with lipid metabolism, wherein the protein or polypeptide associated with lipid metabolism induces adipogenesis, enhances the accumulation of cellular lipid droplets, and/or reduces lipase activity.

11. A transgenic plant cell or algal cell according to claim 10, wherein the protein or polypeptide associated with lipid metabolism is selected from fat specific protein 27 (FSP27), PLIN1, PLIN2, SEIPIN, FIT1,FIT2, acyl-CoA:diacylglycerol acyltransferase 1 (DGAT-1), phospholipid:diacylglycerol acyltransferase 1 (PDAT-1), adipose triglyceride lipase (ATGL), cell death activator (Cidea), LEC2, and WRINKLED1 (WRI1) protein or polypeptide.

12. A transgenic plant cell or algal cell according to claim 11, wherein the protein associated with lipid metabolism is a fat specific protein 27 (FSP27) protein or polypeptide.

13. A transgenic plant cell according to claim 10, wherein the transcenic plant cell is in a plant or plant part.

14. A transgenic plant cell according to claim 13, which is a non-seed cell.

15. A transgenic plant cell according to claim 14, wherein the non-seed cell is a leaf, root, stem, shoot, bud, tuber, fruit, or flower cell.

16. A transgenic plant cell according to claim 13, wherein the cell is a seed cell of a plant.

17. A transgenic plant cell or algal cell according to claim 11, wherein the exogenous protein or polypeptide associated with lipid metabolism is selected from *A. thaliana* SEIPIN1, SEIPIN2, SEIPIN3, or LEC2.

18. A transgenic plant cell or algal cell according to claim 17, wherein lipid droplet size is enhanced as compared to lipid droplet size of a wild-type cell of the same type.

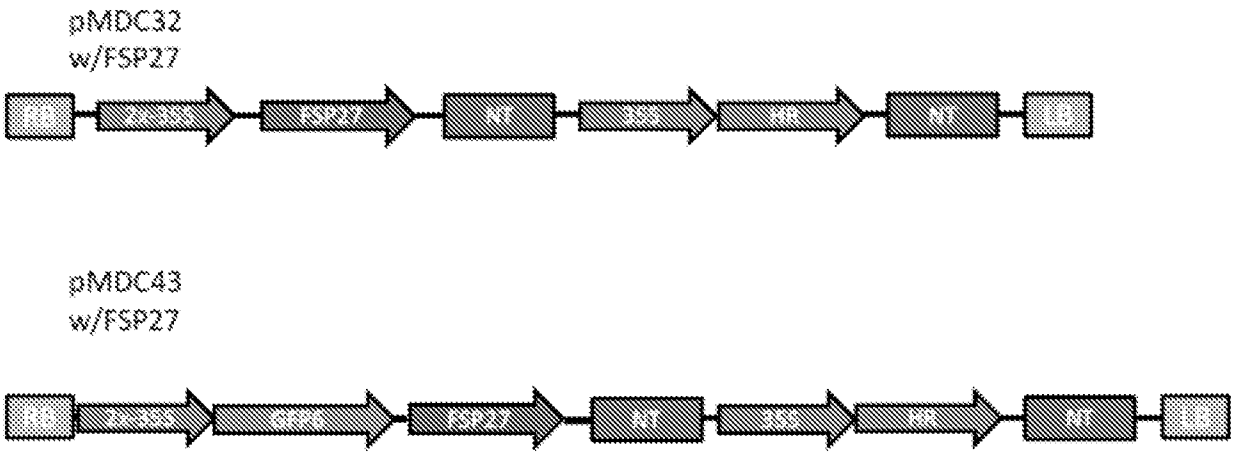
19. A method for screening for a functional protein or polypeptide associated with lipid metabolism for elevating lipid content and/or inducing lipid droplet accumulation in a plant or algal cell, wherein the method comprises:

obtaining a test plant cell or algal cell genetically-modified to express a candidate exogenous protein or polypeptide associated with lipid metabolism; and

growing the genetically-modified test cell and selecting the genetically-modified test cell having elevated lipid content and/or increased lipid droplet size or number, when compared to a wild-type cell of the same type.

20. A method according to claim 19, wherein the cell is a plant cell and the method further comprises regenerating the genetically-modified cell into a plant.

21. The method according to claim 20, further comprising obtaining progeny of said plant.



35S=CaMV 35S promoter
HR=hygromycin resistance
RB=right border
LB=left border
NT=nos terminator

FIG. 1

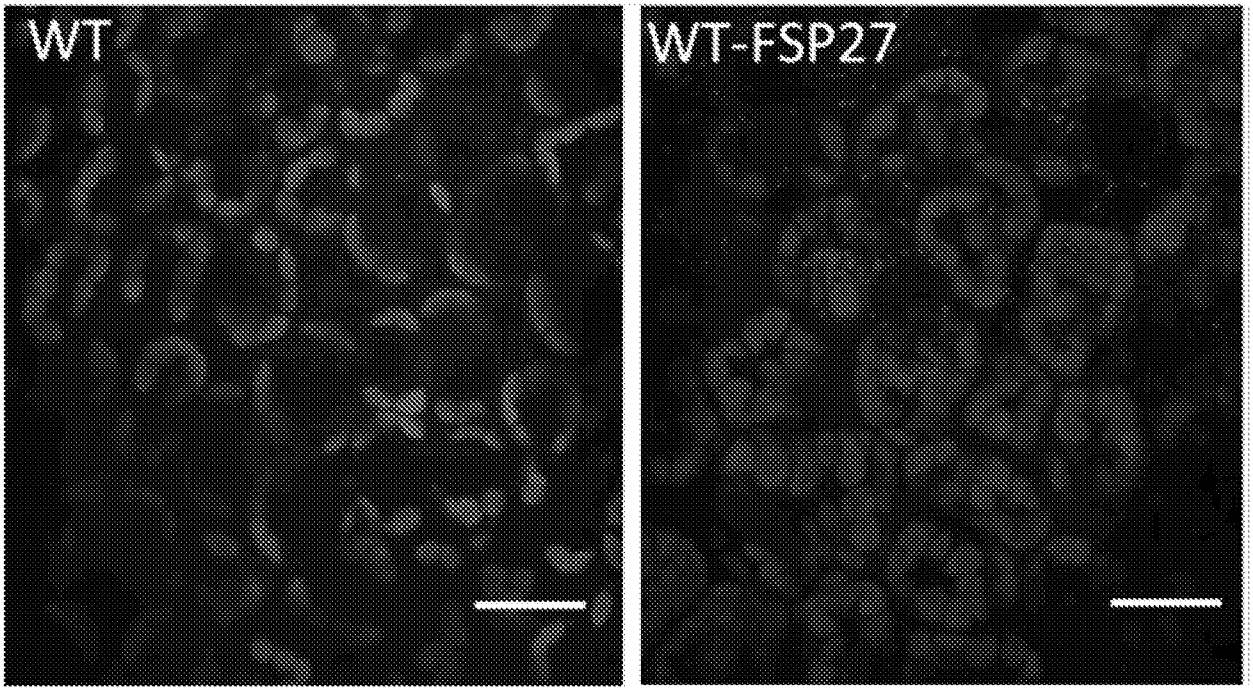


FIG. 2

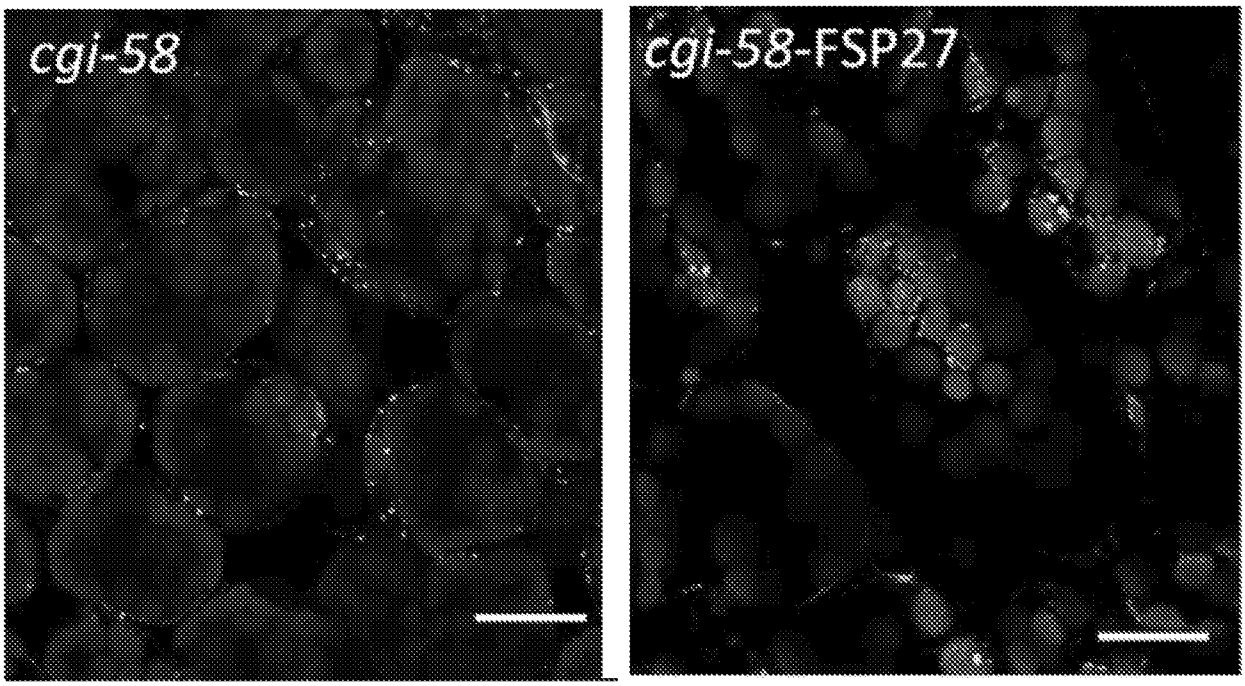


FIG. 3

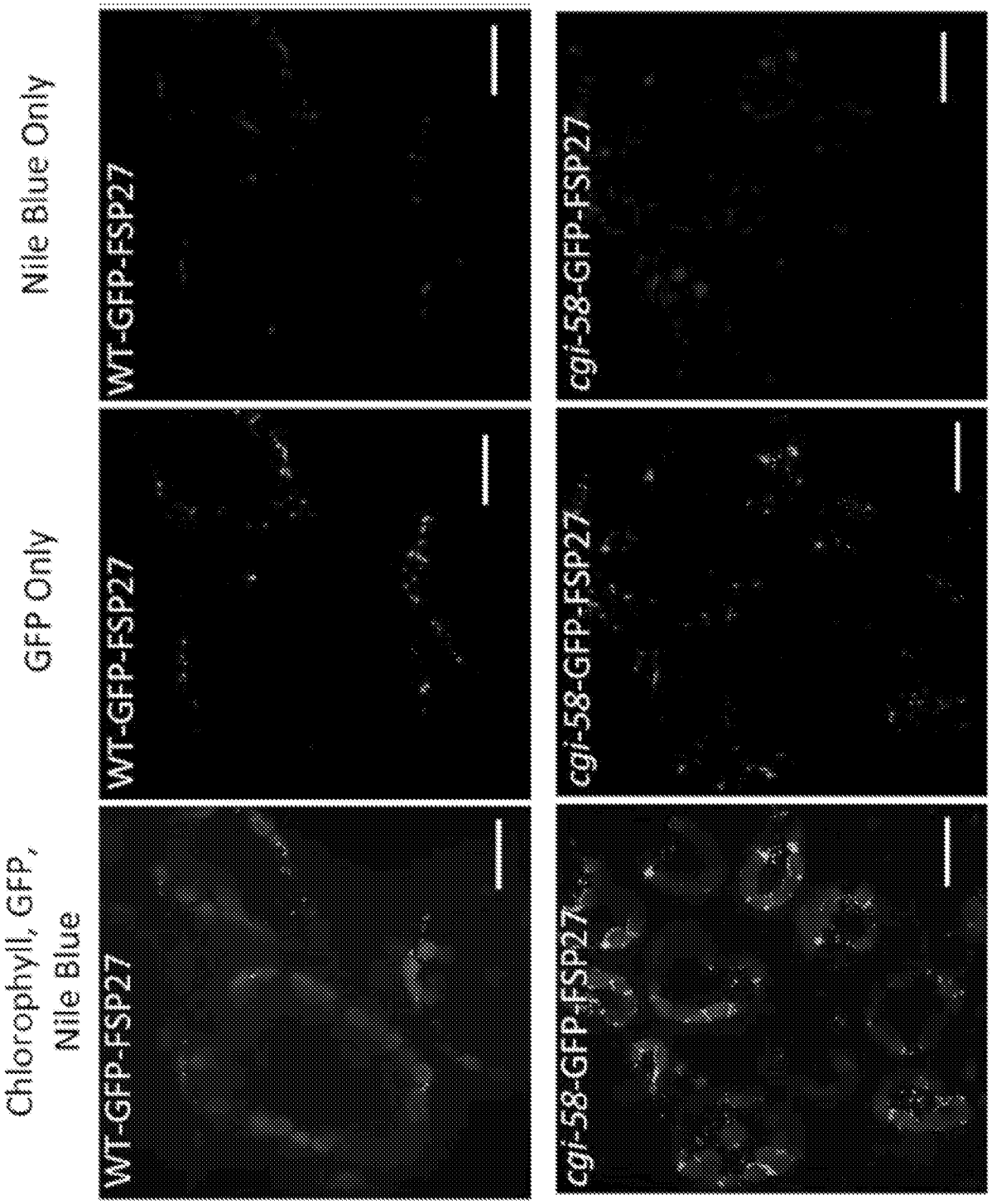

Chlorophyll, GFP, Nile Blue
GFP Only
Nile Blue Only
WT-GFP-FSP27
WT-GFP-FSP27
WT-GFP-FSP27
cgi-58-GFP-FSP27
cgi-58-GFP-FSP27
cgi-58-GFP-FSP27

FIG. 4

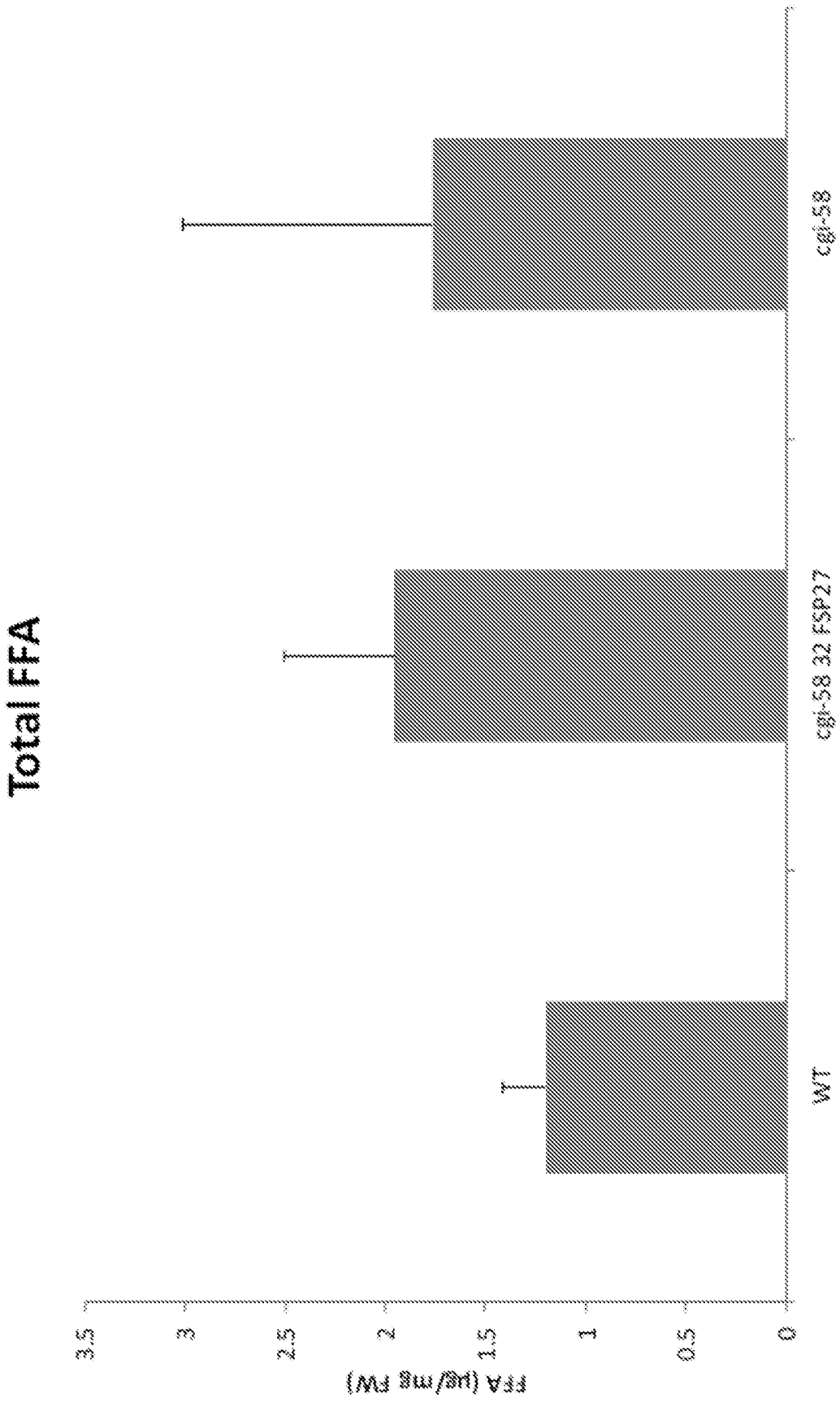

Total FFA
FFA (μg/mg FW)
3.5
3
2.5
2
1.5
1
0.5
0
WT
cgi-58 32 FSP27
cgi-58
Genotype

