

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
6 February 2003 (06.02.2003)

PCT

(10) International Publication Number
WO 03/010507 A2

- (51) International Patent Classification⁷: **G01N**
- (21) International Application Number: PCT/IB02/04257
- (22) International Filing Date: 16 July 2002 (16.07.2002)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
10-2001-0042886 16 July 2001 (16.07.2001) KR
- (71) Applicant (for all designated States except US): **AHRAM BIOSYSTEMS INC.** [KR/KR]; 405 Jangan Bldg., 464-1 Jangan 3-dong, Seoul 130-103 (KR).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): **HWANG, Inhwan** [KR/KR]; 9-1702 Gyosu Apt., Jigok-dong, Nam-gu, Pohang-si, Gyeongsangbuk-do, Seoul 790-390 (KR). **KIM, Dae, Heon** [KR/KR]; 812 Cheonpyeong-ri, Sicheon-myeon, Sancheong-gun, Gyeongsangnam-do, Seoul 666-931 (KR). **LEE, Yong, Jik** [KR/KR]; 3-404, Seungri Apt., Hyoja-dong, Nam-gu, Pohang-si, Gyeongsangbuk-do, Seoul 790-330 (KR). **JIN, Jing, Bo** [CN/KR]; Center for Plant Intracellular, Trafficking, Postech Foundation, San 31, Hyoja-dong, Nam-gu, Pohang-si, Gyeongsangbuk-do, Seoul 790-784 (KR).
- (74) Agent: **KIM, Dong, Jin**; Ahram Biosystems Inc., 405 Jangan Bldg., 464-1 Jangan-3-dong, Seoul 130-103 (KR).
- (81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- Published:**
— without international search report and to be republished upon receipt of that report
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: SYSTEM FOR INTRACELLULAR PROCESS MONITORING AND *IN VIVO* DRUG SCREENING

Names	Schematic presentation	Localization
F1-H ⁺ -ATPase:GFP		Mitochondria
RbcS:GFP		Chloroplast (stroma)
RbcS:RFP		Chloroplast (stroma)
Cab:GFP		Chloroplast (stroma)
RA:GFP		Chloroplast (stroma)
AtOEP7:GFP		Chloroplast (outer envelop membrane)
AtOEP7:RFP		Chloroplast (outer envelop membrane)
NLS:GFP		Nucleus
NLS:RFP		Nucleus
GFP:SKL		Peroxisome

(57) Abstract: This invention relates to a method for detecting characteristics related to intracellular trafficking and subcellular localization of a selected reporter protein. The invention has a wide spectrum of important applications including use in screens to detect compounds that modulate protein trafficking *in vivo*.

WO 03/010507 A2

5

10

FIELD OF THE INVENTION

15

25

30

35

for developing new drug screening systems have been actively carried out recently.

Development of a new drug screening system depends basically on (1) discovery and establishment of a specific target related to a disease or a biological regulation process, and (2) development and optimization of an assay method to
5 determine the effects of drug candidates on a specific target. In order to develop efficient drug screening systems, therefore, selective detection techniques must be developed first. Such detection techniques can be used to identify detailed causes of biological phenomena from the vast amount of the genetic and proteomic information, or to detect specific intracellular or intercellular processes related to such causes. These
10 selective detection techniques provide fundamental basis not only for identifying specific targets for drug screening, i.e., biological factors causing specific phenomena, but also for establishing systematic assay methods that allow more selective drug screening.

The assay method to evaluate the effects of drug candidates on a specific target
15 can be categorized into three classes. The first is a protein- or a molecule-based method in which the effect of the drug candidate on the activity of an enzyme or a receptor critical in the induction mechanism of a disease is directly observed to screen the drug candidates. The second is an organism-based method in which the effect of the drug candidate on the morphological or biochemical characteristics of an organism is
20 observed to screen the drug candidates. The last is a cell-based method in which the effect of the drug candidate on the morphology of the cell or subcellular organelles, and expression, trafficking, and metabolism of proteins in the cell is observed to screen the drug candidates. The protein-based method has the advantage that the assay can be performed for a specific target protein. However, because the assay is conducted *in vitro*,
25 it is not possible to examine various complex cellular factors. Therefore, it requires many additional time-consuming experiments before examining the drug candidate to a living organism. On the other hand, the organism- or cell-based method can detect the composite changes of the organism or the cell induced by the drug candidates. But it also requires additional time-consuming experiments to confirm that the drug candidate
30 specifically affects the target gene or protein.

To overcome such problems of the prior drug screening methods, new technologies are recently under development using genetic transformation techniques. Genetic functions of a specific gene or biological functions of a protein expressed from the specific gene can be elucidated by examining the phenotype of the transformed
35 organism or cell. Drug candidates can be more selectively screened by examining their effects on the specific phenotype of the transformed organism or cell.

Previously, the radioisotope labeling method has been frequently used to selectively observe a specific phenomenon in a cell or an organism. Recently, fluorescent proteins such as green fluorescent protein (GFP), red fluorescent protein (REF) (Moris et al., 1974), and their derivatives (Evans, WO98/21355) that can be expressed in the cell have been developed, and various researches have been reported using expression of these fluorescent proteins. Examples of using the fluorescent proteins include observation of the gene expression and subcellular protein localization (Chalfie and Prasher, US5491084), visualization of the subcellular organelles (Kost et al., 1998), visualization of the protein trafficking in the secretory pathway (Kaether and Gerdes, 1995), and visualization of the protein expression pattern in a plant cell (Hu and Cheng, 1995) and in *Drosophila* embryo (Davis and Viestra, 1998). The fluorescent proteins have been also used to develop new drug screening methods. Examples include a method using mutated organisms transformed to express a modified GFP (Ward and Chalfie, WO95/21191), a method using a cell transformed to express a chimeric protein comprising a fluorescent protein and a transcription factor related to the activity of a cell surface receptor which regulates signal transduction (Harpold et al., US5401629), and a method using a cell transformed to express a chimeric protein comprising a fluorescent protein and the active site of the protein kinase (Thastrup et al., WO96/23898).

Among these studies that utilize the expression of the fluorescent proteins, several examples showed that intracellular processes related to specific proteins can be selectively observed among the complex cellular processes by visualizing the expression and trafficking of a selected protein (Ward and Chalfie, WO95/21191; Kost et al., 1998; Gilooly et al., 2000; Pih et al., 2000). Various researches have reported signal proteins that have trafficking signals targeting to specific subcellular organelles. For example, the chlorophyll-binding proteins are transported to chloroplast, the nuclear localization signal domain (NLS domain) of SV40 is targeted to the nucleus (Goldfarb et al., 1986), the peroxisome targeting motif SKL is targeted to peroxisome (Davis and Viestra, 1998), and F1- H^+ -ATPase is transported to mitochondria (Niwa et al., 1999). As in these examples, most of proteins in the cell are transported to specific subcellular organelles related to their functions. Therefore, it is possible to visualize and observe intracellular trafficking of a specific signal protein (Harpold et al., US5401629; Kost et al., 1998). Phospholipid-specific intracellular processes can also be observed by expressing phospholipid-binding proteins linked to a fluorescent protein. For example, PLC-Delta PH (Stauffer et al., 1998), AtPH, and FAPP1 (Dowler et al., 2000) domains are known to bind specifically to PI(3,4)P2, PI(3)P, and PI(4)P, respectively. A chimeric protein having the PH domain and GFP was shown to be translocated to the plasma

membrane when expressed in a cell (Kost et al., 1998).