FIG. 5

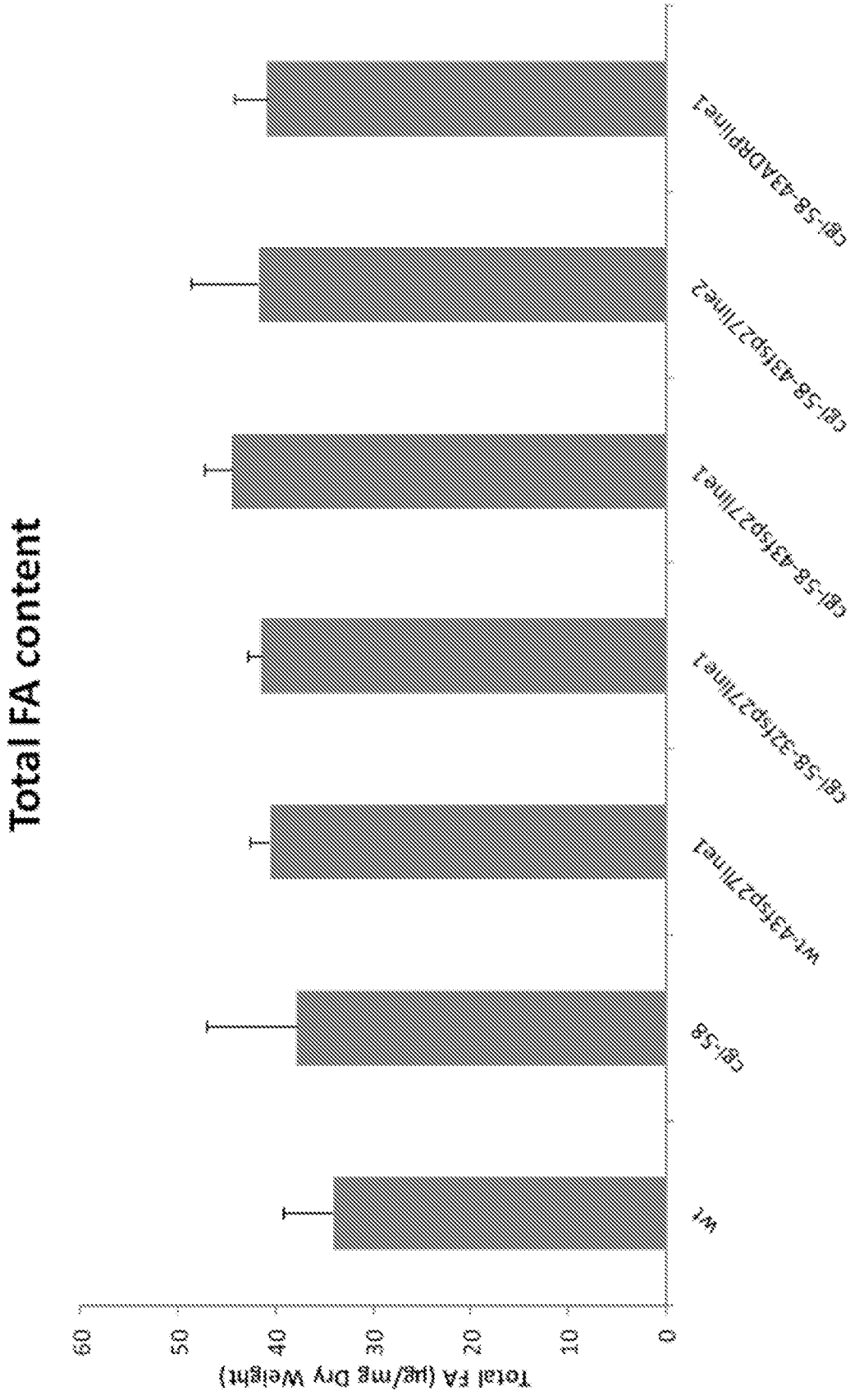

Total FA content
Total FA (µg/mg Dry Weight)
0
10
20
30
40
50
60
Wt
Cgi-58
Wt-43Isp2Tline1
Cgi-58-32Isp2Tline1
Cgi-58-43Isp2Tline1
Cgi-58-43Isp2Tline2
Cgi-58-43ADRPline1