It would be useful to have methods for detecting compounds that modulate intracellular movement of proteins (trafficking). Particular methods would be able to monitor changes in protein localization *in vivo*. It would be especially useful to have methods that express the protein and provide for reliable *in vivo* imaging of the protein inside living cells.

SUMMARY OF THE INVENTION

The invention generally relates to methods for detecting intracellular movement (trafficking) of at least one selected protein inside cells. Additional invention methods involve screening compounds for capacity to modulate at least one of expression and trafficking of the protein *in vivo*. The invention has many important uses including detecting compounds that modulate protein trafficking to at least one pre-selected subcellular organelle.

The invention provides useful methods for detecting trafficking of at least one selected (reporter) protein inside living cells. Particular reporters of interest include at least one "signal" sequence and at least one detectable tag. Preferred signal sequences help shuttle the protein to at least one other intracellular location. More particular reporters include at least one detectable tag, usually a fluorescent (or potentially fluorescent) sequence, that is operably and covalently attached to the signal sequence in-frame, thereby labeling the reporter protein. Particular signal sequences of interest include or consist of amino acid residues that are known to help shuttle the reporter protein to at least one subcellular organelle. As discussed in more detail, such tagged reporter proteins can be used to screen compounds for capacity to modulate (increase or decrease) trafficking to one or more of the organelles.

Practice of the invention generally involves detecting trafficking of at least one of the selected reporter proteins inside cells. For instance, and in one embodiment, the signal sequence of a reporter include or consist of recognition site(s) that facilitate movement of the reporter protein to a sub-cellular organelle. An example of such proteins are characterized herein as Group I signal proteins (see Figure 1). In another embodiment, the proteins include or consist of amino acid sequences that are recognized by and typically contacted by endoplasmic reticulum (ER). Preferred examples of such proteins are featured herein as Group II signal proteins (see Figure 4). These reporter protein groups, as well as others discussed below, can be used as important tools to detect compounds that modulate such trafficking *in vivo*. Such proteins can also be used to confirm trafficking activity in cases in which such function is known or

suspected.

The present invention provides important advantages.

For instance, the invention provides, for the first time, recognition that operably linked fusions of a signal protein and a detectable tag can be used to screen compounds
5 for capacity to modulate intracellular trafficking *in vivo*. That is, the invention is one of general application in which a wide spectrum of suitable fusion proteins can be made and used to detect compounds for capacity to increase, decrease or preserve existing intracellular protein trafficking to one or more pre-selected subcellular organelles. Thus in one embodiment, any of the Group I or II proteins can be fused to at least one
10 detectable tag to create useful reporter protein fusions. Such proteins can be used, for instance, to detect compounds for ability to modulate trafficking to at least one of the nucleus, chloroplast, mitochondria, and endoplasmic reticulum. Most reporter proteins in accord with the invention will include signaling features that provide for trafficking of the reporter to a single pre-selected subcellular organelle. Also envisioned however
15 are reporters with multiple signal proteins (eg., two or three) that can provide trafficking to more than one of pre-selected organelle.

Additional benefits are provided by the invention.

For instance, and in one embodiment, each reporter protein includes at least one fluorescent sequence preferably bound in-frame, and either directly or indirectly, to at
20 least one signal protein. In such embodiments, the fluorescent sequence is intended to provide a highly sensitive and reproducible tag that can be used to detect minute amounts of trafficked protein inside cells. Such advantages enhance the sensitivity, selectivity and reproducibility of the invention particularly in screens that are designed to detect compounds with potential trafficking activity.

25 Additionally, the invention is fully compatible with existing screening platforms including what is often referred to as "high" or "ultra-high" throughput strategies. For instance, the invention can provide an important and valuable *in vivo* confirmation of such strategies where trafficking activity of a subject compound is known or suspected.

Accordingly, and in one aspect, it is an objective of the present invention to
30 provide a method for detecting trafficking and localization of a selected protein inside a cell. In one embodiment, such a method includes introducing and expressing at least one type of, preferably less than about five, more preferably one type of reporter protein inside cells and detecting the cellular distribution of that reporter to detect trafficking and localization of the desired protein within the cells.

35 The general detection method can be used as a screening platform to detect and optionally quantify trafficking activity of one or more compounds to be tested. Thus in

one invention embodiment, the cell expressing the reporter protein can be contacted with at least one desired compound to be tested, preferably less than about fifty of such compounds, more preferably less than about ten to fifteen of such compounds, even more preferably one compound to be tested. The compound to be tested is added to the cells under conditions conducive to providing a protein trafficking effect. That is, the compound can be added to cultured cells before, during or after introduction of the reporter protein. In this example of the invention, cellular distribution of the reporter protein is visualized in the presence of the compound to detect any modulation of protein trafficking inside the cells. Typical practice of the invention involves comparison with a suitable control eg., as when water or buffer is added to the cells in lieu of the tested compound.

The invention further encompasses diverse and complex intracellular processes and a detailed method for selectively screening compounds (eg., chemicals) that affect a specific intracellular process based on the above detection method.

To accomplish this object, the present invention provides:

- (1) an efficient method for preparing a transformed cell to express a reporter protein comprising a signal protein having a trafficking signal targeting to a specific intracellular organelle and a fluorescent protein linked thereto,
- (2) detailed methods for preparing various recombinant genes, each comprising a gene encoding a signal protein and a gene encoding a fluorescent protein linked thereto, and
- (3) methods for constructing recombinant plasmids that can be used to efficiently express the reporter proteins in a cell.

The present invention also provides a systematic method for transforming a cell with the recombinant plasmid and expressing efficiently the reporter protein in the transformed cell. Furthermore, the present invention provides a detailed method for monitoring details of the intracellular distribution of the reporter protein in a living cell by observing the fluorescence image either continuously or step by step during the expression and trafficking processes.

In addition to providing the efficient systematic method for measuring the spatial distribution specific to the expression and trafficking processes of a selected protein in the transformed cell, it is another objective of the present invention to provide a method for selectively screening drug candidates that affect a specific intracellular process in the transformed cell.

In order to expand and systemize the usages of the selective detection method for the intracellular processes and the drug screening method based on this detection

method, the present invention provides recombinant genes for diverse signal proteins that direct targeting to specific intracellular organelles such as nucleus, mitochondria, chloroplast, peroxisome, plasma membrane, endoplasmic reticulum, Golgi apparatus, storage vacuole, lytic vacuole, prevacuolar compartment, etc. In addition, the present invention aims to provide an optimized overall procedure from the construction of the recombinant plasmids to the preparation of the transformed cells.