FIG.6

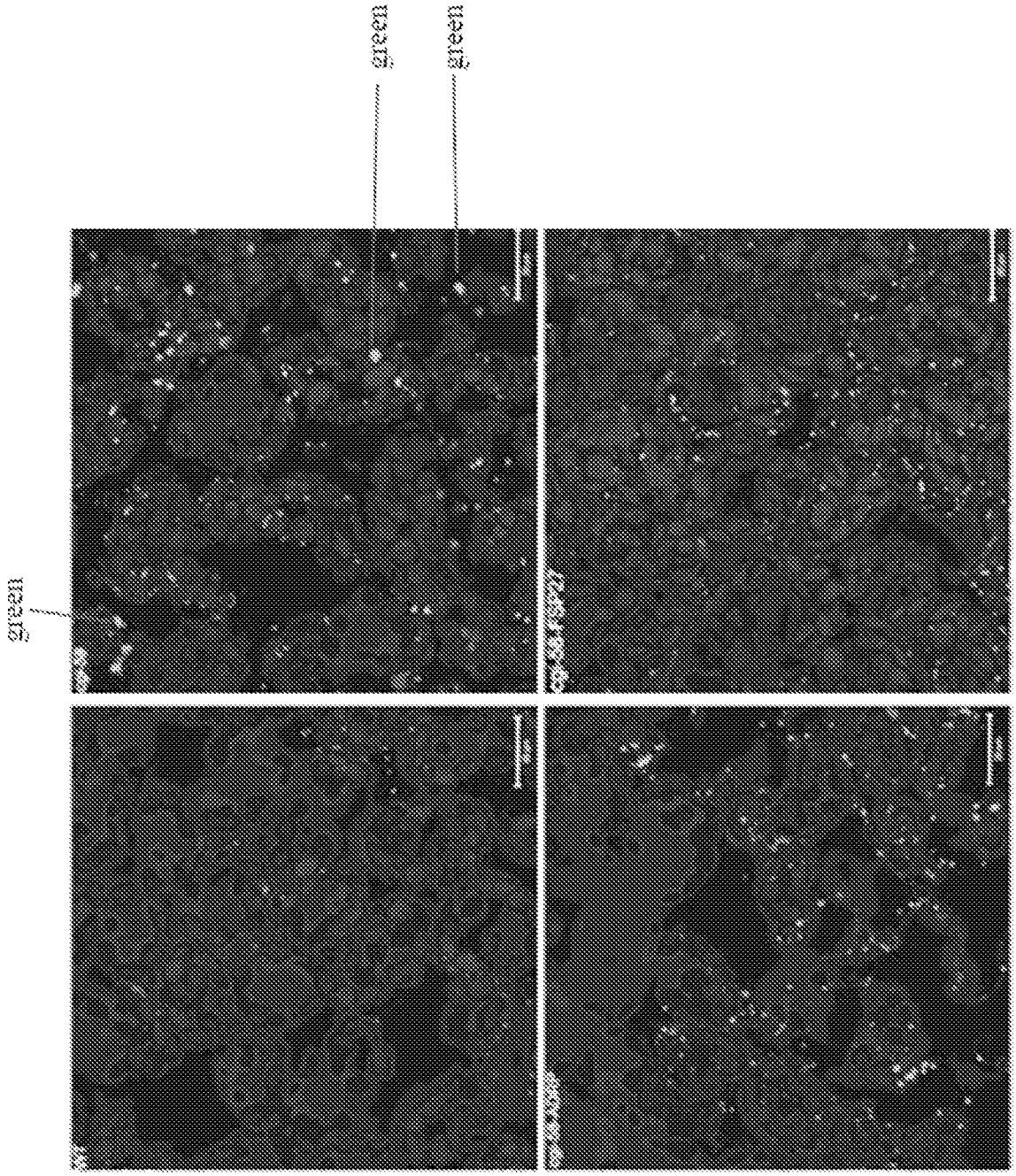


FIG. 7A

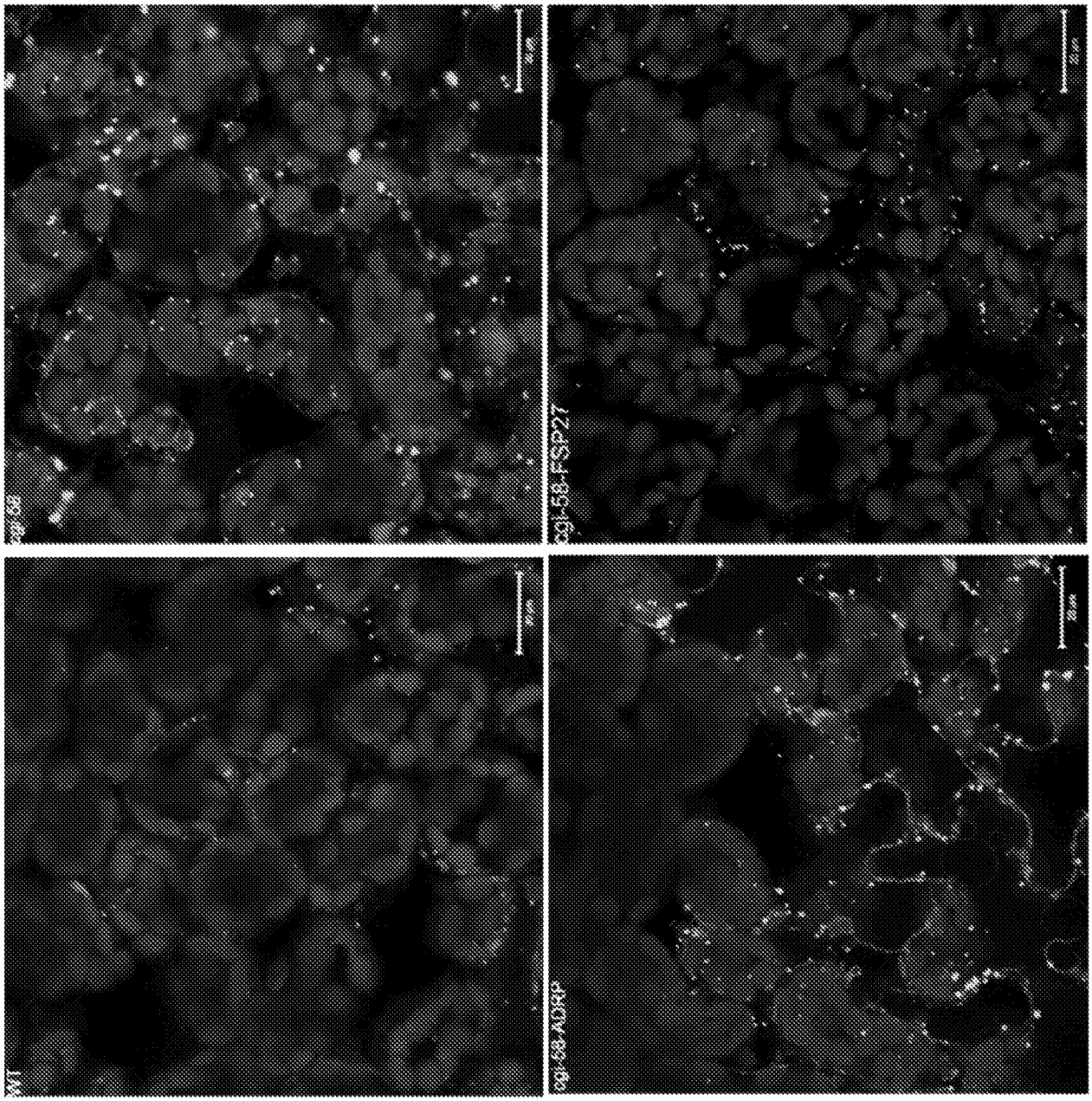


FIG. 7B

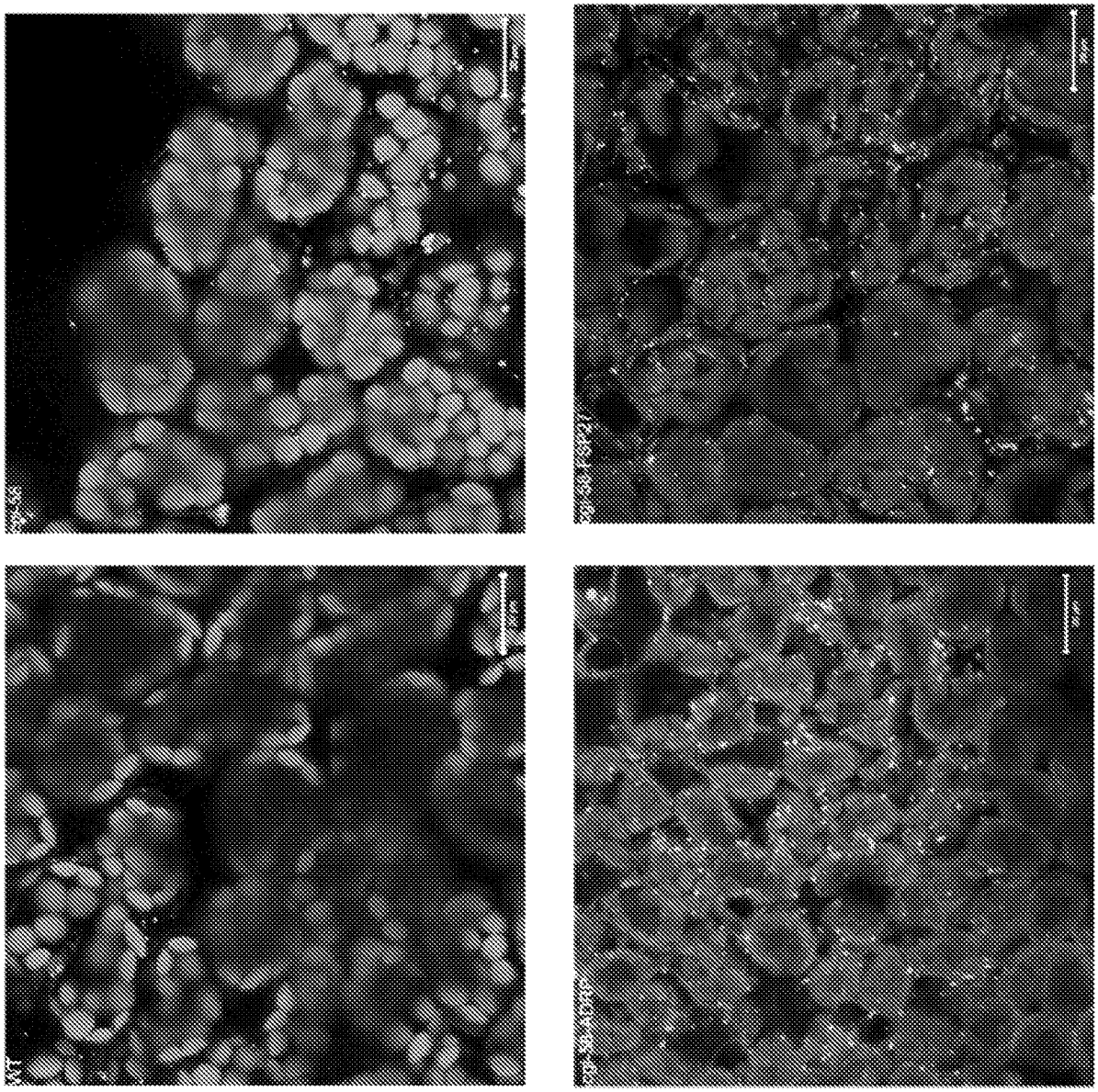


FIG. 7C

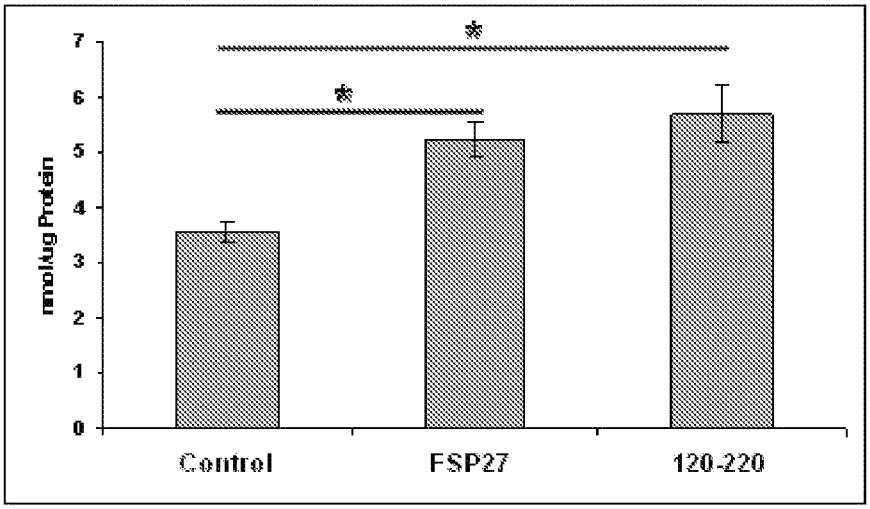


FIG. 8

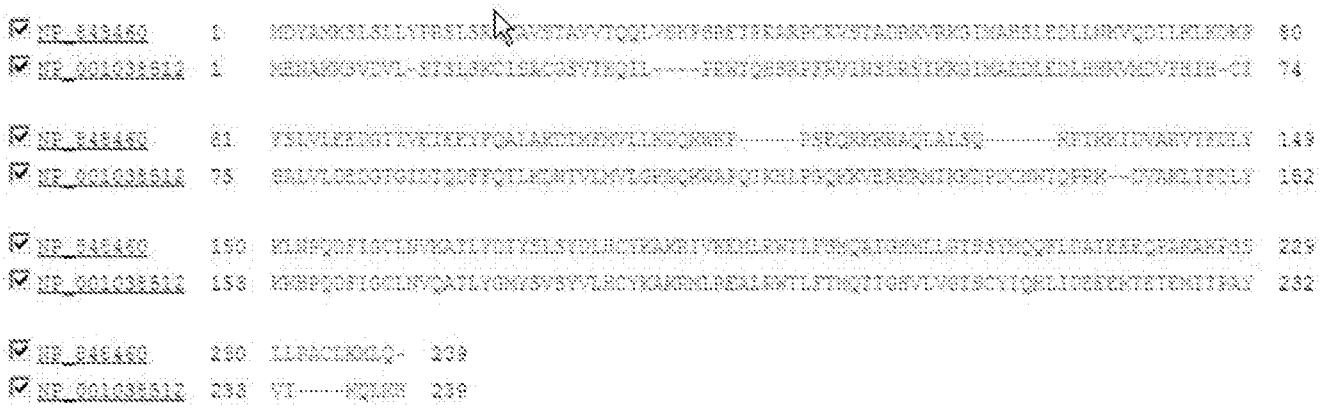


FIG. 9

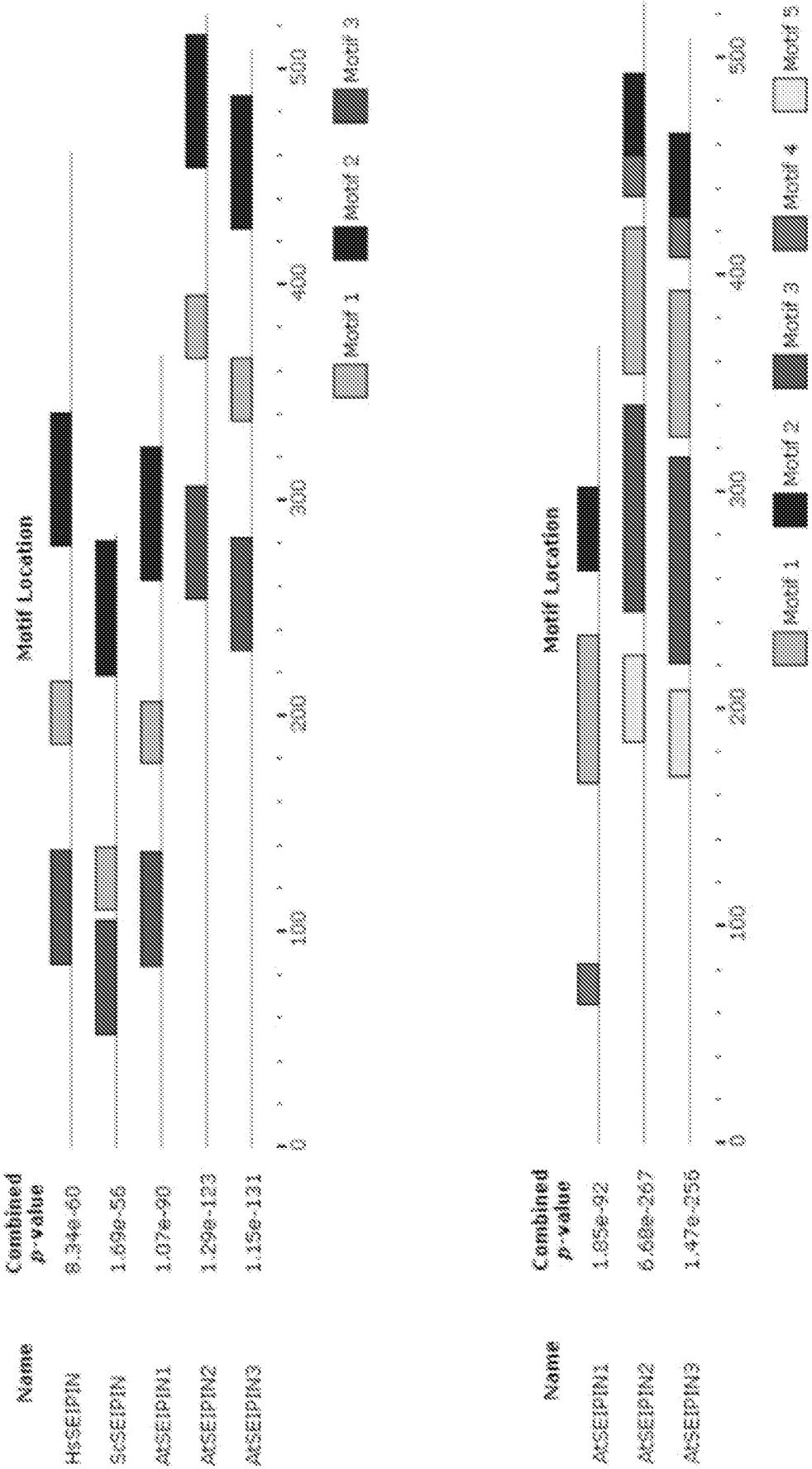

Name
Combined p-value
HsSEIPIN
8.34e-60
SxSEIPIN
1.69e-56
AtSEIPIN1
1.07e-90
AtSEIPIN2
1.29e-123
AtSEIPIN3
1.15e-131
Motif Location
0
100
200
300
400
500
Motif 1
Motif 2
Motif 3
Name
Combined p-value
AtSEIPIN1
1.85e-92
AtSEIPIN2
6.68e-267
AtSEIPIN3
1.47e-256
Motif Location
0
100
200
300
400
500
Motif 1
Motif 2
Motif 3
Motif 4
Motif 5

FIG. 10

Motif 1 E-value=6.5e^{-9}	
HsSEIPIN	GMFLVTISCYTRGGRIISTSSRSVMLHY
ScSEIPIN	LYCQSKTDHLNLDNLLIDVYRGIGFLLG
AtSEIPIN1	GVFQLKVELLSLKGETIARSSQICMLRF
AtSEIPIN2	GMFQVRVDFLSVDGQTIASIRRICMLRF
AtSEIPIN3	GMFQVRVDFLSASGHVLASSRRFCMVKF
Consensus Sequence	GMFQVKVDFLSLDGHTIASSRRICMLHF
Motif 2 E-value=4.2e^{-15}	
HsSEIPIN	IHAVFTGLRYLLYNFIMTCAFIGVASNFTFLSVIVLFSYMQWVWGGIWFRHRFSLQVNIRK
ScSEIPIN	IHFESGIKFRMNFEQGLRMLMLRKRFLSYIIGISIFHCIICVLFFITHCTAFIFVRKGQEK
AtSEIPIN1	LFQLYEAEIVINSKGIWIKRMAYNWKWTLCVWTSMYLYVAILTALLWCFRIVLFIYTSSRT
AtSEIPIN2	IFELYDASLSVESGLIFFRKIIWKWRKTLFVWISMSLFITELLFTLVCCRILIIFRTQFRD
AtSEIPIN3	IFEIYDASLFLESKLIFLKRIIWNWRKTLFVWISMSLFIMELLFALVFFRILIIFRTGQRT
Consensus Sequence	IHEIYDAKLYMNYKLIFIKRMIWNWRKTLFVWISMYIFIMCLLFALVCCRILIFFRTGQRK
Motif 3 E-value=2.4e^{-6}	
HsSEIPIN	GRARRLLLQFGVLFCTILLLLWVSVFLIGSFYYSYMITVSHLSIVIFIYRID
ScSEIPIN	VILNTFNILNGVQFGIKFFQSIKSIFVGTDLFQTIDNGLSQLIFMRDNMEYK
AtSEIPIN1	VRRTWFGILGACIVSMVMVLALILAVVIGVGIVSLYVEKFVVVRDRLFFDYT
AtSEIPIN2	VCKFGWGMFWAVYVGIVLFGLLVSSLMIGGYVINRIADKIFEVKETLNFDYT
AtSEIPIN3	VRRVTWGLFCAVYVGIMLFALLVSAFMISGFVITILANEFLVIKESLNFDYT
Consensus Sequence	VRFNWWGIFWAVFVGIVLFQLLVSAFMIGGFVITIYMAIKFFVIKEFLNFDYT

FIG. 11

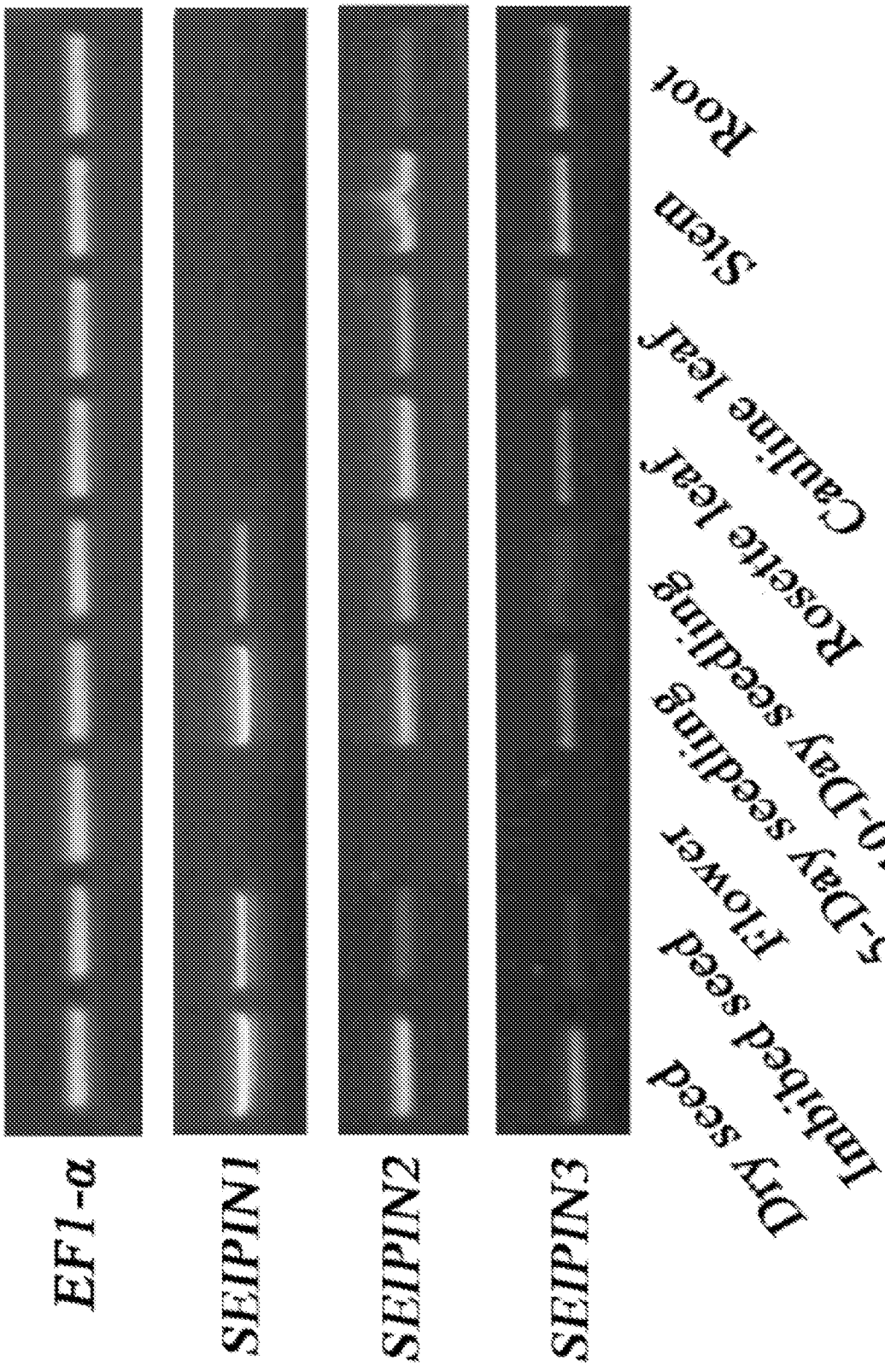

EF1-α
SEIPIN1
SEIPIN2
SEIPIN3
Dry seed
Imbibed seed
Flower
5-Day seedling
10-Day seedling
Rosette leaf
Cauline leaf
Stem
Root

FIG. 12

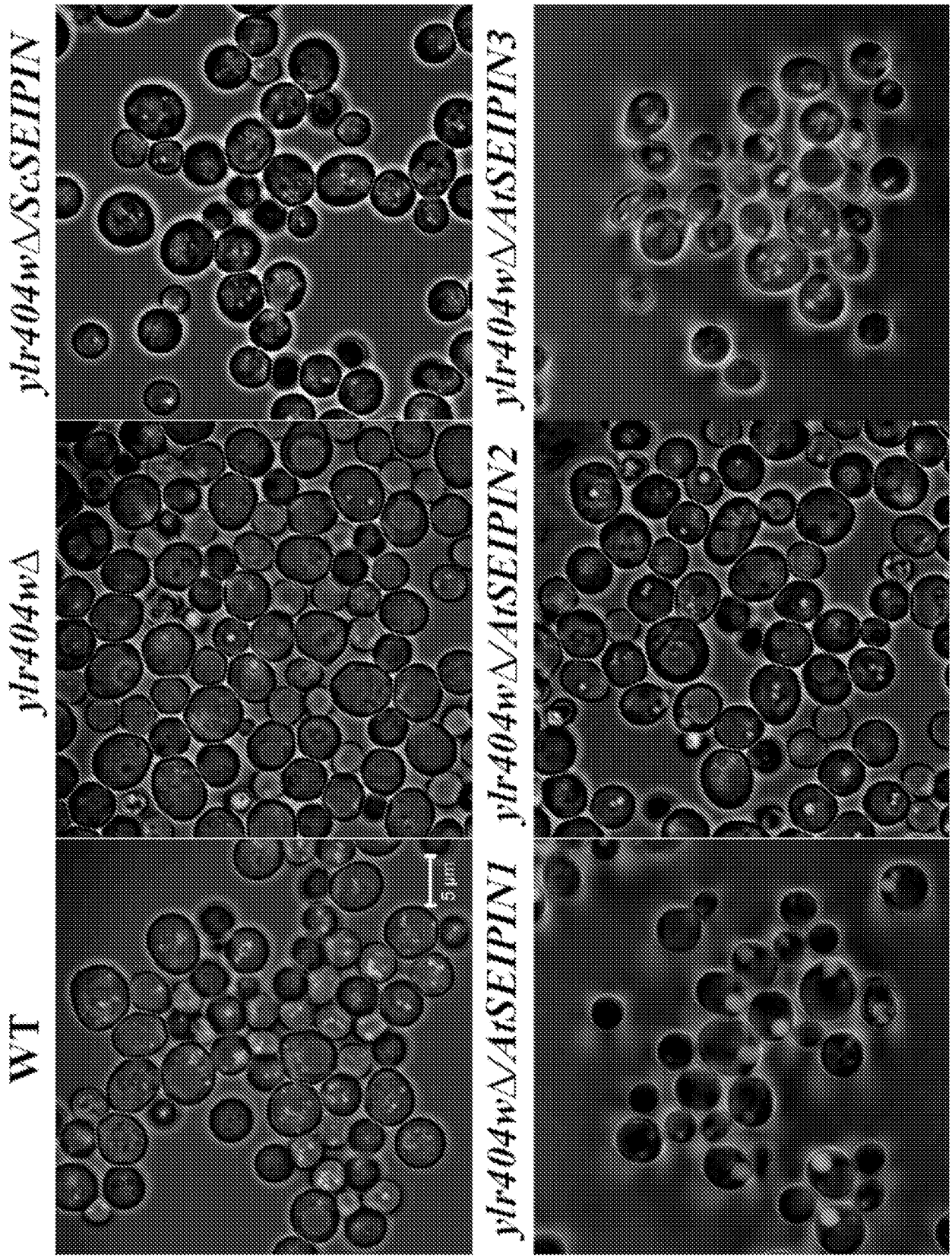

WT
ylr404wΔ
ylr404wΔ/ScSEIPIN
ylr404wΔ/AtSEIPIN1
ylr404wΔ/AtSEIPIN2
ylr404wΔ/AtSEIPIN3
5 μm