Practice of the invention provides additional benefits.

For instance, it establishes a method for selectively detecting specific characteristics related to intracellular trafficking and localization of a selected protein and a method for selectively screening chemicals that affect such specific cellular characteristics of the selected protein.

Also, it provides detailed methods for preparing the transformed cells that can express diverse reporter proteins targeting to various subcellular organelles. Therefore, a fundamental basis is provided for using more systematically and more thoroughly the selective detection method for the intracellular processes and the selective drug screening method based on this detection method.

The invention also establishes a method for determining the detailed effects of chemicals by examining the effects of several known chemicals on the intracellular processes. In further detail, the present invention provides (1) the method for selectively screening chemicals that affect the intracellular trafficking and localization of a selected protein, (2) the method for screening chemicals that inhibit or enhance the transcription or translation process by monitoring the effects of the chemicals on the expression of the reporter protein, and (3) the method for screening cytotoxic chemicals by observing the effect of chemicals that cause deformation, damage, or disruption of subcellular cellular organelles.

Other uses and advantages of the invention are disclosed below.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows schematic diagrams of the reporter proteins targeting to the subcellular organelles enclosed by membrane, which are classified as Group I proteins in the present invention.

Figure 2 shows fluorescence images showing the expression of the Group I reporter proteins in the cell, wherein

(a) shows that AtOEP7:GFP is localized to the envelope of chloroplast, wherein the red fluorescent signal is the auto-fluorescence signal of chloroplast and the yellow fluorescence signal is an merged image of the auto-fluorescence signal of chloroplast

and the green fluorescence signal of the reporter protein;

(b) is a photograph where the red auto-fluorescence of chloroplast in (a) is eliminated by using a filter;

(c) shows that Rubisco small subunit (RbcS) is localized in the stroma of chloroplast;

(d) shows that Cab:GFP emits in the chloroplast;

(e) shows that Rubisco Activase (RA) linked to GFP is localized in the chloroplast;

(f) shows that F1- H^+ -ATPase linked to GFP is localized in the mitochondria;

(g) shows that peroxisome targeting signal (SKL) linked to GFP is localized in the peroxisome envelope; and

(h) shows that nuclear localization signal (NLS) linked to GFP is localized in the nucleus.

Figure 3 shows a photograph of a Western blot analysis of the expressed AtOEP7:GFP, wherein T, S, and M indicate the total protein extract, the soluble fraction, and the membrane fraction, respectively.

Figure 4 shows schematic diagrams of the reporter proteins targeting to the subcellular organelles by endosomal trafficking, which are classified as Group II proteins in the present invention.

Figure 5 shows fluorescence images showing the expression of the Group II reporter proteins in the cell, wherein

(a) shows that H^+ -ATPase:GFP is localized in the plasma membrane;

(b) shows the fluorescence image of sialtransferase (ST) translocated in the Golgi apparatus, wherein the red fluorescence signal is the auto-fluorescence signal of chloroplast;

(c) shows that the reporter protein BiP:RFP is localized in the lumen of the endoplasmic reticulum;

(d) shows that 500:GFP:KKXX is localized in the membrane of the endoplasmic reticulum;

(e) shows that 526:GFP is localized in the storage vacuole;

(f) shows that Chi-n:RFP:Chi-c is localized in the storage vacuole;

(g) shows that Chi-n:GFP, in which the carboxyl terminus of the signal protein, chitinase, is not present, fails to translocate in the storage vacuole but appears as speckles in the endoplasmic reticulum; and

(h) shows that SPO:GFP is distributed evenly in the lytic vacuole.

Figure 6 shows schematic diagrams of the reporter proteins binding specifically

to phospholipids, which are classified as the Group III proteins in the present invention.

Figure 7 shows fluorescent images showing the expression of the Group III reporter proteins, wherein

5 (a) shows that the reporter protein GFP:EBD emits fluorescence at the outer membrane of vacuole, indicating that phosphatidylinositol 3-phosphate (PI(3)P) is present on the outer membrane of vacuole;

(b) shows that GFP:FAPP1 emits fluorescence in the plasma membrane indicating that phosphatidylinositol 4-phosphate (PI(4)P) is present in the plasma membrane; and

10 (c) shows that GFP:PH emits fluorescence in the plasma membrane, indicating that phosphatidylinositol 4,5-diphosphate (PI(4,5)P₂) is present in the plasma membrane.

Figure 8 shows photographs representing the inhibitory effect of wortmannin on the intracellular trafficking of a Group I reporter protein RbcS:GFP. Wortmannin is
15 known to be an inhibitor of PI(3)P and PI(4)P.

(a) shows that RbcS:GFP is targeted correctly to the chloroplast in the control protoplast, as expected.

(b) and (c) show that in the presence of wortmannin, the green fluorescence signal is not located in the chloroplast but appears as speckles or aggregates.

20 Figure 9 shows photographs representing the effect of a chemical on the intracellular trafficking of a Group II reporter protein, wherein

(a) and (b) show that the green fluorescence signal of 500:GFP:KKXX is observed as numerous networks in the control protoplast; and

25 (c) and (d) show that in the presence of BafA1, the green fluorescent signal is observed as ring patterns in the plasma and vacuolar membranes.

Figure 10 shows photographs showing the change induced by a specific inhibitor upon co-expression of two reporter proteins, wherein

(a) and (b) show the green and red fluorescent signals in the control protoplast, respectively;

30 (d) shows that treatment with brefeldin A does not affect BiP;

(e) shows that treatment with brefeldin A disrupts the Golgi apparatus and thus ST is transported to the endoplasmic reticulum instead of the Golgi apparatus; and

(c) and (f) show that the control protoplast and the brefeldin A-treated protoplast represent distinctive difference in their fluorescence images.

35 Figure 11 shows photographs that visualize the effects of brefeldin A disrupting the subcellular organelles, wherein

(a) shows that BiP:RFP is distributed along the structure of the endoplasmic reticulum in the control protoplast;

(c) visualizes that the disrupted structure of the endoplasmic reticulum in the presence of brefeldin A; and

5 (b) and (d) show the protoplasts in (a) and (c) observed under bright field.

Figure 12 shows photographs representing the inhibitory effect of chemicals on Group III reporter proteins, wherein

(a) shows that the green fluorescence signal in the control protoplast is located on the vacuolar membrane along the distribution of PI(3)P which binds to EBD, but in
10 the presence of wortmannin or LY294002, the green fluorescence signal is distributed throughout the cytosol; and

(b) shows that the distribution of the fluorescence signal is not affected by the above chemicals when the same experiments were performed with GFP:EBDC1358S, in which the amino acid residue 1358 of EBD was mutated.