FIG. 13

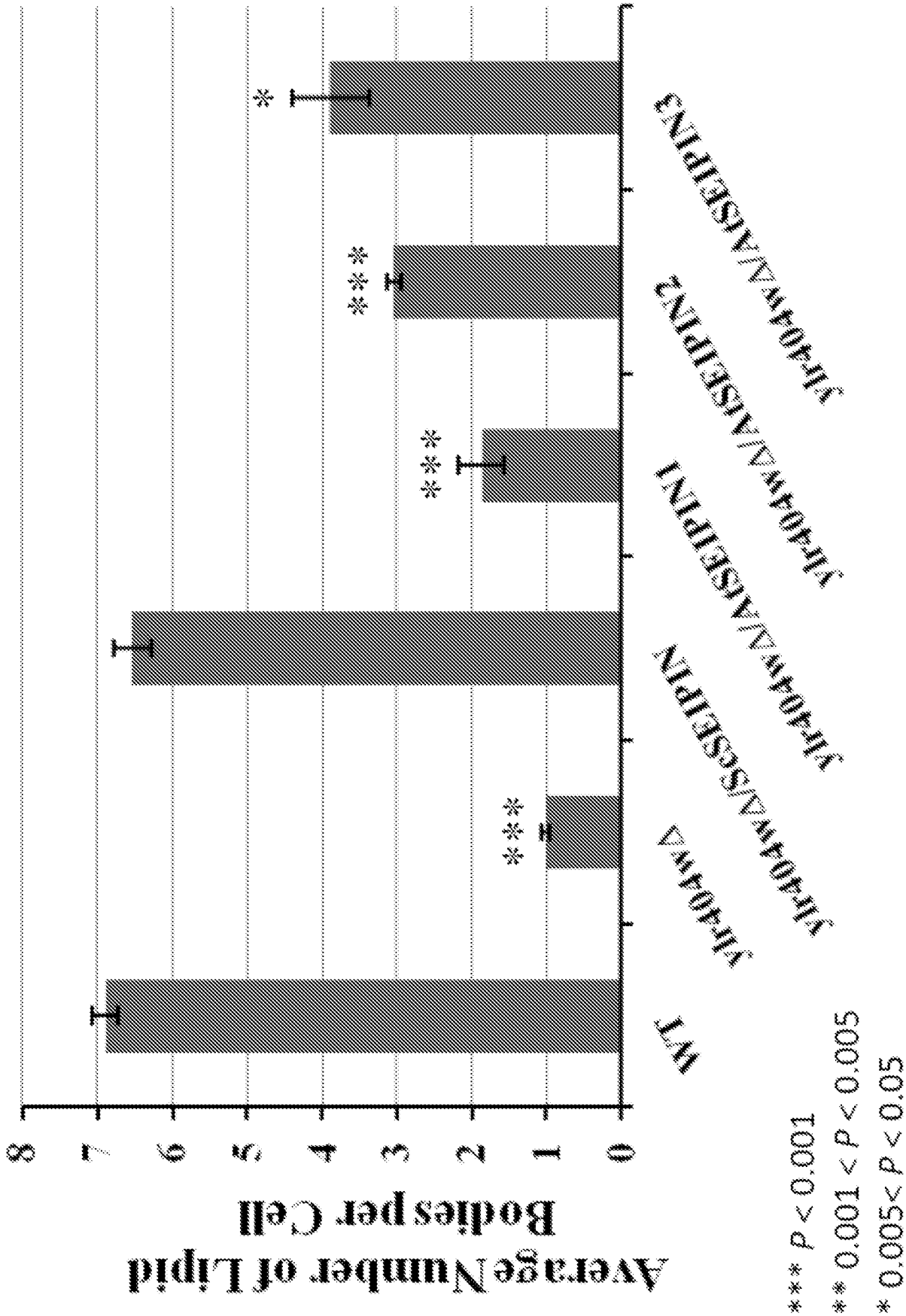

Average Number of Lipid Bodies per Cell
0
1
2
3
4
5
6
7
8
WT
yif40Δ
yif40Δ/SCSEIPIN
yif40Δ/ASEIPIN1
yif40Δ/ASEIPIN2
yif40Δ/ASEIPIN3
*** P < 0.001
** 0.001 < P < 0.005
* 0.005 < P < 0.05

FIG. 14

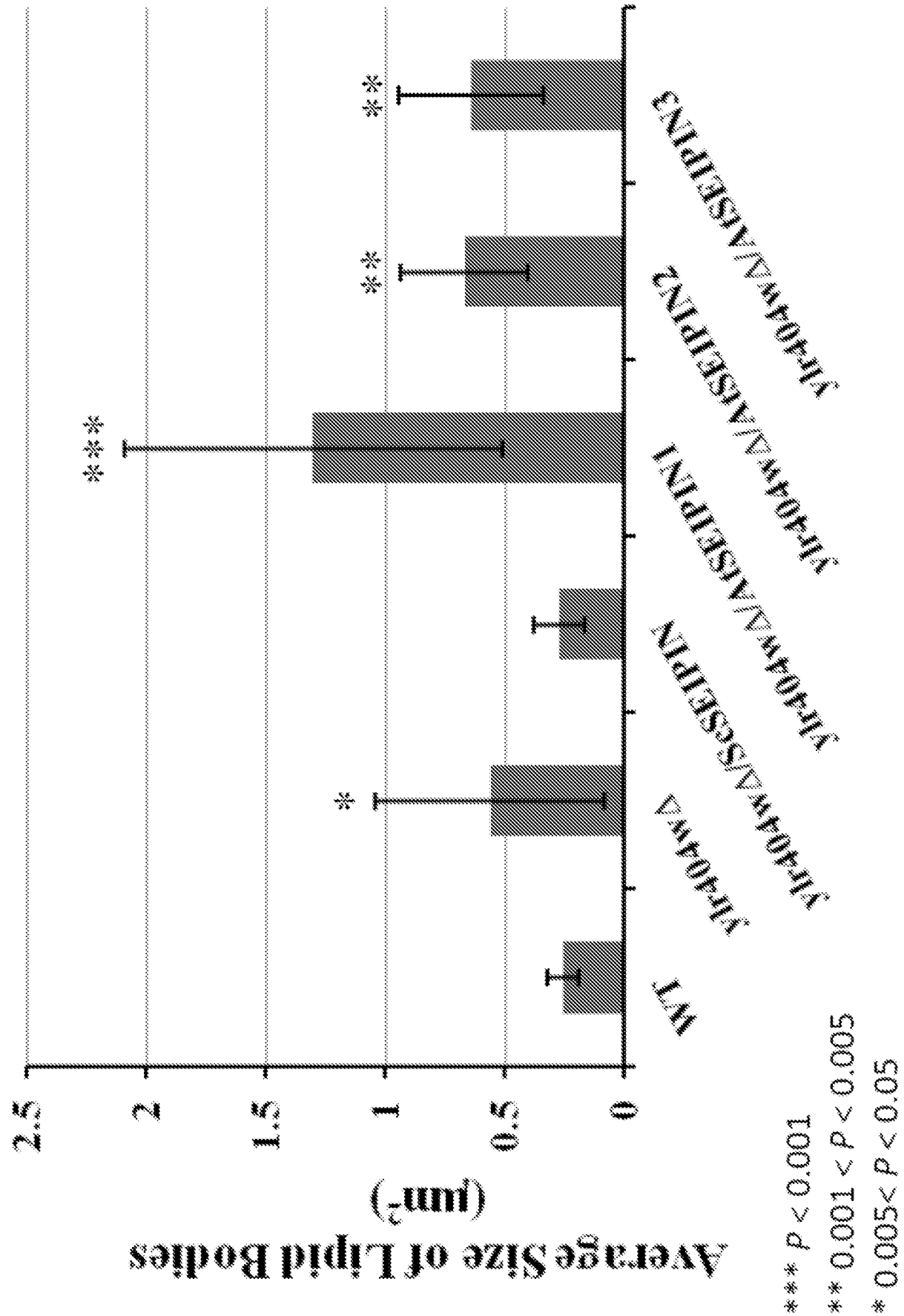

Average Size of Lipid Bodies
(μm²)
2.5
2
1.5
1
0.5
0
WT
YJF404WΔ
YJF404WΔ/ScSEIPIN
YJF404WΔ/ASEIPIN1
YJF404WΔ/ASEIPIN2
YJF404WΔ/ASEIPIN3
*
***
**
**
*** P < 0.001
** 0.001 < P < 0.005
* 0.005< P < 0.05

FIG. 15

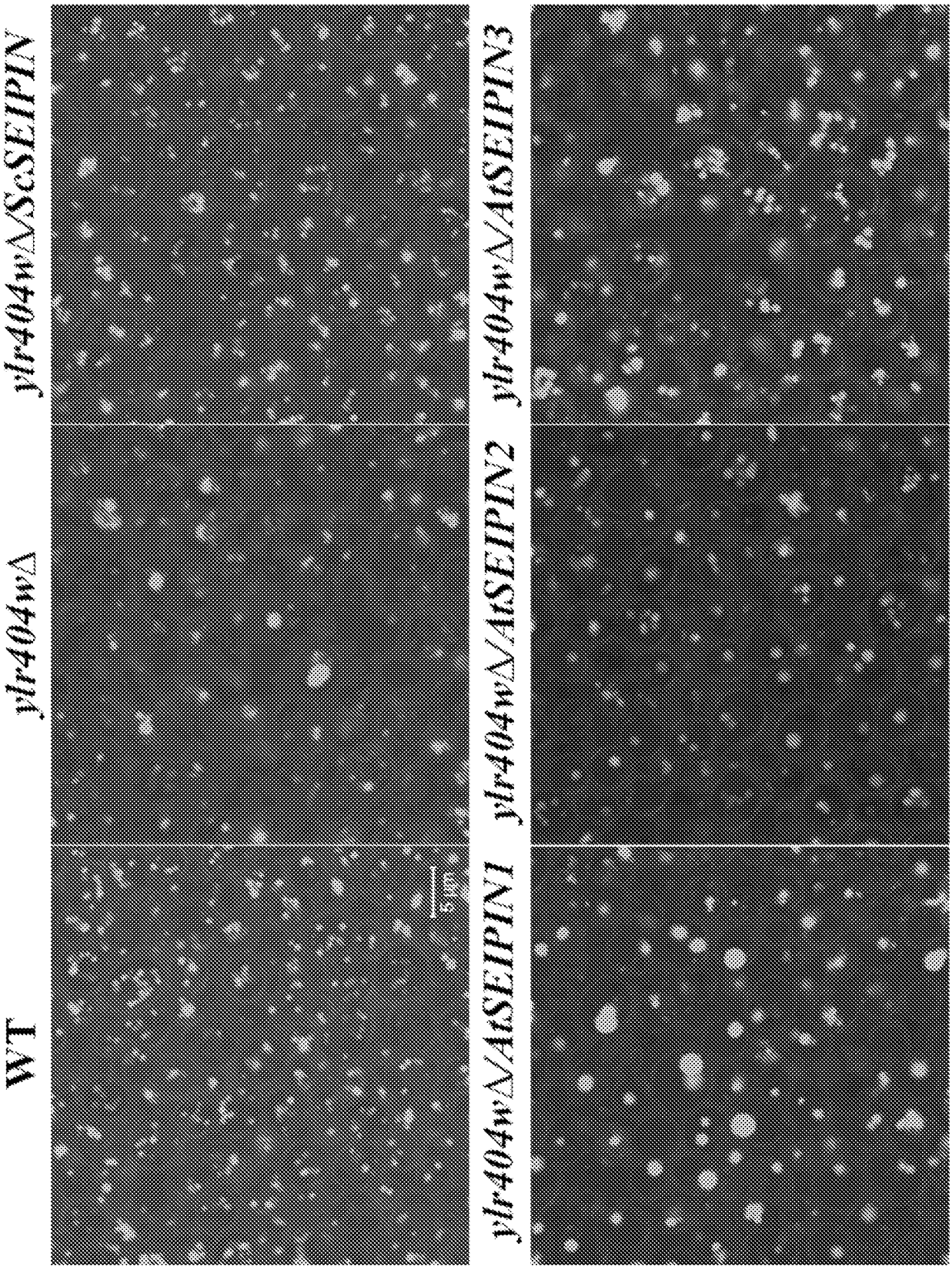

WT
ylr404wΔ
ylr404wΔ/ScSEIPIN
5 μm
ylr404wΔ/AtSEIPIN1
ylr404wΔ/AtSEIPIN2
ylr404wΔ/AtSEIPIN3

FIG. 16

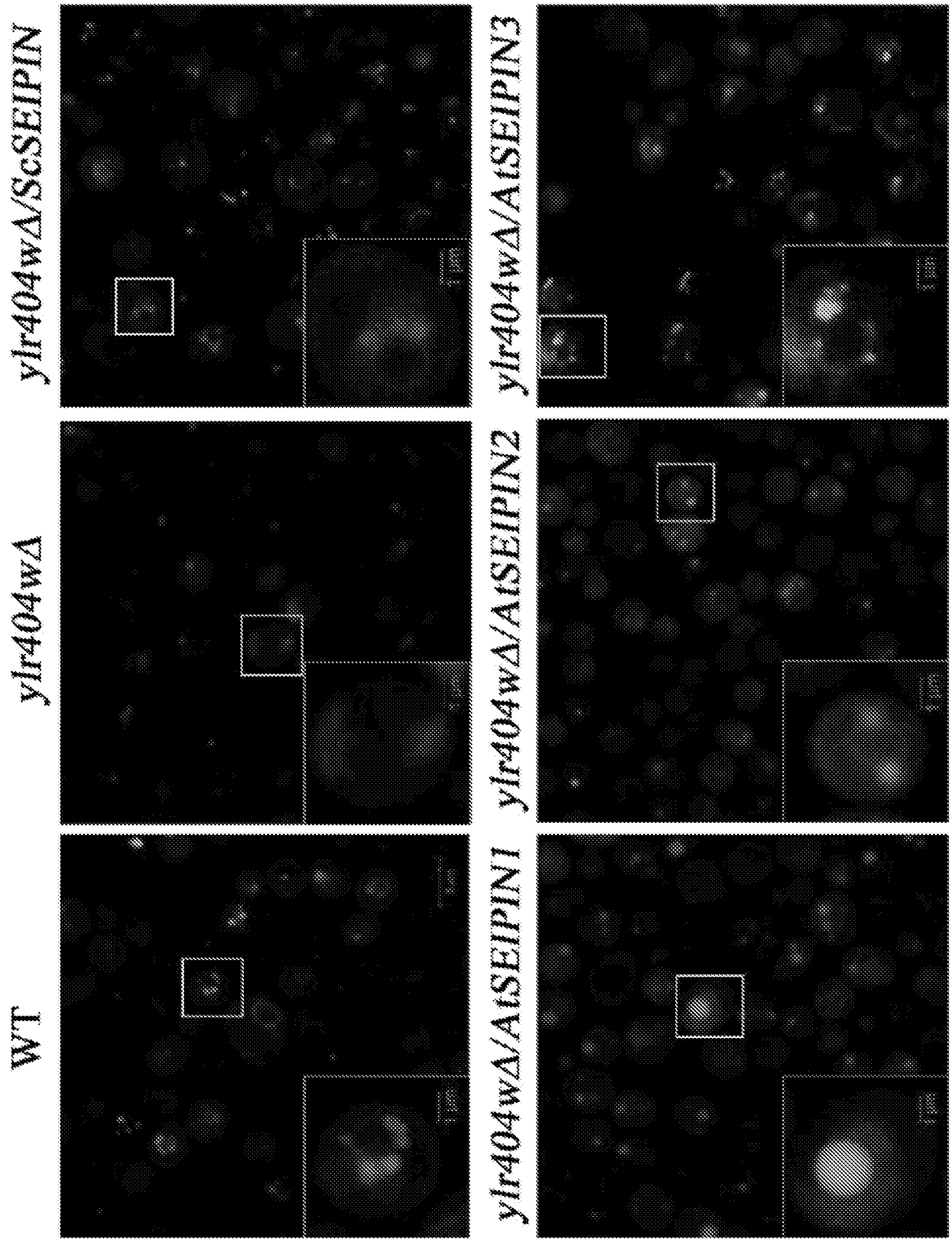

WT
ylr404wΔ
ylr404wΔ/ScSEIPIN
ylr404wΔ/AtSEIPIN1
ylr404wΔ/AtSEIPIN2
ylr404wΔ/AtSEIPIN3

FIG. 17

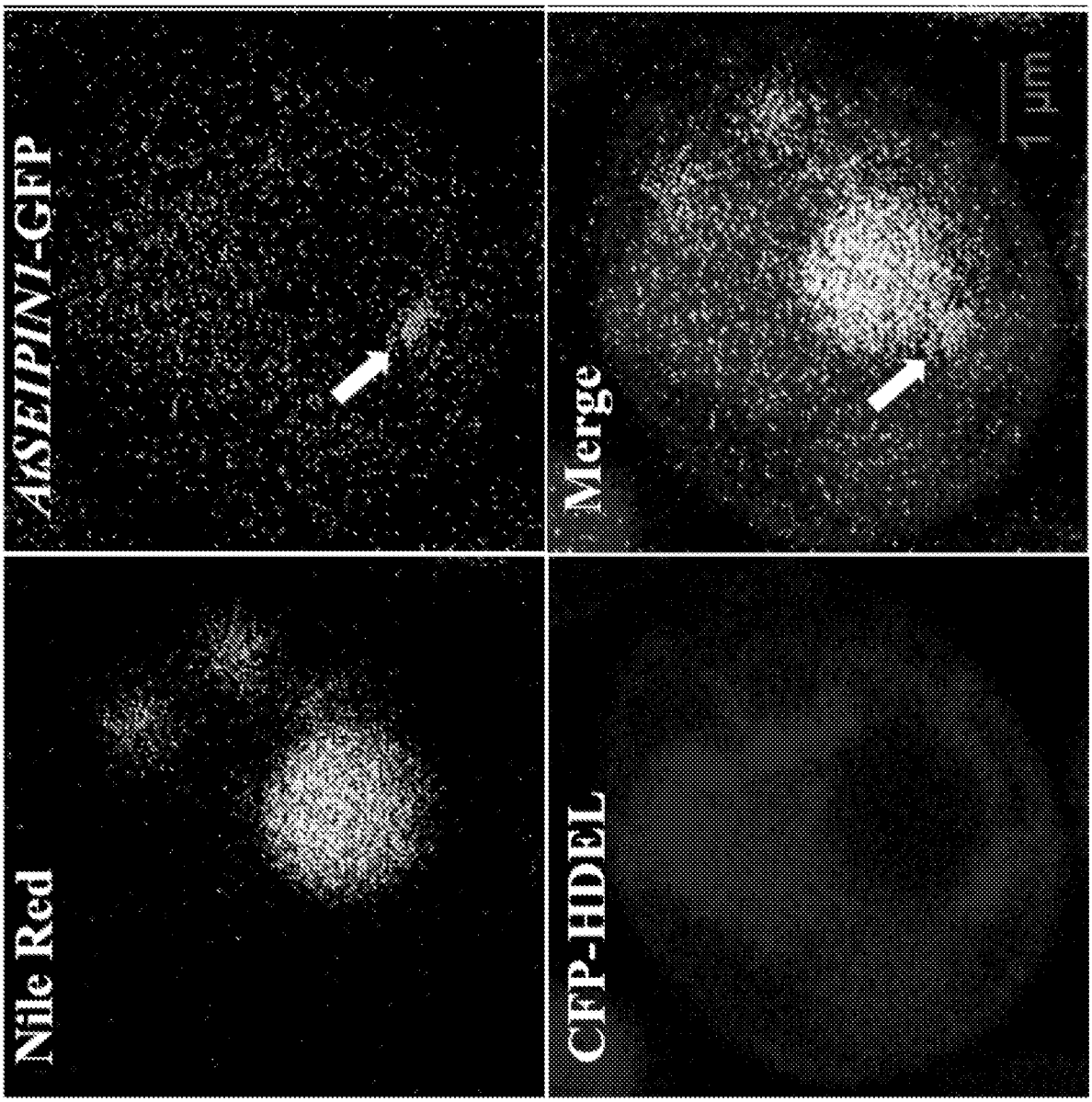


FIG. 18

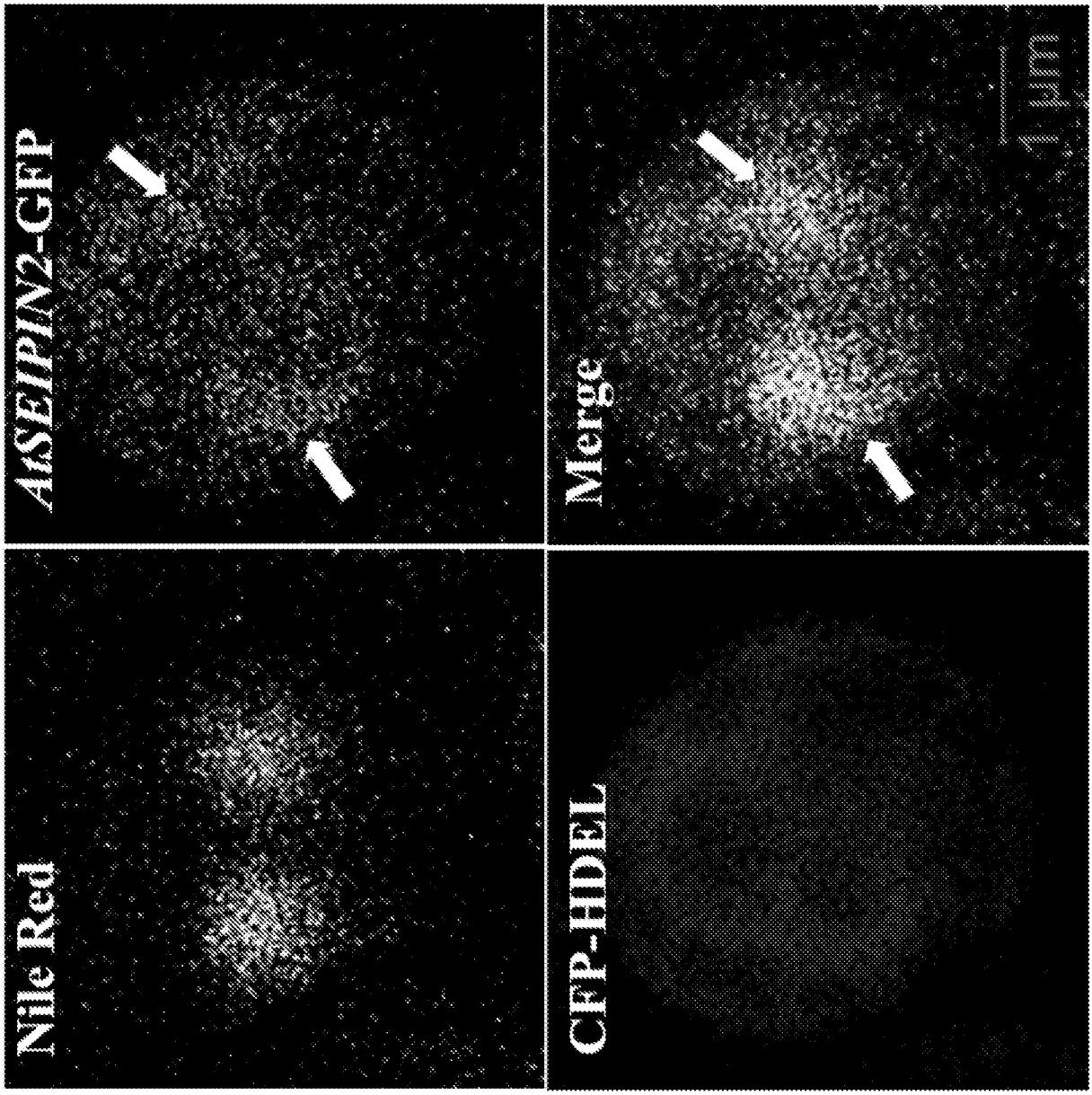


FIG. 19

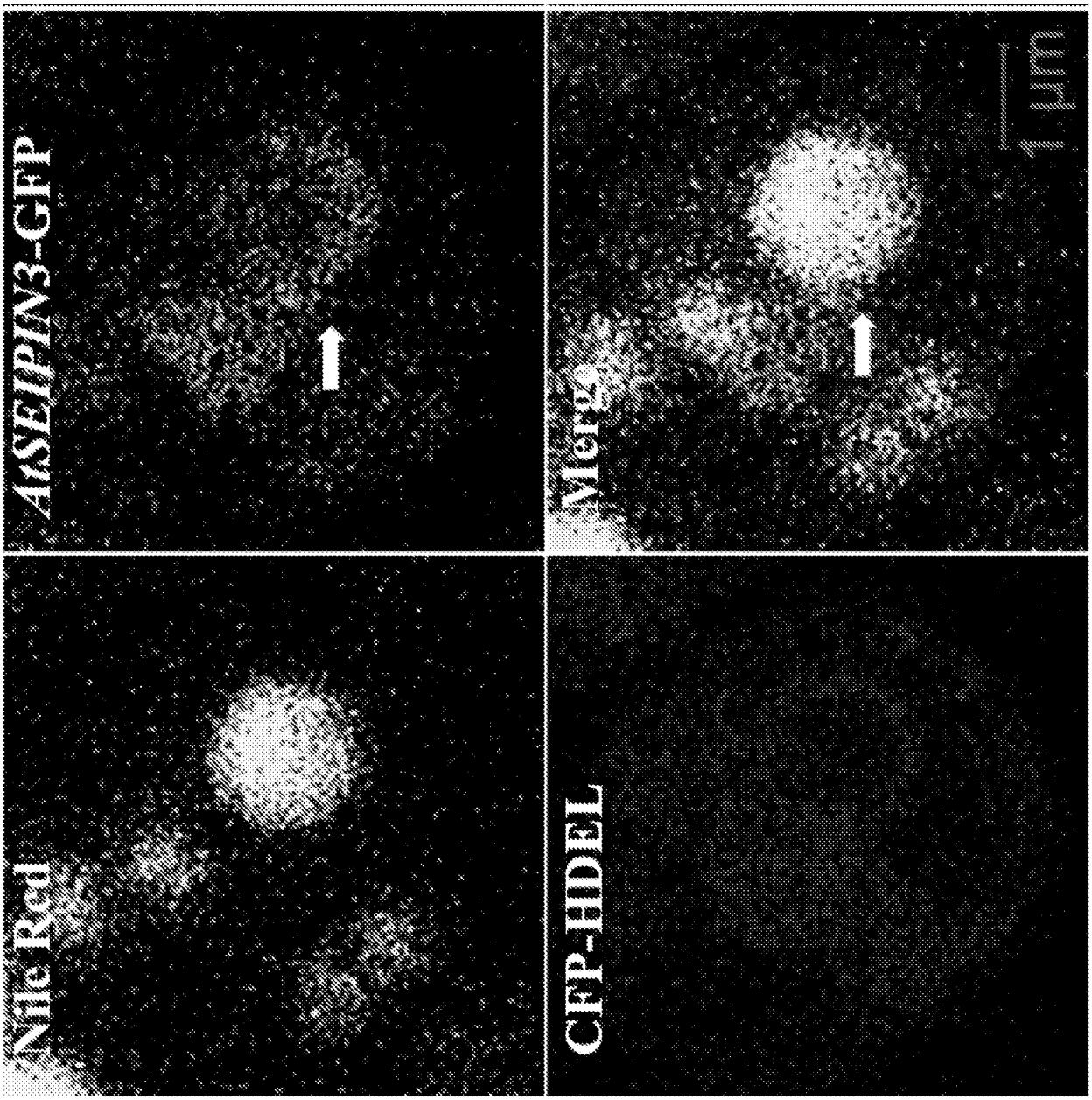


FIG. 20

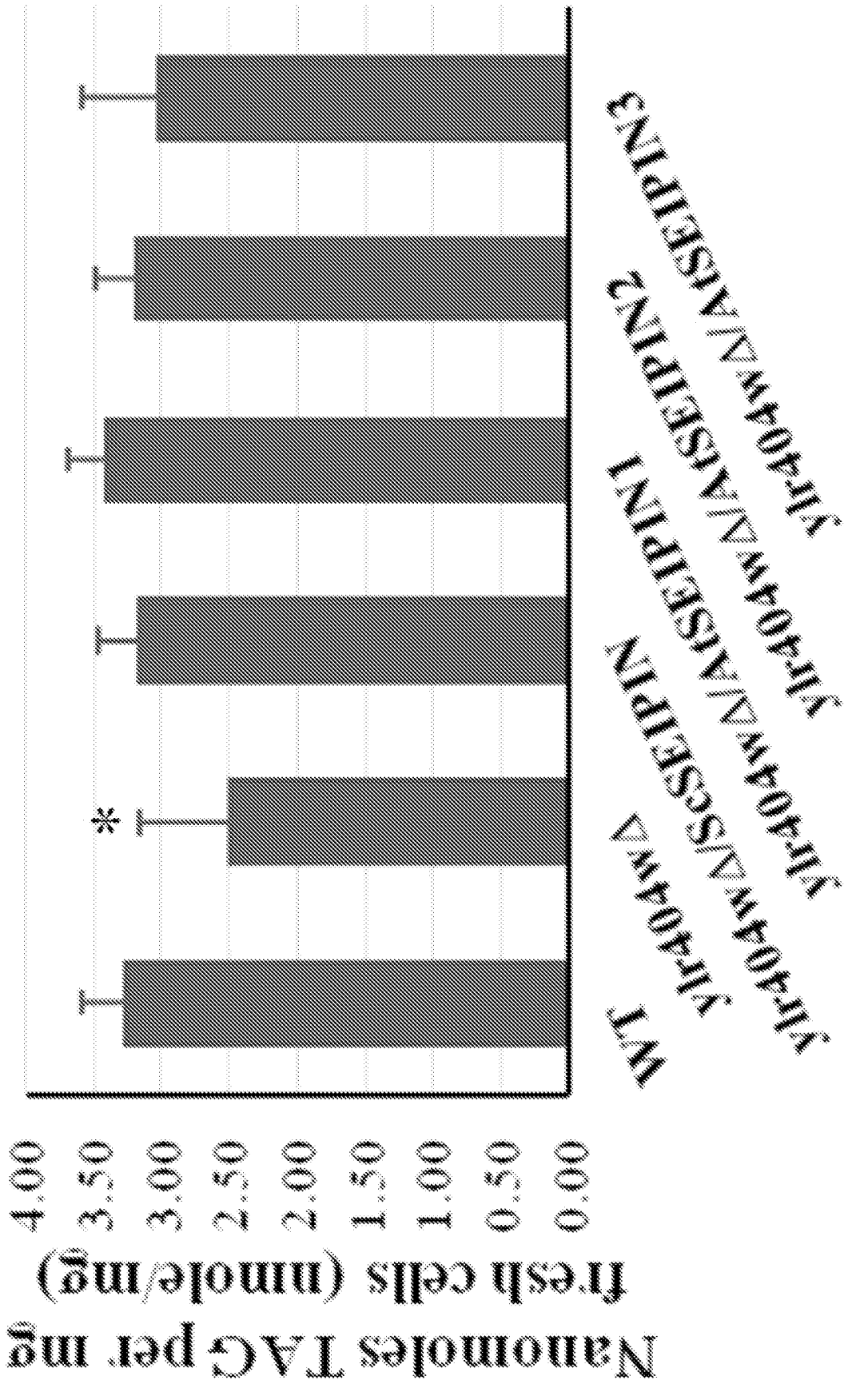

Nanomoles TAG per mg fresh cells (mole/mg)
4.00
3.50
3.00
2.50
2.00
1.50
1.00
0.50
0.00
*
WT
yl402w Δ
yl402w Δ/sCSF1PN
yl402w Δ/sCSF1PN1
yl402w Δ/sCSF1PN2
yl402w Δ/sCSF1PN3

FIG. 21

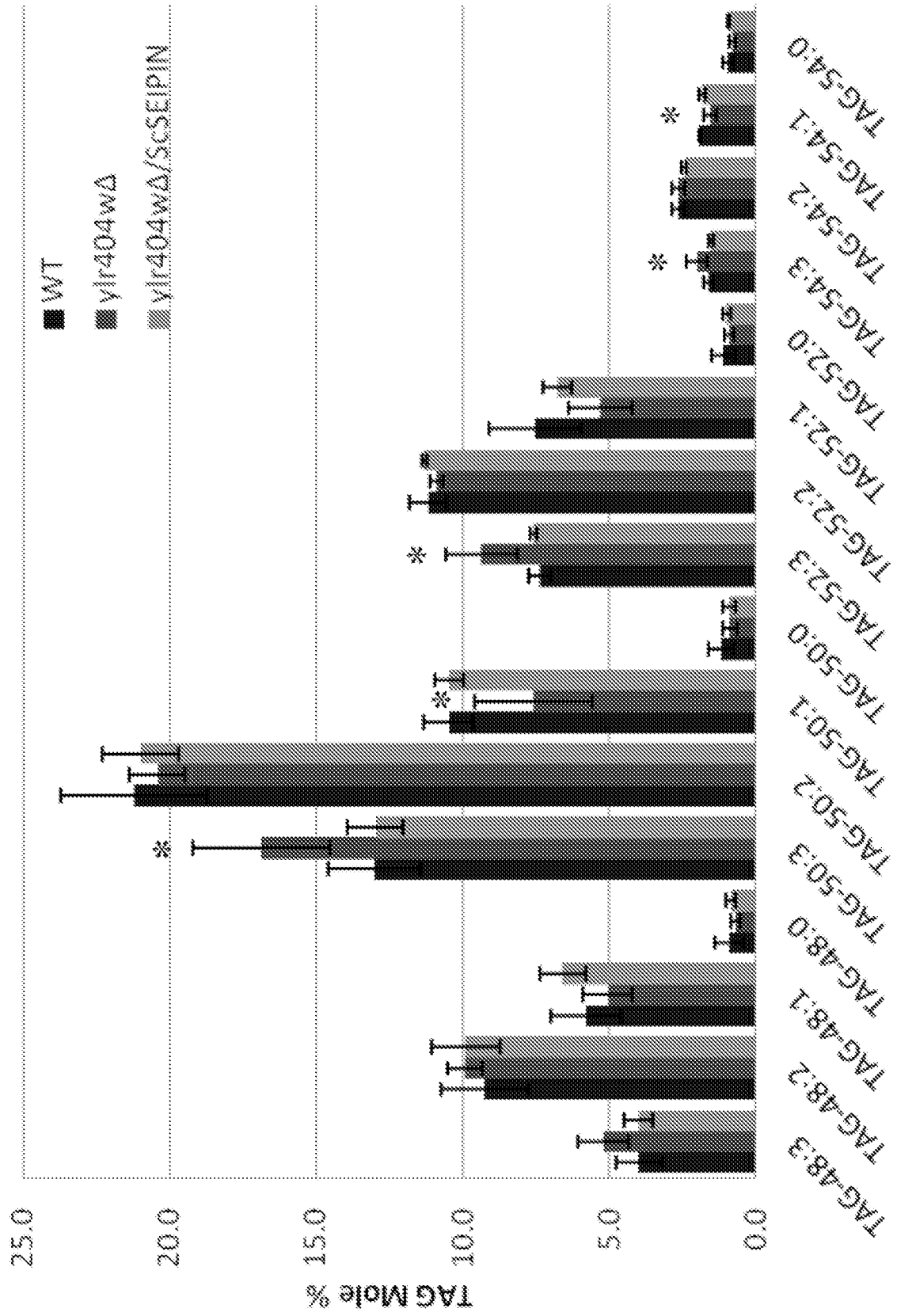

WT
Ylr404wΔ
Ylr404wΔ/ScSEIPIN
TAG Mole %
25.0
20.0
15.0
10.0
5.0
0.0
TAG-48:3
TAG-48:2
TAG-48:1
TAG-48:0
TAG-50:3
TAG-50:2
TAG-50:1
TAG-50:0
TAG-52:3
TAG-52:2
TAG-52:1
TAG-52:0
TAG-54:3
TAG-54:2
TAG-54:1
TAG-54:0
* P = 0.05