15 Figure 13 shows photographs showing the variation in the intensity of the fluorescent signal induced by an expression inhibitor, wherein

(a) shows that the intensity of the green fluorescence is decreased in the cycloheximide-treated protoplast compared to that in the control protoplast of (b); and

(c) and (d) show the auto-fluorescence signal of the chloroplast presented for
20 relative comparison of the cell growth and the metabolism conditions.

DETAILED DESCRIPTION OF THE INVENTION

As discussed, the invention relates to a method detecting intracellular movement (trafficking) of a selected reporter protein inside cells. In one example, the
25 invention provides methods for screening compounds for capacity to modulate at least one of expression and trafficking of the protein *in vivo*. Particular use of the invention involves screening compounds for capacity to modulate protein trafficking to a subcellular organelle.

As also discussed, the present invention provides a method for detecting
30 specific characteristics related to trafficking and localization of a selected protein in a cell. In one embodiment, the method comprises at least one of:

- (a) preparing a recombinant gene comprising a gene encoding a signal protein which includes a trafficking signal targeting to a specific subcellular organelle and a gene encoding a fluorescent protein linked thereto;
- 35 (b) preparing a recombinant plasmid comprising the recombinant gene with a promoter and a terminator operably linked thereto so that the recombinant

gene can be expressed in the cell;

- (c) transforming the cell with at least one recombinant plasmid prepared in step (b);
- (d) expressing the reporter protein comprising the signal protein and the fluorescent protein in the transformed cell; and
- (e) detecting the cellular distribution of the reporter protein by monitoring the fluorescence image of the transformed cell during the expression or trafficking, or thereafter.

In a second aspect, the present invention provides a method for screening chemicals that affect specific characteristics related to trafficking and localization of a selected protein in a cell. In one embodiment, the method comprises at least one of:

- (a) preparing a recombinant gene comprising a gene encoding a signal protein which includes a trafficking signal targeting to a specific subcellular organelle and a gene encoding a fluorescent protein linked thereto;
- (b) preparing a recombinant plasmid comprising the recombinant gene with a promoter and a terminator operably linked thereto so that the recombinant gene can be expressed in the cell;
- (c) transforming the cell with at least one recombinant plasmid prepared in step (b);
- (d) expressing the reporter protein comprising the signal protein and the fluorescent protein in the transformed cell, while treating the transformed cell with a chemical before, after, or at the same time as the expression;
- (e) detecting the cellular distribution of the reporter protein by monitoring the fluorescence image of the transformed cell treated with the chemical during the expression or trafficking, or thereafter; and
- (f) determining the effect of the chemical by comparing the fluorescence image obtained in step (e) with that of a control transformed cell which is not treated with the chemical.

In the third aspect, the present invention provides recombinant genes encoding the reporter proteins that are used in the above methods to visualize the trafficking of the reporter proteins and their distributions in subcellular organelles. The recombinant gene comprises a gene encoding a signal protein which includes a trafficking signal targeting to a specific subcellular organelle and a gene encoding a fluorescent protein linked thereto.

Details of the compositions of the present invention are described below. In the present invention, the signal protein is selected from the proteins that have trafficking

signals targeting to nucleus, mitochondria, chloroplast, peroxisome, plasma membrane, endoplasmic reticulum, Golgi apparatus, storage vacuole, lytic vacuole, and prevacuolar compartment, and also those proteins targeting to 3 classes of phospholipids. Examples of the signal proteins include NLS (nuclear localization sequence), AtOEP7, Cab
5 (chlorophyll a/b binding protein), SKL (peroxisome targeting motif), RbcS (rubisco small subunit), RA (rubisco activase), F₁-H⁺-ATPase, H⁺-ATPase, BiP (chaperone binding protein), ST (Sialyltransferase). Chi (chitinase), recombinant clone 526, clone 491, clone 500, AtVTI1a, SPO (sporamin), EBD, AtPH, FAPP, PH, etc. The present invention provides methods for preparing transformed cells that can express reporter
10 proteins, each comprising one of the signal proteins described above and a fluorescent protein label linked thereto. The present invention also provides a systematic method for selectively detecting the intracellular processes using the transformed cells and a systematic method for selectively screening drug candidates based on this detection method.

15 More particularly, the methods of the present invention include a step of selecting a specific protein that has a property of translocating to a specific subcellular location, a step of synthesizing the whole gene of the selected signal protein or a portion thereof encoding the trafficking signal of the selected protein, a step of synthesizing a gene encoding a fluorescent protein that can be linked to the signal protein to
20 fluorescently visualize the subcellular localization, and a step of constructing a recombinant gene comprising a gene encoding the signal protein and a gene encoding the fluorescent protein linked thereto. The function of the signal protein can change depending on the way that the fluorescent protein is linked to the signal protein. The present invention therefore provides compositions of the reporter proteins whose signal
25 proteins can correctly direct the trafficking, and also construction methods thereof.

The present invention also provides a procedure for constructing recombinant plasmids that can be used to express the recombinant genes in a cell. The recombinant plasmid can be constructed by ligating a recombinant gene into a vector containing a promoter, a terminator, and other necessary factors. As well known to the persons
30 having ordinary skills in the art to which the present invention pertains, methods for transforming a cell by introducing the recombinant plasmid include, but are not limited to, chemical-mediated methods using PEG (polyethylene glycole), potassium phosphate, or DEAE-dextran, cationic lipid-mediated lipofection, microinjection, electroporation, and electrofusion. In the transformation, one type of the recombinant plasmid could be
35 introduced, or else two or more types of the recombinant plasmids can be introduced to express two or more reporter proteins simultaneously. The conditions need to be

optimized to efficiently express the reporter protein comprising the signal protein and the fluorescent protein in the transformed cell. The persons skilled in the art can select appropriate conditions depending on the signal protein and the fluorescent protein used. Detailed structures of the recombinant plasmids are explained in the examples of the present invention to present the amino acid sequences of the reporter proteins or corresponding nucleic acid sequences.

In constructing the reporter proteins according to the present invention, the signal proteins were classified into three classes based on the protein trafficking mechanisms, in order to show that the methods of the present invention for constructing the reporter proteins can be commonly used for various mechanisms of protein trafficking to different subcellular organelles. Signal proteins targeting to nucleus, chloroplast, mitochondria, etc., correspond to the case that a specific portion of the signal protein directly acts as a recognition signal to direct the intracellular trafficking. These signal proteins are classified as Group I for convenience in the specification of the present invention (see Figure 1). Signal proteins targeting to endoplasmic reticulum, Golgi apparatus, lytic vacuole, storage vacuole, plasma membrane, etc., correspond to the case of the endosomal trafficking in which a specific portion of the signal protein acts as a signal to be captured by endoplasmic reticulum so that the signal protein is translocated as enclosed in endoplasmic reticulum. These signal proteins are classified as Group II for convenience in the specification of the present invention (see Figure 4). In the present invention, detailed methods are provided for visualizing the trafficking processes and the cellular distributions of these classes of the signal proteins. In addition, signal proteins related to intracellular signal transduction via specific binding to phospholipids are selected in the present invention, and methods are provided for observing the subcellular organelles that contain specific phospholipid. These signal proteins are classified as Group III for convenience in the specifications of this invention (see Figure 6).