FIG. 22

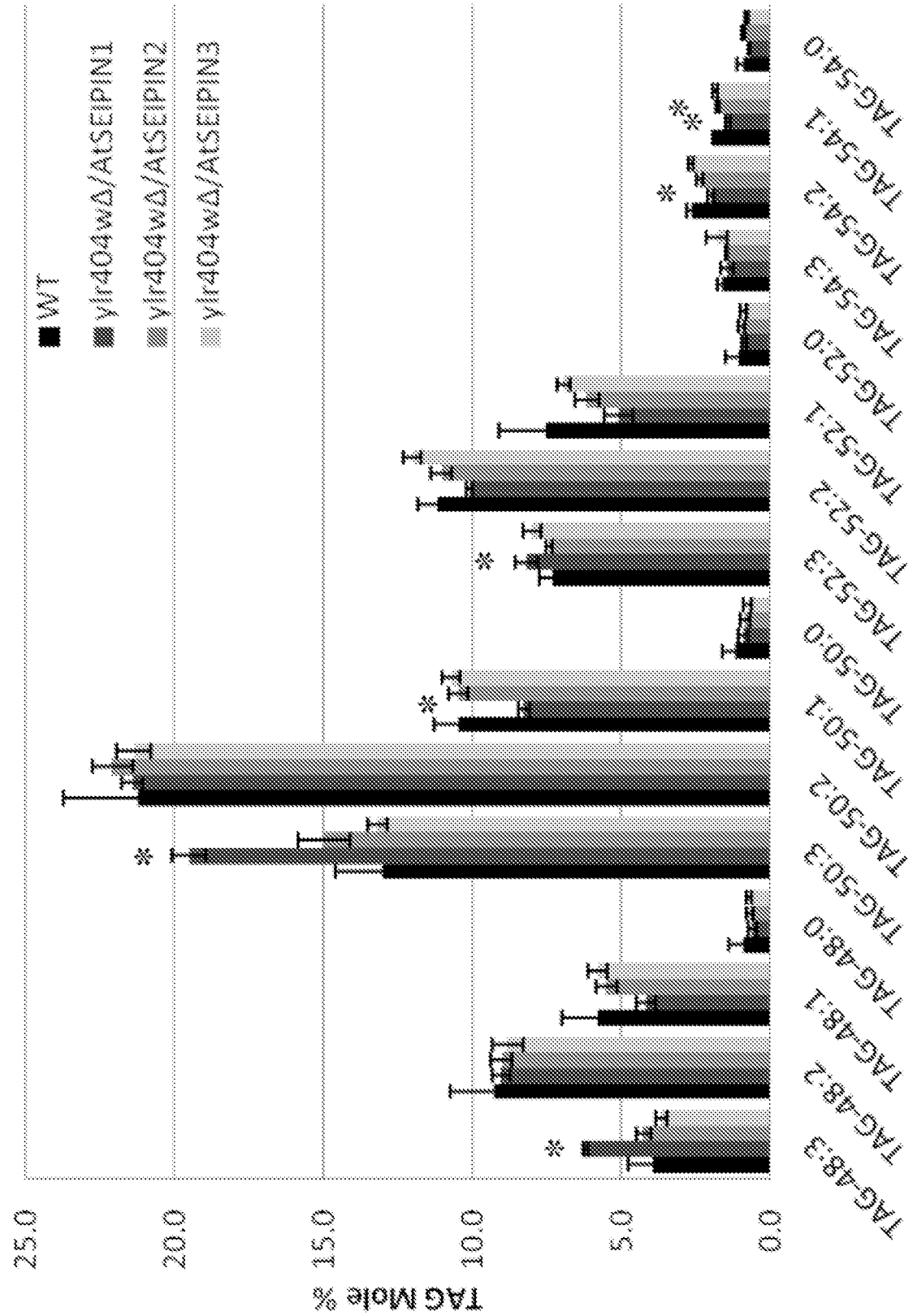

WT
ylr404wΔ/AtSEIPIN1
ylr404wΔ/AtSEIPIN2
ylr404wΔ/AtSEIPIN3
TAG Mole %
25.0
20.0
15.0
10.0
5.0
0.0
TAG-48:3
TAG-48:2
TAG-48:1
TAG-48:0
TAG-50:3
TAG-50:2
TAG-50:1
TAG-50:0
TAG-52:3
TAG-52:2
TAG-52:1
TAG-52:0
TAG-54:3
TAG-54:2
TAG-54:1
TAG-54:0
P = 0.05

FIG. 23

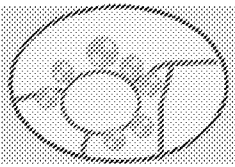
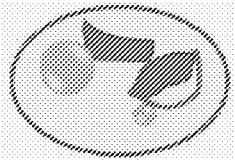
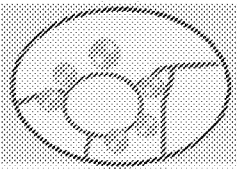
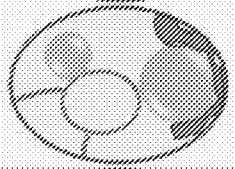
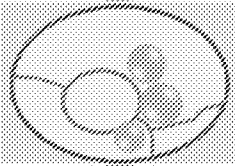
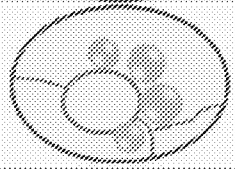
Yeast Strain	LD Phenotype		ER Organization	
Wild type	6.9 per cell	0.3 μm²		Normal ER
SEIPIN-deletion mutant	1.0 per cell	0.5 μm²		Aggregated ER overlapped with neutral lipid
ScSEIPIN	6.5 per cell	0.3 μm²		Normal ER
AtSEIPIN1	Not numbers 1.9 per cell	↑ Size 1.3 μm²		Both normal and aggregated ER
AtSEIPIN2	↑ Numbers 3.0 per cell	Not size 0.6 μm²		Normal ER
AtSEIPIN3	↑ Numbers 3.9 per cell	Not size 0.6 μm²		Normal ER

FIG. 24

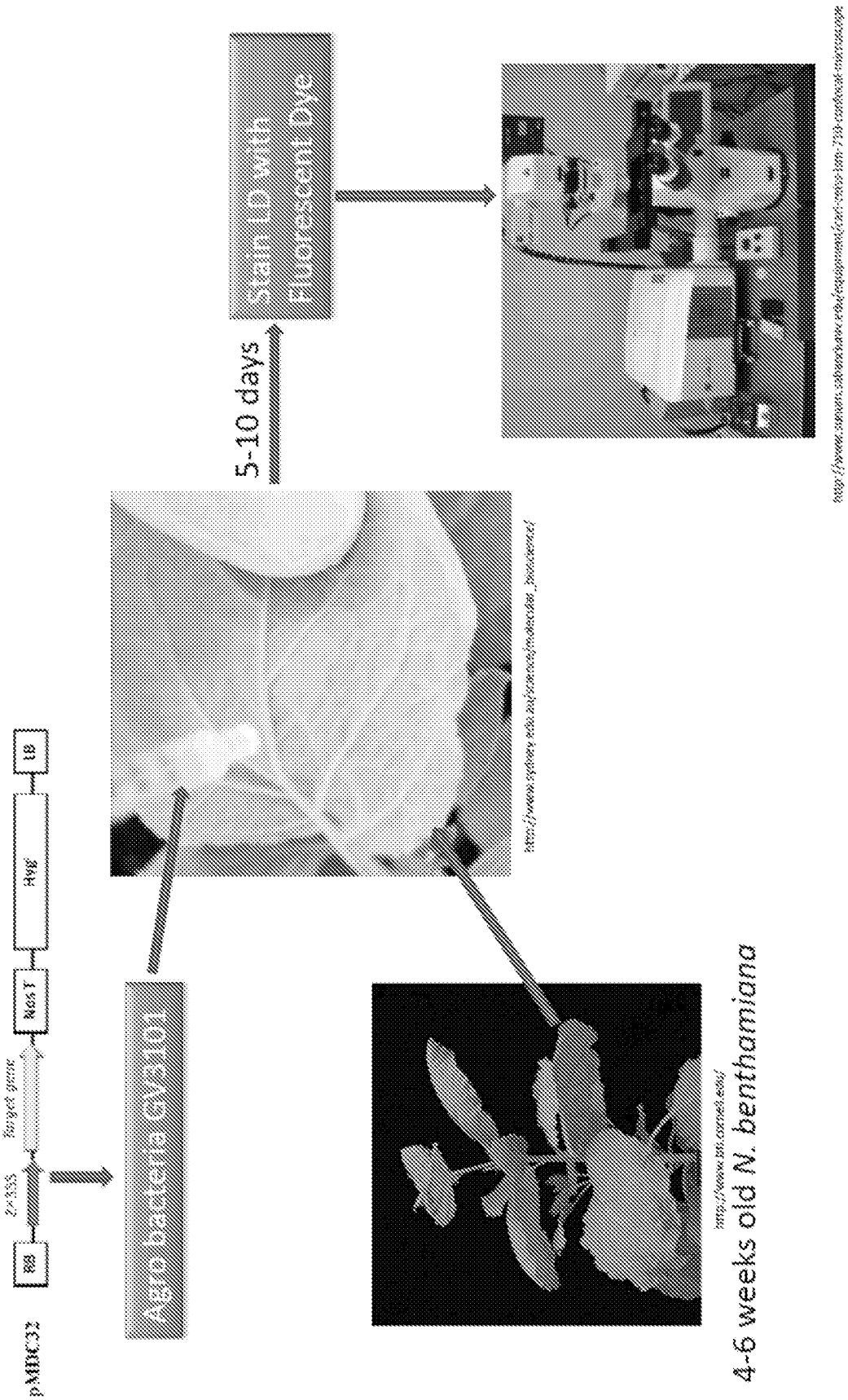

pMBC32
RB
2×35S
Target gene
NosT
Hyg
tB
Agro bacteria GV3101
4-6 weeks old N. benthamiana
5-10 days
Stain LD with Fluorescent Dye

FIG. 25

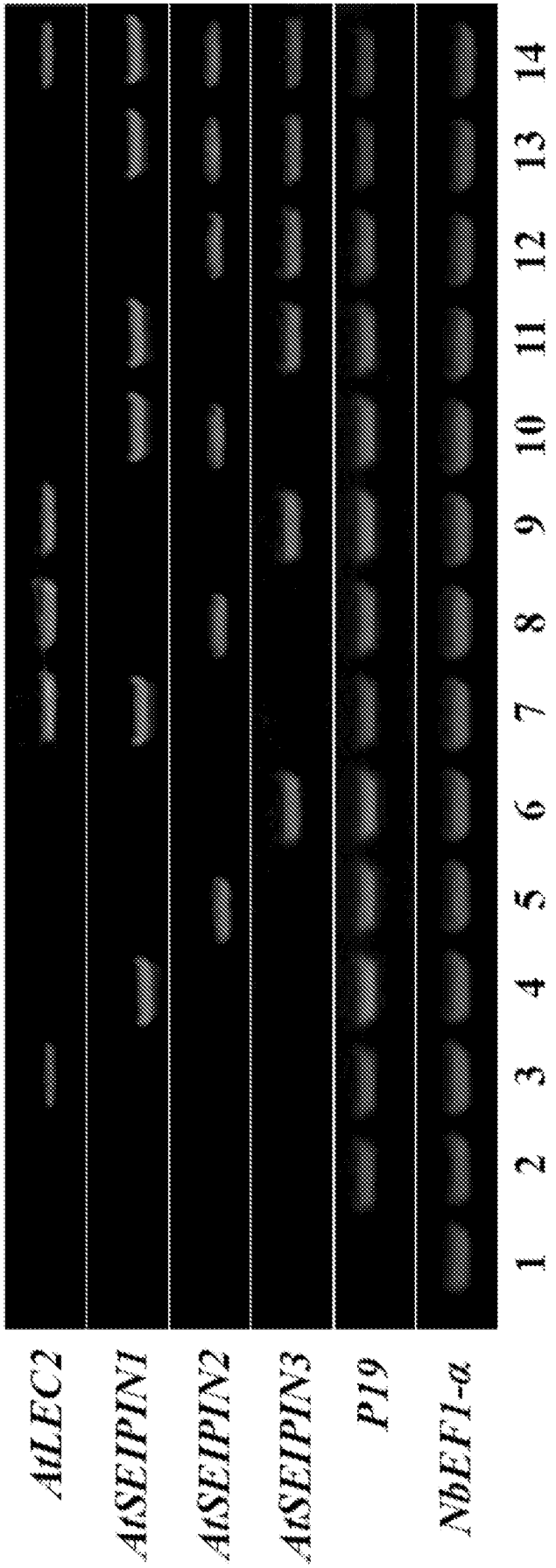
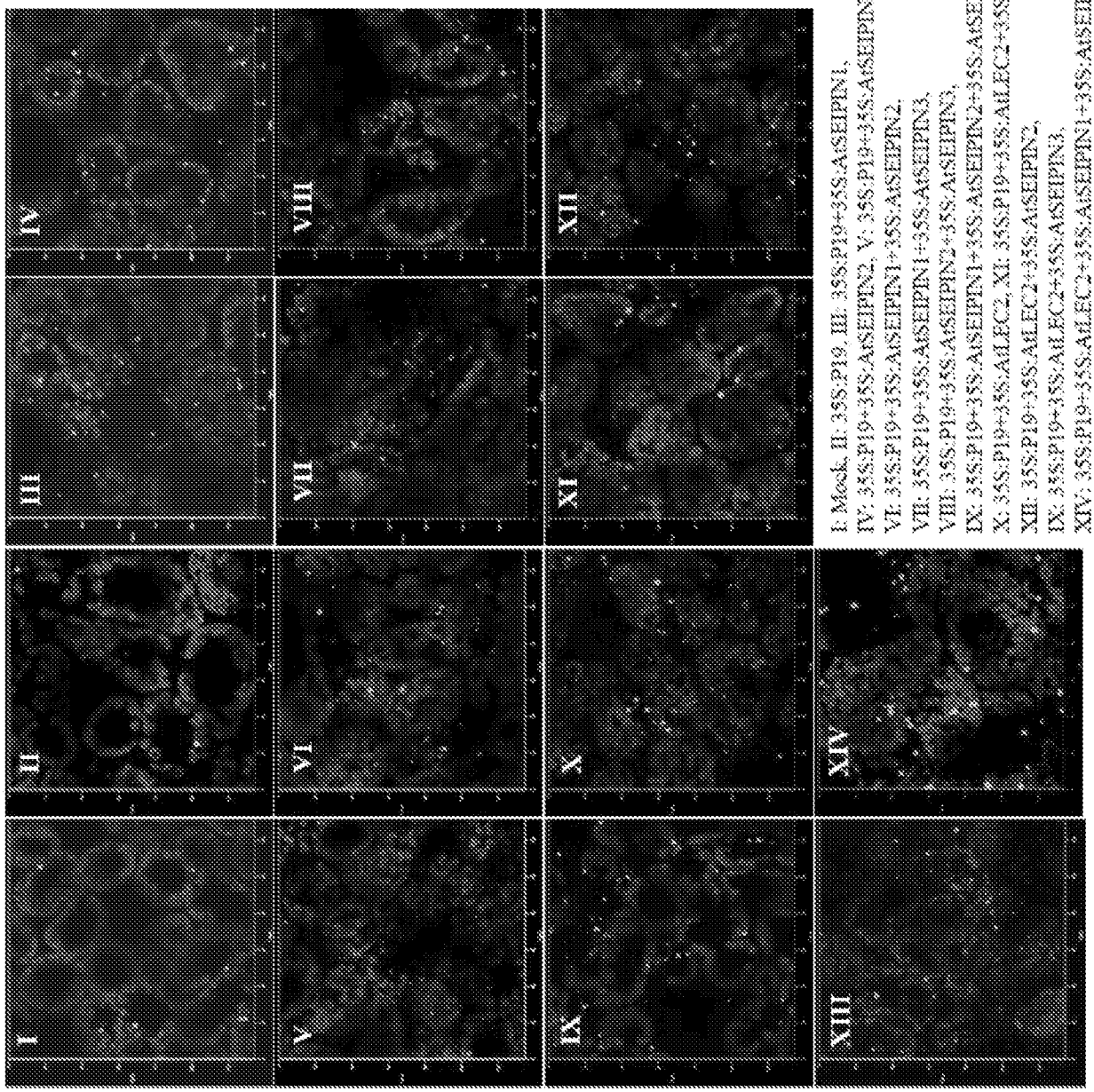

AtLEC2
AtSEIPIN1
AtSEIPIN2
AtSEIPIN3
P19
NbEF1-α
1 2 3 4 5 6 7 8 9 10 11 12 13 14

FIG. 26



I: Mock, II: 35S:P19, III: 35S:P19+35S:AtSEIPIN1,
IV: 35S:P19+35S:AtSEIPIN2, V: 35S:P19+35S:AtSEIPIN3,
VI: 35S:P19+35S:AtSEIPIN1+35S:AtSEIPIN2,
VII: 35S:P19+35S:AtSEIPIN1+35S:AtSEIPIN3,
VIII: 35S:P19+35S:AtSEIPIN2+35S:AtSEIPIN3,
IX: 35S:P19+35S:AtSEIPIN1+35S:AtSEIPIN2+35S:AtSEIPIN3,
X: 35S:P19+35S:AtLEC2, XI: 35S:P19+35S:AtLEC2+35S:AtSEIPIN1,
XII: 35S:P19+35S:AtLEC2+35S:AtSEIPIN2,
IX: 35S:P19+35S:AtLEC2+35S:AtSEIPIN3,
XIV: 35S:P19+35S:AtLEC2+35S:AtSEIPIN1+35S:AtSEIPIN2+35S:AtSEIPIN3