It will often be useful to practice the invention with reporter proteins having specific trafficking features. Thus for example in embodiments in which it is desirable to detect trafficking and localization of reporters with Group I or Group II signal proteins (or both Group I and Group II signal proteins), it will not be as useful to detect trafficking of report proteins having specific Group III signal proteins.

Thus in an embodiment along these lines, the invention can be employed to detect trafficking and localization of a selected reporter protein in which the reporter includes a Group I or Group II signal protein. Preferably, such a reporter does not include a Group III signal protein as defined herein. In a more specific embodiment, the

reporter protein of interest does not include the endosome binding domain (EBD) of human early endosome antigen 1 (a Group III signal protein) as defined by Kim D.H et al. (2001) in *Plant Cell* 13: 287 ie., amino acids 1257 to 1411 of EBD. In another embodiment, the reporter protein does not include one or more of the following signal proteins as defined in Figure 6: AtPH; FAAP1; and PH. In still another embodiment, the reporter protein for use with the present invention does not include a recognition site(s) that binds specifically to subcellular organelles that contain specific phospholipid as defined herein for the Group III family of signal proteins.

The present invention thus provides a method for detecting trafficking and localization of a selected reporter protein inside a cell in which the protein does not include one or more of the EBD signal protein, AtPH; FAAP1; or PH as defined in Figure 6. Thus in this illustration of the invention, the method includes introducing and expressing the reporter protein inside cells and detecting the cellular distribution of that reporter to detect trafficking and localization of the desired protein within the cells. Such a method can also be used in accord with the invention as a screening platform to detect and optionally quantify trafficking activity of one or more compounds to be tested. Thus in one invention embodiment, the cell expressing the reporter protein (without one or more of the EBD, AtPH; FAAP1; or PH signal proteins as defined in Figure 6) can be contacted with at least one desired compound to be tested. The compound to be tested can be added to cultured cells before, during or after introduction of the reporter protein as discussed previously. Such a specific invention method is useful in a variety of settings eg., where detection of Group I, Group II, or both Group I and II trafficking is desired.

In other embodiments, it will often be useful to practice the invention methods with reporter proteins that do not include Group II signal proteins. For instance, it will be useful to detect trafficking with reporters that do not include one or more of the following Group II signal proteins: ST; BiP; Sporamin; and H⁺-ATPase as defined by Kim D.H et al. (2001) in *Plant Cell* 13: 287. In other embodiments, it will be more useful to detect trafficking with reporters in which the signal protein does not include a recognition site(s) that is involved with endosomal trafficking such that a specific portion of the signal protein acts as a signal to be captured by endoplasmic reticulum. More particular signal proteins along these lines, and which would not be as useful in this invention embodiment, are translocated as enclosed in endoplasmic reticulum to eg., the Golgi apparatus, lytic vacuole, storage vacuole, and plasma membrane.

Accordingly, the invention provides a method for detecting trafficking and localization of a selected reporter protein inside a cell in which the protein does not

include one or more of the ST; BiP; Sporamin; and H⁺-ATPase signal proteins as defined eg., in Figure 4. In this illustration of the invention, the method includes introducing and expressing the reporter protein inside cells and detecting the cellular distribution of that reporter to detect trafficking and localization of the desired protein within the cells. Importantly, the method can also be used in accord with the invention as a screening platform to detect and optionally quantify trafficking activity of one or more compounds to be tested. Thus in one invention embodiment, the cell expressing the reporter protein (without one or more of the ST; BiP; Sporamin; and H⁺-ATPase signal proteins as defined in Figure 4) can be contacted with at least one desired compound to be tested. The compound to be tested can be added to cultured cells before, during or after introduction of the reporter protein as discussed previously. Such a specific invention method has a variety of important uses such as in settings in which detection of Group I, Group III, or both Group I and III trafficking is desired.

Using the trafficking properties of these signal proteins, details of the subcellular localization of the signal proteins can be specifically visualized, for example, on the membrane, or inside or outside of an organelle. The present invention also provides a method for visualizing the localization of two or more proteins simultaneously by using two or more fluorescent proteins with different colors.

In the embodiments of the present invention, green fluorescent protein (GFP, Davis and Viestra, 1998) and red fluorescent protein (RFP) are used for constructing the reporter proteins to visualize their cellular localization. However, the person having ordinary skill in the art to which the present invention pertains can fully understand that the reporter proteins of the present invention can be constructed by using fluorescent proteins other than GFP and RFP.

The expression and trafficking processes of the reporter protein can be visualized in details for each stage of the processes by continuously monitoring the images of the fluorescence emitted by the reporter protein expressed in the transformed cell, using a fluorescence microscope at a specific wavelength.

As it becomes possible to detect the expression and trafficking processes of a specific protein and its cellular distribution, a selective drug screening system for identifying chemicals inhibiting or enhancing the intracellular trafficking of the selected protein is established using this detection method. More particularly, it is demonstrated that chemicals affecting the intracellular trafficking can be identified by treating the transformed cell with a chemical before, after, or at the same time as the expression of the reporter protein, and then identifying the effect of the chemical by comparing the cellular distribution of the reporter protein in the transformed cell treated with the

chemical with that in the control transformed cell which is not treated with the chemical.

The same method can be used to screen chemicals inhibiting or enhancing the transcription or the translation of proteins because the level of the protein expression can be examined from decrease or increase in the intensity of the fluorescence signal from the reporter protein. This is also demonstrated in the examples of the present invention.

Morphological changes induced by a chemical, such as modification, damage, or destruction of the subcellular organelles can be detected by observing the distribution or pattern of the fluorescence signal from the reporter protein. Therefore, it is also possible to screen cytotoxic chemicals that cause alteration of the subcellular organelles. This is also demonstrated in the examples of the present invention.

In order to show that the selective drug screening system provided by the present invention can be used practically, known inhibitors such as bafilomycin A1, wortmannin, and brefeldin A are examined to confirm that inhibition of the intracellular trafficking and expression processes and also cytotoxicity causing morphological changes of the subcellular organelles can be detected practically. This is described in detail in the examples of the present invention. As described in the examples, the signal proteins included in Group I, II, and III are examined to check the difference arising from the protein trafficking mechanisms.

Compositions and usages of the present invention are described in detail in the embodiments with the attached drawings. The embodiments are to explain, but not limit, the present invention. The person having ordinary skill in the art to which this invention pertains can easily recognize other objects and advantages of the present invention from the attached drawings, the detailed description, and the claims of the present invention.

25

The following Examples are illustrative of the present invention.