FIG. 27

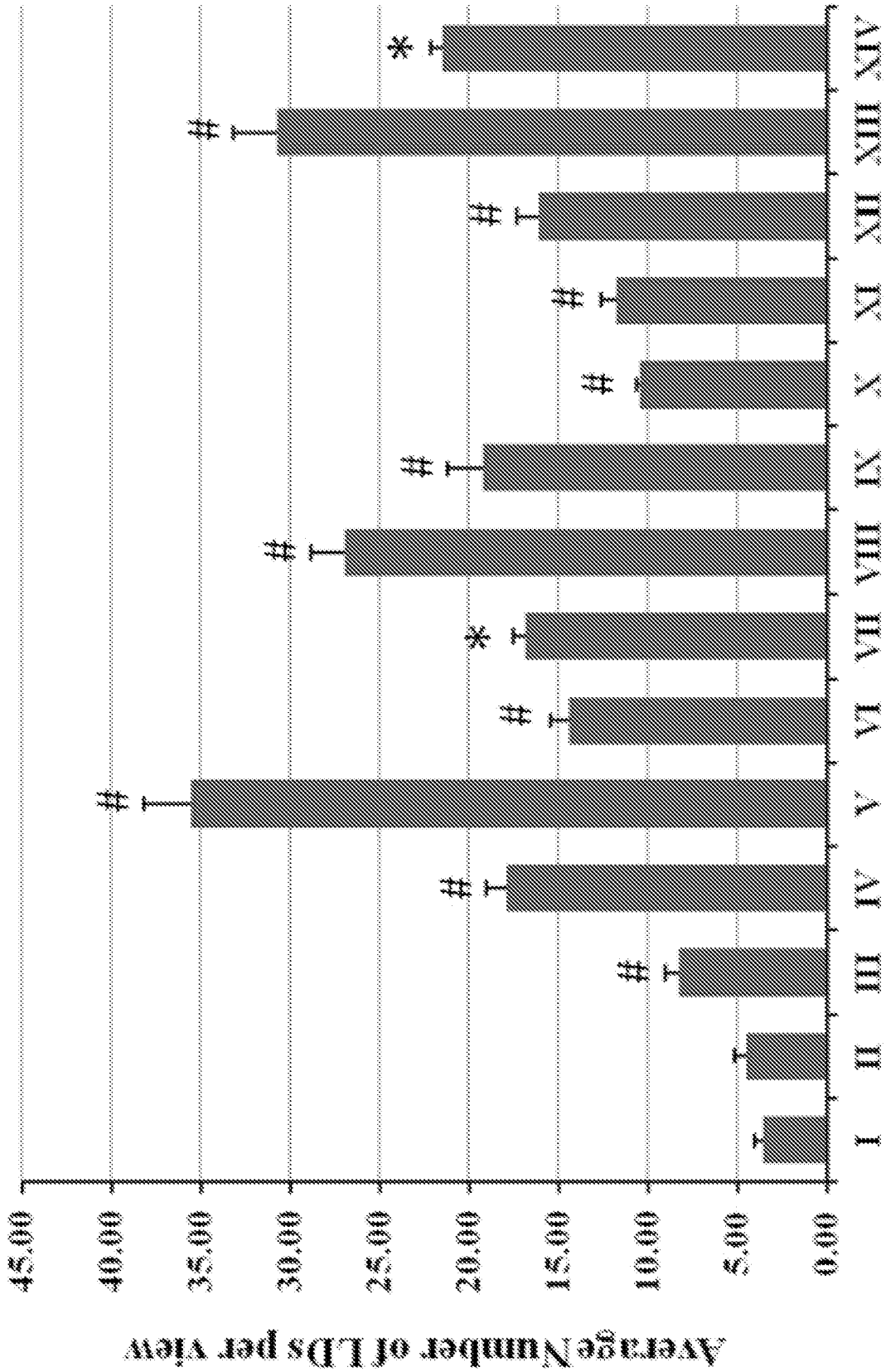

Average Number of LDs per view
45.00
40.00
35.00
30.00
25.00
20.00
15.00
10.00
5.00
0.00
I
II
III
IV
V
VI
VII
VIII
IX
X
XI
XII
XIII
XIV

FIG. 28

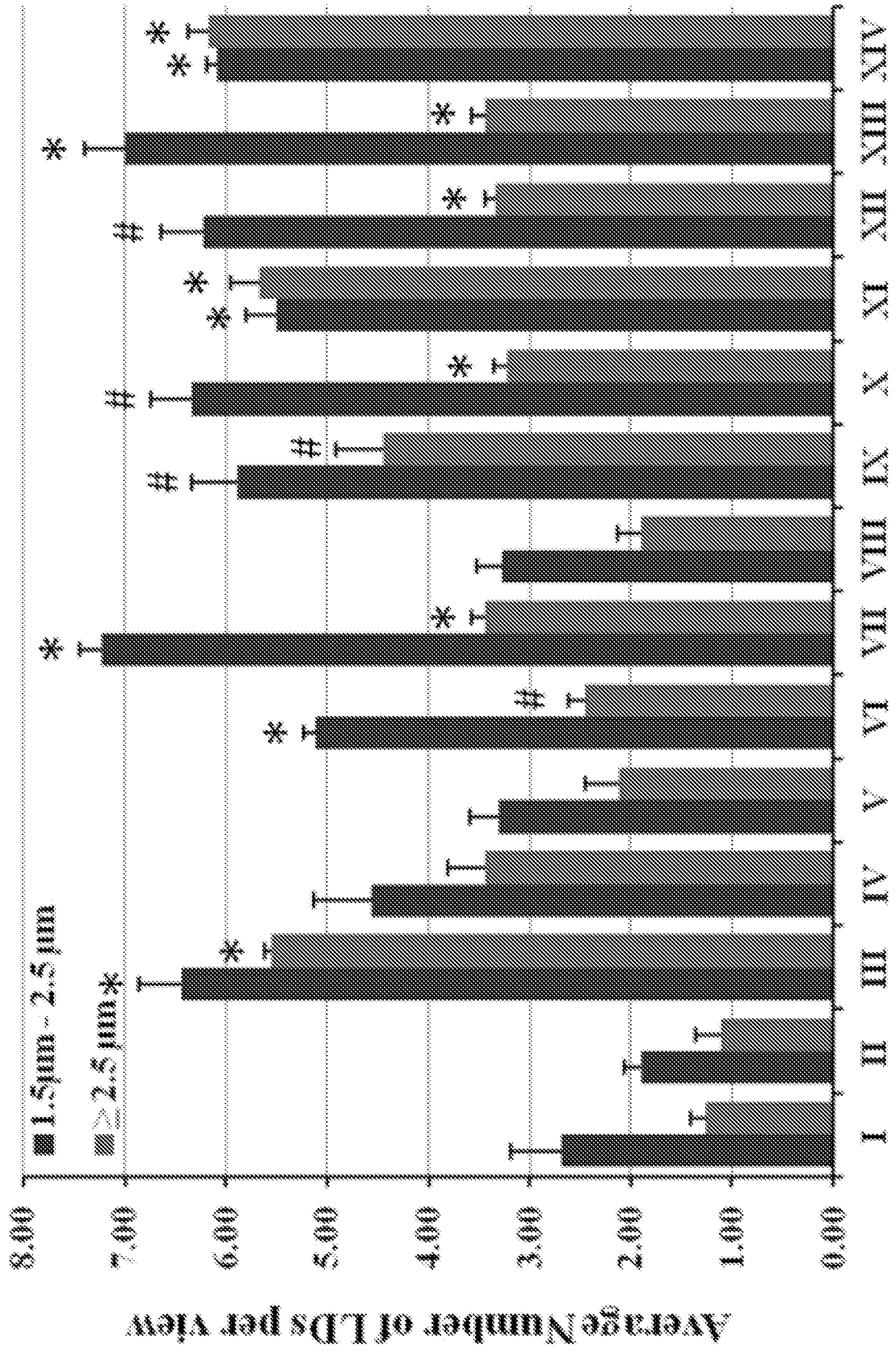

Average Number of LDs per view
8.00
7.00
6.00
5.00
4.00
3.00
2.00
1.00
0.00
1.5μm - 2.5 μm
≥ 2.5 μm
I
II
III
IV
V
VI
VII
VIII
XI
X
IX
XII
XIII
XIV

FIG. 29

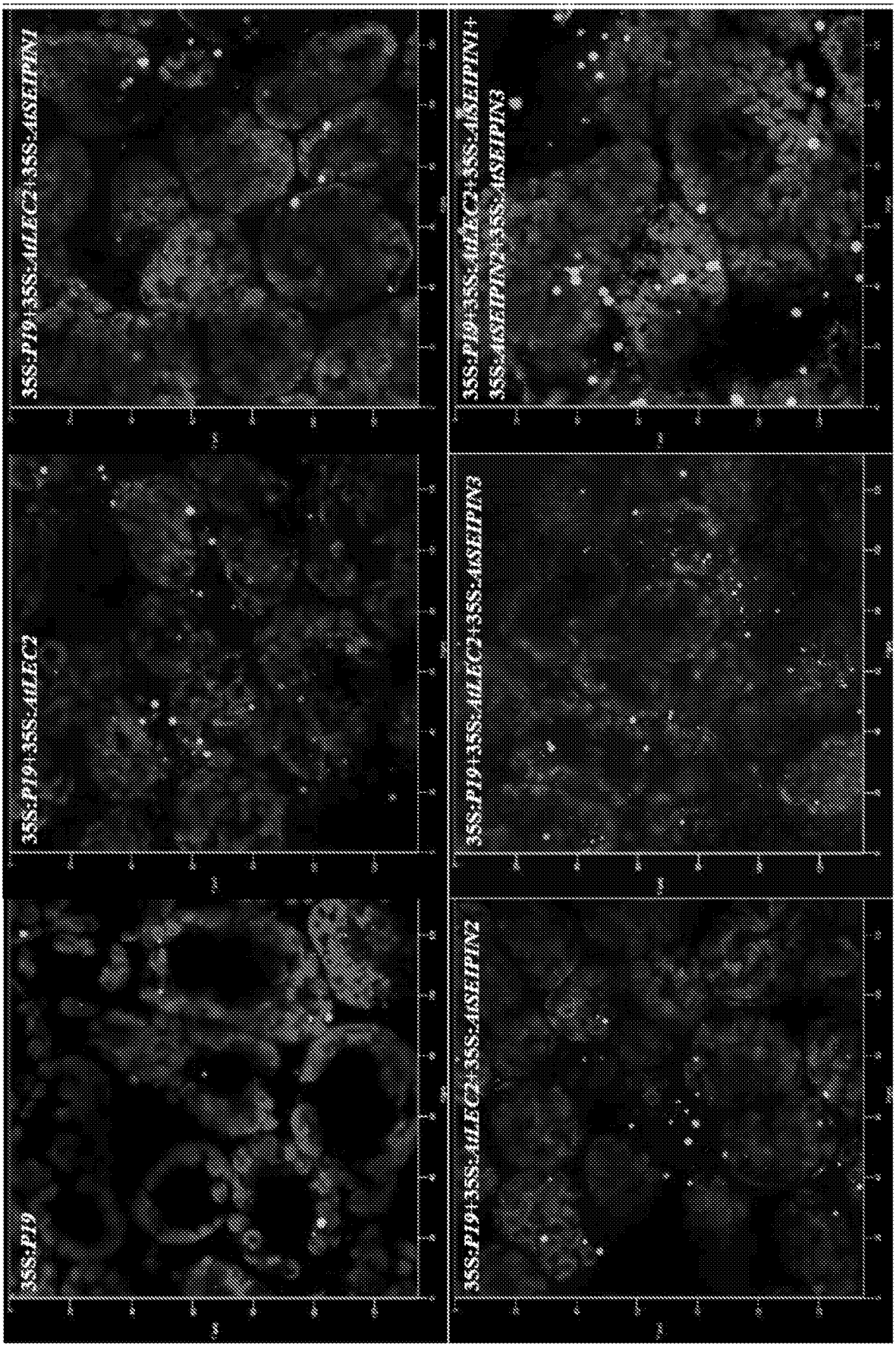

35S-P19
35S-P19+35S-AtLEC2
35S-P19+35S-AtLEC2+35S-AtSEIPIN1
35S-P19+35S-AtLEC2+35S-AtSEIPIN2
35S-P19+35S-AtLEC2+35S-AtSEIPIN3
35S-P19+35S-AtLEC2+35S-AtSEIPIN1+
35S-AtSEIPIN2+35S-AtSEIPIN3