Example 1. Construction of recombinant plasmids for expression of Group I proteins targeting to the organelles across the membrane.

The coding region for the transit peptide of F1-H⁺-ATPase (access number D88374) was amplified by polymerase chain reaction (PCR) from a λ ZAPII cDNA library using two specific primers (5'-CTTTAATCAATGGCAATG and 5'-CCATGGCCTGAACTGCTCTAAGCTT) and ligated in-frame to the 5' end of the coding region of the green fluorescent protein to generate a recombinant gene for F1-H⁺-ATPase:RFP (Niwa et al., 1999). The recombinant gene was subcloned into pUC

35

under the control of the 35S promoter to construct a recombinant plasmid for ATPase:RFP. The same method was used for construction of other recombinant plasmids.

To express the reporter protein of the ribulose biphosphate carboxylase (Rubisco) complex, the coding region for the transit peptide of the small subunit of the Rubisco complex was PCR amplified from a λ ZAPII cDNA library using two specific primers (5'-CCTCAGTCACACAAAGAG and 5'-ACTCGAGGGAATCGGTAAGGTCAG). The resulting PCR product was subcloned into pBluescript and subsequently ligated in-frame to the 5' end of the coding regions of GFP and RFP to construct recombinant plasmids for RbcS:GFP and RbcS:RFP, respectively.

The coding region of the chloroplast a/b binding protein was PCR amplified from a λ ZAPII cDNA library using two specific primers (5'-TAGAGAGAAACGATGGCG and 5'-GGATCCCGTTTGGGAGTGGA ACTCC) and used to construct a recombinant plasmid for Cab:GFP.

The coding regions of the transit peptide of rubisco activase (RA) was PCR amplified from a λ ZAPII cDNA library using two specific primers (5'-TCTAGAATGGCCGCCGCAGTTTCC and 5'-GGATCCATCTGTCTCCATCGGTTTG) and ligated to the 5' end of the coding regions of GFP and RFP to construct recombinant plasmids for RA:GFP and RA:RFP, respectively.

The coding region of the *Arabidopsis* outer envelope membrane protein, AtOEP7, a homolog of OEP14 of pea was PCR amplified from a *Arabidopsis* genomic DNA using two specific primers (OEP7-F: 5'-GACGACGACGCAGCGATG and OEP7-R: 5'-GGATCCCCAAACCCTCTTTGGATGT) which were designed to remove the natural termination codon, and subsequently ligated to the 5' end of the coding regions of GFP and RFP to construct recombinant plasmids for AtOEP7:GFP and AtOEP:RFP, respectively.

A recombinant plasmid for the nuclear localization signal (NLS), NLS:GFP, was constructed as described previously (Pih et al., 2000). A recombinant plasmid for NLS:RFP was constructed by replacing the GFP coding region with the RFP coding region in the recombinant gene for NLS:GFP.

The recombinant plasmid for the peroxisomal reporter GFP:SKL (serine, lysine,

leucine) was constructed by PCR with 326GFP (Davis and Viestra, 1998) as a template using two specific primers (5'-CCGTATGTTACATCACC and 5'-TTATAGCTTTGATTTGTATAGTTCATCCAT).

5 Schematic diagrams of the reporter proteins constructed by the methods described above are presented in Figure 1.

Example 2. Preparation of protoplasts and transformation of protoplasts with recombinant plasmids.

(a) Preparation of protoplasts.

10 Leaf tissues (5g) of 3-4 week-old *Arabidopsis* plants grown on soil in a green house were cut into small squares (5-10 mm²) with a new razor blade and incubated with 50 ml of the enzyme solution (0.25% Macerozyme R-10, 1.0% Cellulase R-10, 400 mM mannitol, 8 mM CaCl₂, 5 mM Mes-KOH, pH 5.6) at 22°C with gentle agitation (50-75rpm). After incubation, the protoplast suspension was filtered through a 100 µm
15 mesh and protoplasts were collected by centrifugation at 46xg for 5 min. The pelleted protoplasts were resuspended in 5 to 10 ml of the W5 solution (154 mM NaCl, 125 mM CaCl₂, 5 mM KCl, 5 mM glucose, 1.5 mM Mes-KOH, pH 5.6), overlaid on top of 20 ml of 21% sucrose, and centrifuged for 10 min. at 78xg. The intact protoplasts at the interface were transferred to a new tube containing 20 ml of the W5 solution. The
20 protoplasts were pelleted again by centrifugation at 55xg for 5 min and resuspended in 20 ml of the W5 solution. The protoplasts were incubated on ice for 30 minutes.

(b) Isolation of recombinant plasmid DNAs and transformation of protoplasts.

25 Recombinant plasmids were purified using Qiagen columns (Valencia, CA) according to the manufacture's protocol. To transform the protoplasts with the DNA, the protoplasts were pelleted again by centrifuge at 46xg for 5 min and resuspended in the MaMg solution (400 mM Mannitol, 15 mM MgCl₂, 5 mM Mes-KOH, pH 5.6) at a density of 5x10⁶ protoplasts/ml. The recombinant plasmids were introduced into *Arabidopsis* protoplasts by PEG-mediated transformation (Kim et al., 2001; Jin et al.,
30 2001). Plasmid DNA (about 20-50 µg at a concentration of 2 µg/µl) was added to 300 µl of the protoplast suspension, and subsequently 325 µl of the PEG solution (400 mM Mannitol, 100 mM Ca(NO₃)₂, 40% PEG 4000) was added. The mixture was gently mixed and incubated for 30 min at room temperature. After incubation, the mixture was diluted with 10 ml of the W5 solution. Protoplasts were recovered by centrifugation at
35 50xg for 5 min and resuspended in 3 ml of the W5 solution and incubated at 22°C in the

dark.

Example 3. Expression of Group I reporter proteins and observation of their expression and localization.

5 The recombinant plasmids constructed in Example 1 were used to transform the protoplasts according to the method described in Example 2. The expression of the reporter proteins after the transformation was monitored as a function of time by capturing images using a fluorescence microscope (Axioplan fluorescence microscope, Zeiss, Germany) equipped with a cooled charge-coupled device (CCD) camera. The
10 filter sets used were XF116 (exciter: 474AF20, dichroic: 500DRLP, emitter: 510AF23), XF33/E (exciter: 535DF35, dichroic: 570DRLP; emitter: 605DF50), and XF137 (exciter: 540AF30; dichroic, 570DRLP, emitter: 585ALP) (Omega, Inc, Brattleboro, VT) for GFP, RFP, and auto-fluorescence of chlorophyll, respectively. Data were then processed using Adobe (Mountain View, CA) Photoshop software, and the images were rendered
15 in pseudo-color.

(a) Localization of the chloroplast targeting reporter proteins.