FIG. 30

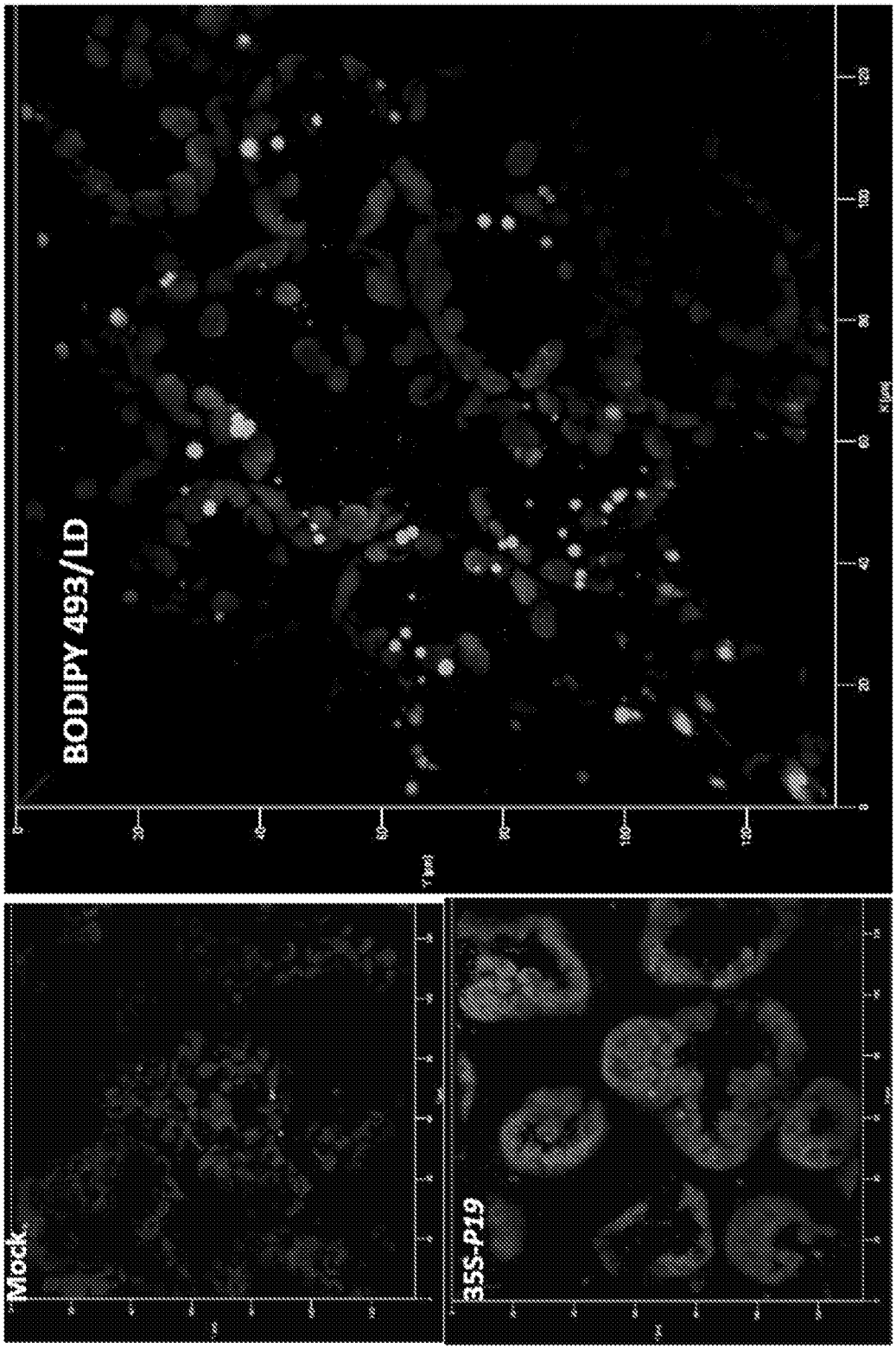


FIG. 31

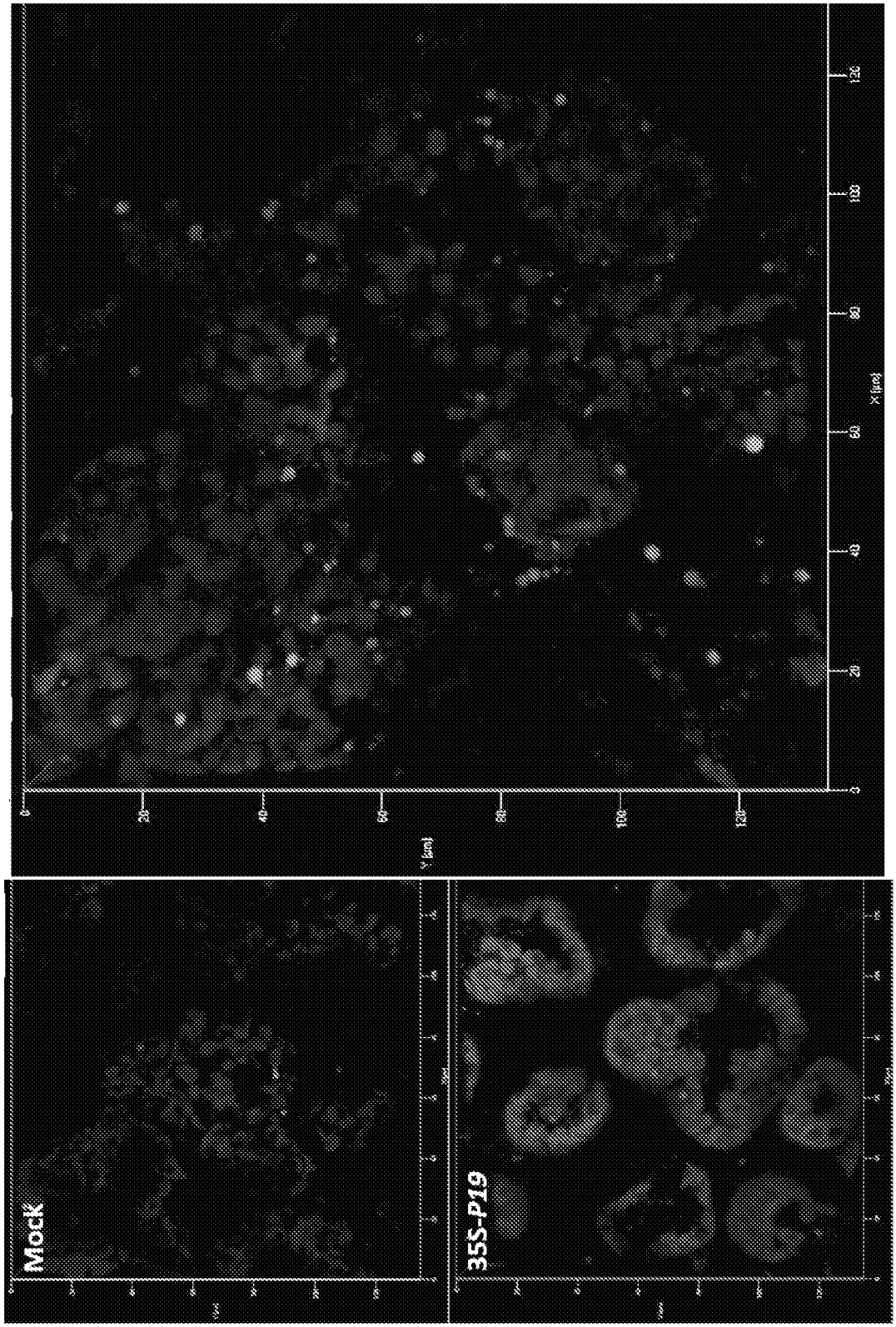


FIG. 32

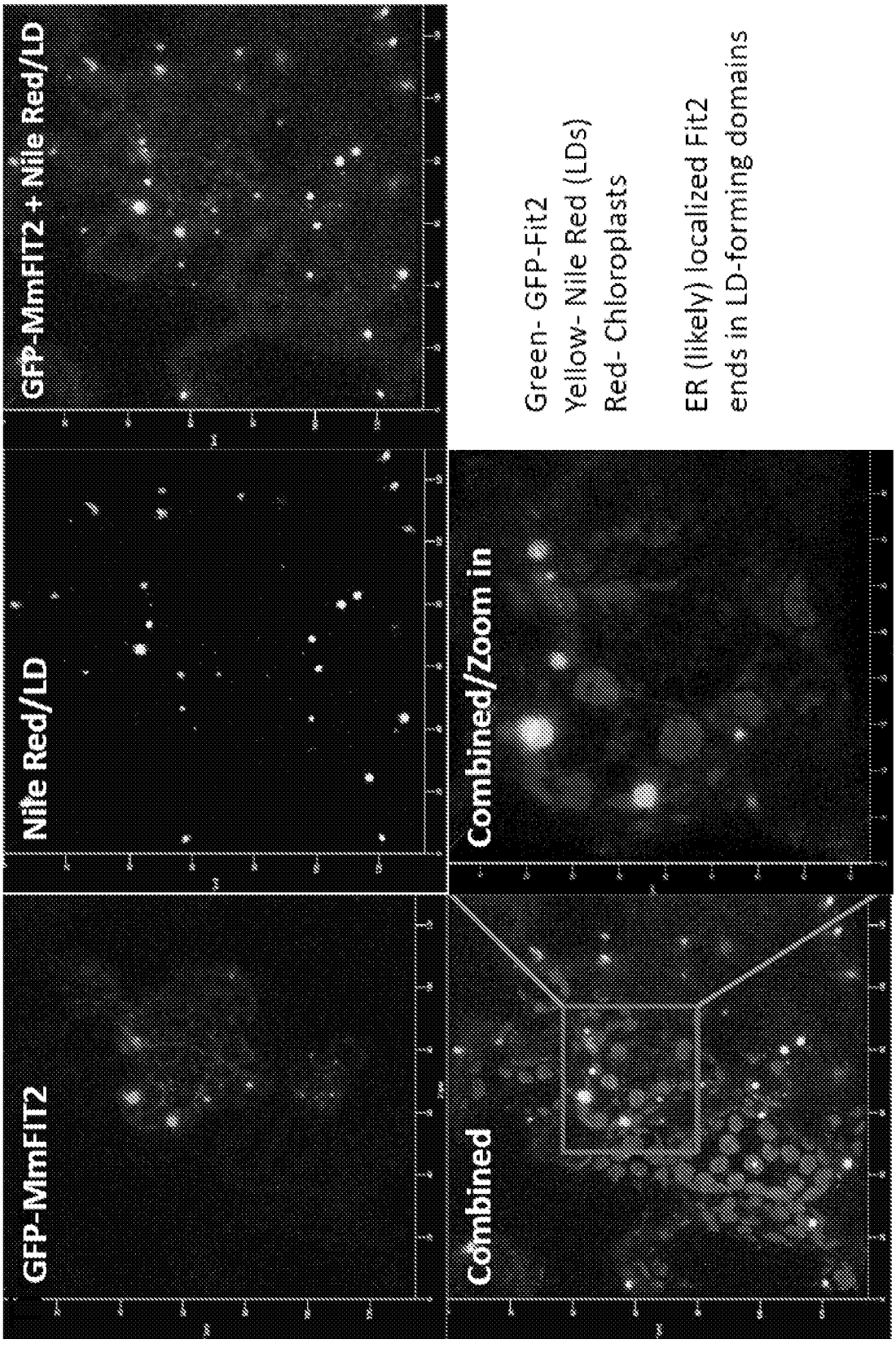

GFP-MmFIT2
Nile Red/LD
GFP-MmFIT2 + Nile Red/LD
Combined
Combined/Zoom in
Green- GFP-Fit2
Yellow- Nile Red (LDs)
Red- Chloroplasts
ER (likely) localized Fit2
ends in LD-forming domains

FIG. 33

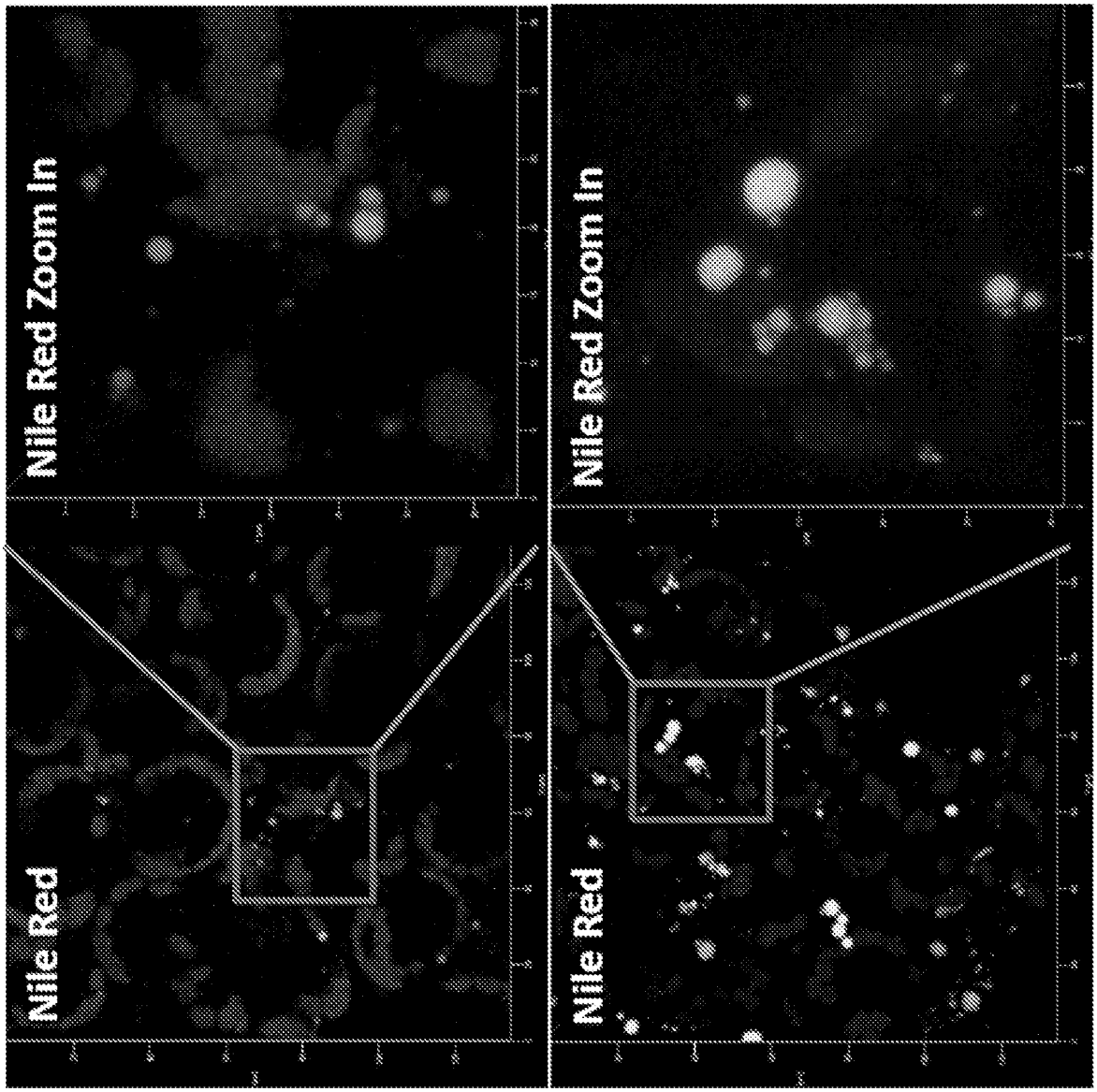


FIG. 34

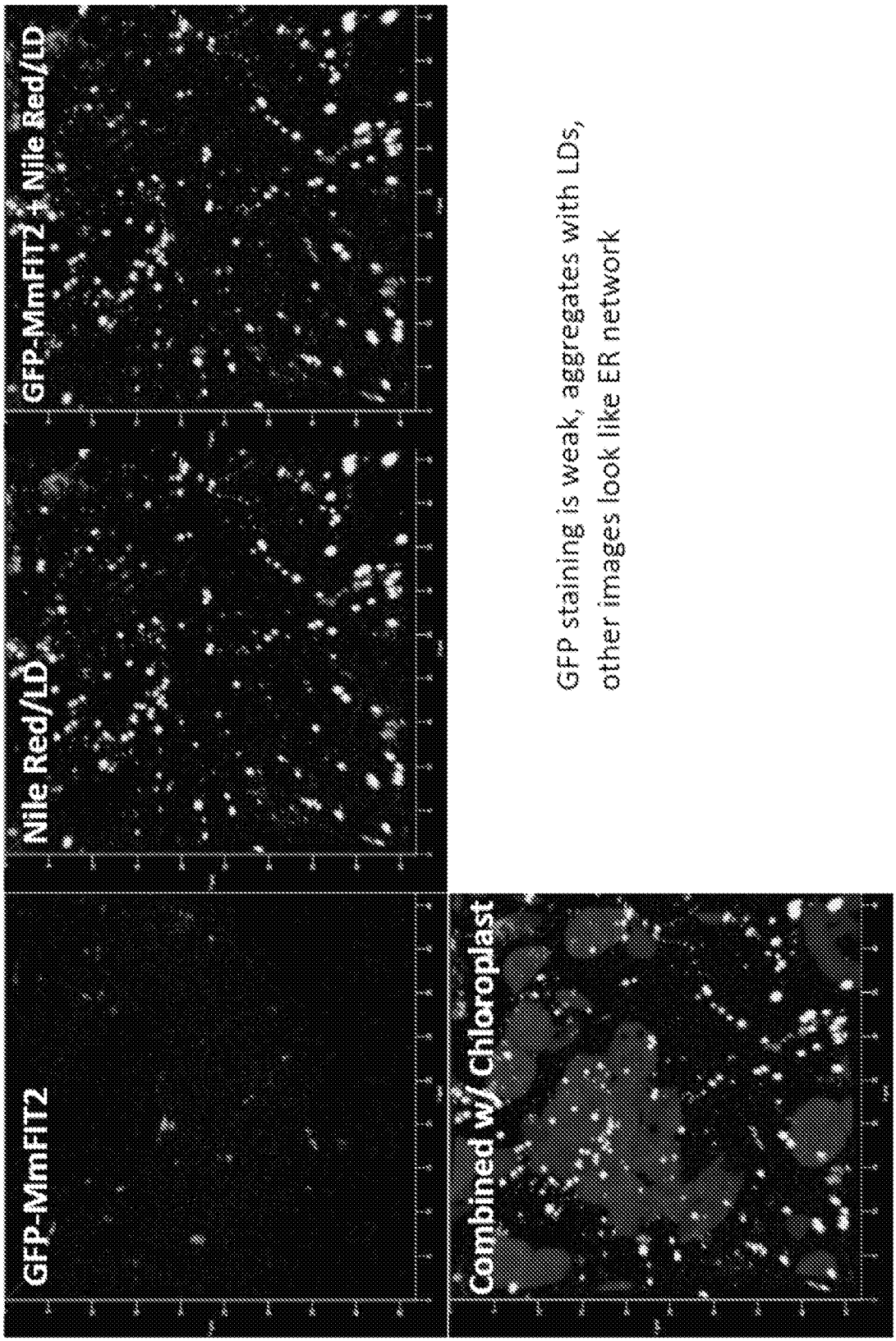

GFP-MmFIT2
Nile Red/LD
GFP-MmFIT2 + Nile Red/LD
Combined w/ Chloroplast
GFP staining is weak, aggregates with LDs,
other images look like ER network

FIG. 35

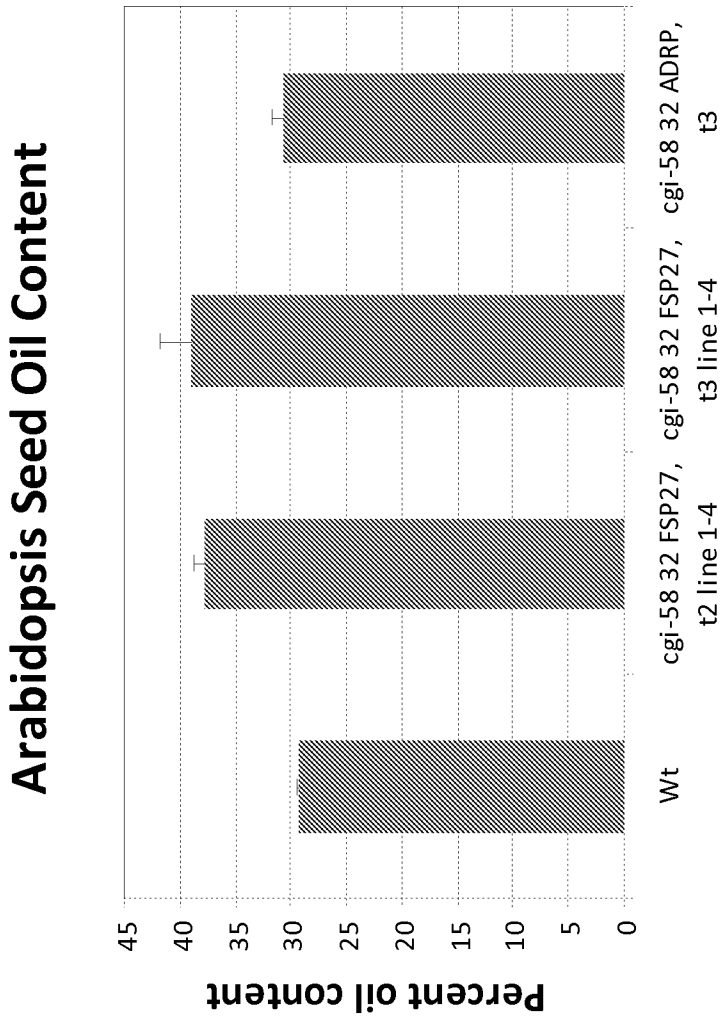

Arabidopsis Seed Oil Content
Percent oil content
45
40
35
30
25
20
15
10
5
0
Wt
cgi-58 32 FSP27, t2 line 1-4
cgi-58 32 FSP27, t3 line 1-4
cgi-58 32 ADRP, t3

FIG. 36

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2013/076672

A. CLASSIFICATION OF SUBJECT MATTER
IPC(8) - C12N 15/82 (2014.01)
USPC - 435/419
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC(8) - A01H 3/00; C12N 15/10, 63, 82 (2014.01)
USPC - 435/320.1, 419, 468, 470; 800/278, 281

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
CPC - C12N 15/8201, 15/8202, 15/8216, 15/8223, 15/8247 (2014.01)

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
PatBase, Google Patents, Google Scholar, ProQuest

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X ---- Y	US 2011/0314725 A1 (PETRIE et al) 29 December 2011 (29.12.2011) entire document	1-3, 5-7, 10, 11, 13-16 ---------- 4, 8, 9, 12, 17, 18
X	WO 2012/075543 A1 (YANG et al) 14 June 2012 (14.06.2012) entire document	19-21
Y	PURI et al. 'Fat-specific Protein 27, a Novel Lipid Droplet Protein That Enhances Triglyceride Storage. Journal of Biological Chemistry. 282(47). Pages 34213-34218. 19 September 2007. entire document	4, 12
Y	MEINKE et al. 'Leafy Cotyledon Mutants of Arabidopsis.' The Plant Cell. Vol. 6, No. 8. Pages 1049-1064. August 1994. entire document	8, 9, 17, 18
A	US 2010/0021912 A1 (FARESE, JR. et al) 28 January 2010 (28.01.2010) entire document	1-21
A	US 2010/0221400 A1 (CHAPMAN et al) 02 September 2010 (02.09.2010) entire document	1-21

☐ Further documents are listed in the continuation of Box C. ☐

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
11 February 2014	28 FEB 2014

Name and mailing address of the ISA/US
Mail Stop PCT, Attn: ISA/US, Commissioner for Patents
P.O. Box 1450, Alexandria, Virginia 22313-1450
Facsimile No. 571-273-3201

Authorized officer:
Blaine R. Copenheaver
PCT Helpdesk: 571-272-4300
PCT OSP: 571-272-7774