The green fluorescence of the reporter protein AtOEP7:GFP was observed at the outer envelop membrane of the chloroplast. The red fluorescence in Figure 2a is the
20 auto-fluorescence of chloroplasts. Figure 2b shows the image obtained by eliminating this auto-fluorescence by using a filter. This result indicates that the chimeric protein comprising the signal protein with the chloroplast envelope targeting signal and the fluorescent protein label was correctly targeted to the chloroplast envelope membrane.

Localization of the green fluorescence of the chimeric proteins, RbcS:GFP, Cab:GFP, and RA:GFP are presented in Figure 2(c), 2(d), and 2(e), respectively. As
25 shown in the figures, RbcS:GFP was targeted to the stroma of chloroplast, and Cab:GFP and RA:GFP also emitted the fluorescence in the chloroplast. These results indicate that the chimeric proteins comprising the trafficking signal of RbcS, Cab, or RA and fluorescent reporter protein were targeted to the chloroplast.

30

(b) Mitochondria targeting of F1-H⁺-ATPase:RFP.

The red fluorescence of the reporter protein F1-H⁺-ATPase:RFP was observed in the mitochondria (Figure 2(f)). This result indicates the chimeric protein comprising the signal protein with the mitochondria targeting signal and the fluorescent protein was
35 transported into the mitochondria.

(c) Peroxisome targeting of GFP:SKL.

By analyzing the location of the green fluorescence from the reporter protein GFP:SKL, it was observed that the green fluorescence of GFP:SKL was transported to the peroxisome as shown in Figure 2(g). This result indicates that the reporter protein comprising the peroxisome targeting signal, SKL (serine, lysine, leucine), and the fluorescent protein label was translocated to the peroxisome.

(d) Nuclear targeting of NLS:GFP.

By analyzing the location of the green fluorescence from the reporter protein, it was observed that the green fluorescence of NLS:GFP was transported to the nucleus as shown in Figure 2(h). This result indicates that the reporter protein comprising the nuclear localization signal and the fluorescent protein label was translocated to the nucleus.

Example 4. Confirmation of the chloroplast envelope targeting of AtOEP7:GFP by Western blot analysis.

The recombinant plasmid for AtOEP7:GFP was constructed according to the method in Example 1. This recombinant plasmid was used to transform protoplasts according to the method in Example 2, and the transformed protoplasts were incubated for 24 hrs at 22°C. The total protein extract was prepared as follows. Five ml of cell lysate was centrifuged, suspended in 5 ml of the extraction solution (10 mM EDTA, 50 mM HEPES-KOH, 0.33 M sorbitol, 0.5 g/l BSA, 5 mM sodium ascorbate) at 4°C, and homogenized every three seconds for 20 min. The total protein extract was fractionated by ultra-centrifugation at 100,000xg to separate the soluble and membrane fractions. Both fractions were then electrophoresed on a 7.5% SDS/PAGE gel and transferred onto the PVDF membrane. The blot was probed with a polyclonal anti-GFP antibody.

As presented in Figure 3, the result shows that the expressed signal protein was transported to the chloroplast envelope membrane and not present in the cytosol. This result indicates that localization of proteins, which is conventionally determined by Western blot analysis, can be identified by the method provided by the present invention.

Example 5. Construction of recombinant plasmids for expression of Group II proteins that are transported to subcellular organelles by endosomal trafficking.

The full length coding sequence of H⁺-ATPase (*Arabidopsis* AHA2) was amplified with two specific primers (5'-GAGATGTCGAGTCTCGAA and 5'-CTCGAGCACAGTGTAGTGACTGG) and ligated to the 5' end of the GFP coding

sequence. The ligated recombinant gene was subcloned into the pUC vector under the control of the 35S promoter to construct a recombinant plasmid for H⁺-ATPase:GFP. The same procedure was applied in the following examples.

The coding sequence of the chaperone binding protein (BiP) (access number
5 D82817) was amplified from an *Arabidopsis* cDNA library using two specific primers,
BIP5 (5'-TACGCAAAAGTTTCCGAT-3') and BIP3 (5'-
CTAGAGCTCATCGTGAGA-3'). The amino terminal region (44 amino acids) and the
carboxyl terminal region (80 amino acids) of this gene were ligated to the amino
terminus and the carboxyl terminus of GFP or RFP, respectively, to construct
10 recombinant plasmids for BiP:GFP and BiP:RFP.

The sialtransferase (ST) cDNA was amplified from a λ ZAPII cDNA library
using two specific primers (5'-ATGATTCATACCAACTTGAAG and 5'-
GGATCCACAACGAATGTTCCGGAA). GFP or RFP was ligated in-frame to the
carboxyl terminus of ST to construct ST:GFP or ST:RFP.

15 To construct a recombinant plasmid for the chimeric protein Chi-n:RFP:Chi-c, a
DNA fragment including the RFP coding sequence without the termination codon was
inserted into the *Sma* I and *Eco* RV sites of the chitinase cDNA of pea (access number
M13968).

To express clone 500 encoding the vacuolar sorting receptor protein (BP-80)
20 without the cytoplasmic tail, a recombinant plasmid for 500:GFP was constructed by
inserting the GFP coding region without the termination codon into the *Eco*RI site of
clone 500 (Kim et al., 2001).

For clone 526 encoding BP-80 with its cytoplasmic tail substituted by the
tonoplast intrinsic protein (TIP), a recombinant plasmid for 526:GFP was constructed
25 by inserting the coding region of GFP into the *Eco*RI site of clone 526.

To express the reporter protein for clone 491 encoding the BP-80 protein,
recombinant plasmids for 491:GFP and 491:RFP were constructed by inserting clone
491 into the 5' end of the coding regions of GFP and RFP without the termination
codon, respectively.

30 A recombinant plasmid for 500:GFP:KKXX was generated as follows: The
GFP coding region without the termination codon was inserted into the *Eco*RI site of
clone 500 (Jiang and Rogers, 1998) and KKXX was then added to the C-terminus of
500:GFP by PCR amplification using two specific primers (5'-
GGATCCTCTAGAGGATCGATCCGG and 5'-
35 TTAGATGAGTTTCTTTTCTCAAAGAAAGTTTCAAAGGAATCCCCCTCC).

To express AtVTI1a, a homolog of *Arabidopsis* t-SNARE which is transported from the *trans*-Golgi network to the storage prevacuole (Zheng et al., 1999), a recombinant plasmid for RFP:AtVTI1a was constructed by ligating the coding region of AtVTI1a to the C-terminus of the RFP coding region. A recombinant plasmid for
 5 AtVTI1:GFP was constructed by ligating the coding region of GFP to the C-terminus of the coding region of AtVTI1a.

To express sporamin, a recombinant plasmid for SPO:GFP was constructed by ligating GFP to the carboxyl terminus of the sporamin B gene.

Schematic diagrams of the reporter proteins expressed from the recombinant
 10 plasmids constructed as above are shown in Figure 4.

Example 6. Observation of the expression and localization of Group II reporter proteins.

Recombinant plasmids for H⁺-ATPase:GFP, ST:GFP, BiP:GFP, 526:GFP, Chi-n:RFP:Chi-c, and 500:GFP:KKXX were constructed as described in Example 5 and
 15 used to transform the protoplasts by the method of Example 2. Expression of the reporter proteins was monitored as a function of time using a fluorescence microscope as explained in Example 3. A part of the results is given in the following.

Fluorescence of the reporter protein H⁺-ATPase:GFP was observed in the
 20 plasma membrane (Figure 5(a)). Fluorescence of ST:GFP was observed in the Golgi apparatus (Figure 5(b)). The red fluorescence in these images is the auto-fluorescence of chloroplasts. Reporter proteins, BiP:RFP and 500:GFP:KKXX showed fluorescence in the lumen and the membrane of the endoplasmic reticulum, respectively (Figure 5(c) and 5(d)). 526:GFP showed fluorescence on the membrane of the storage vacuole
 25 (Figure 5(e)) and Chi-n:RFP:Chi-c showed fluorescence in the storage vacuole (Figure 5(f)). When chitinase was used as a signal protein, Chi-n:RFP, in which the carboxyl region of chitinase was not ligated, was not targeted to the storage vacuole, but it was present as speckles in the endoplasmic reticulum (Figure 5(g)). Fluorescence of SPO:GFP was distributed uniformly throughout the lytic vacuole (Figure 5(h)).

30

Example 7. Construction of recombinant plasmids to express Group III proteins that are specific to phospholipids.

To construct the recombinant DNA for GFP:EBD, the C-terminal coding region (amino acid residue 1257 to 1411) of human early endosome antigen 1 (EEA1) was
 35 PCR amplified with two primers 5'-GAATTCGTGGCAATCTAGTCAACGG-3' and 5'-CTAATGTTAGTGTAATATTAC-3', and ligated to the C-terminus of the GFP coding

sequence without the termination codon. This recombinant DNA was inserted to a pUC vector under the control of the 35S promoter to construct a recombinant plasmid. The same cloning procedure was applied in the examples hereafter.

5 A recombinant plasmid for a EBD derivative, GFP:EBDC1358S, was prepared using a primer directing replacement of the amino acid residue 1358 to serine.

A recombinant plasmid for the chimeric protein of *Arabidopsis* Pleckstrin homology (PH) domain, GFP:AtPH, was constructed by PCR amplification using two primers 5'-CCCGGGAAATGGAGAGTATGTGGCGA-3' and 5'-TAATCACCGCCTGTGATCATA-3'. A recombinant plasmid for the chimeric protein
10 of FAPP including the PH domain, GFP:FAPP, was constructed by PCR amplification using two primers 5'-CTCGAGATGGAGGGGGTTCTGTACAAG-3' and 5'-TCACGCTTTGGAGCTCCCAAGGGC-3'. A recombinant plasmid for PH:GFP was constructed by the method of Kost B et al. (1998).

15 Schematic diagrams of the reporter proteins constructed as above are shown in Figure 6.

Example 8. Observation of the expression and localization of Group III reporter proteins.

20 Recombinant plasmids for GFP:EBD, GFP:AtPH, GFP:FAPP, and GFP:PH, were constructed as described in Example 7 and used to transform the protoplasts by the method of Example 2. Expression of the reporter proteins was monitored as a function of time using a fluorescence microscope as explained in Example 3. A part of the results is given in the following.

The reporter proteins, GFP:EBD and GFP:AtPH, showed fluorescence at the
25 outer membrane of vacuole, indicating the presence of phosphatidylinositol 3-phosphate (PI(3)P) in the outer membrane of vacuole. GFP:FAPP and GFP:PH showed fluorescence at the plasma membrane indicating the presence of phosphatidylinositol 4-phosphate (PI(4)P) and phosphatidylinositol 4,5-diphosphate (PI(4,5)P2), respectively, in the plasma membrane. These results show that it is possible to use the phospholipid-
30 specific protein or its portion as a signal protein to target a protein to the phospholipid-containing cellular compartments (see Figure 7).

Example 9. The effect of wortmannin on the intracellular trafficking of RbcS:GFP.

35 (a) Plasmid construction, transformation, and expression of the chimeric protein.

A recombinant plasmid for RbcS:GFP was constructed as described in Example 1. Isolation of the recombinant plasmid, preparation of protoplasts, and transformation of protoplasts were performed as in Example 2. The protoplast suspension was treated with wortmannin at a concentration of 5 µg/ml. Then the protoplasts were transformed and incubated in the dark. Expression of the chimeric protein was observed as in Example 3.

(b) The effect of wortmannin on the intracellular trafficking of the chimeric protein.

The effect of wortmannin on the trafficking of RbcS:GFP was examined. Wortmannin is known as a specific inhibitor of phosphatidyl 3-phosphate (PI(3)P) and phosphatidylinositol 4-phosphate (PI(4)P) (Ui et al., 1995). Localization of the green fluorescence in the wortmannin-treated protoplast was compared with that of the control protoplast that was not treated with wortmannin. As shown in Figure 8, in contrast to the control protoplast in which the green fluorescence of RbcS:GFP was targeted to the chloroplast as expected (Figure 8(a)), the green fluorescence was not translocated to the chloroplast in the presence of wortmannin, but observed as either speckles or aggregates (Figure 8(b) and 8(c)). This result indicates that wortmannin inhibits trafficking of the chloroplast-targeting protein from the cytosol to the chloroplast.

Example 10. The effect of bafilomycin A1 (BafA1), known as an inhibitor of the vacuolar type H⁺-ATPase, on retrograde trafficking of 500:GFP:KKXX.

(a) Plasmid construction, transformation, and expression of the chimeric protein.

The recombinant plasmid for 500:GFP:KKXX was constructed as described in Example 5. Isolation of the recombinant plasmid, preparation of protoplasts, and transformation of protoplasts were performed as in Example 2. The protoplast suspension was treated with bafilomycin A1 at a concentration of 5 µg/ml. Then the protoplasts were transformed and incubated in the dark. Expression of the chimeric protein was observed as in Example 3.

(b) The effect of bafilomycin A1 (BafA1) on the trafficking of the chimeric protein.

Localization of the green fluorescence in the BafA1-treated protoplast was compared with that of the control protoplast that was not treated with BafA1. As shown in Figure 9, while the green fluorescence was observed as numerous networks in the control protoplast (Figure 9(a) and 9(b)), it was observed as a ring pattern on the plasma

membrane (Figure 9c) and the vacuolar membranes (Figure 9d) in the presence of BafA1. This result indicates that BafA1 inhibits the retrograde trafficking of 500:GFP:KKXX in the *Arabidopsis* protoplast and causes transport of the reporter protein to the plasma membrane or the vacuolar membrane.

5

Example 11. The effect of a specific inhibitor on the localization of two reporter proteins.

(a) Plasmid construction, transformation, and expression of the chimeric protein

Recombinant plasmids for the reporter proteins, BiP:GFP and ST:RFP that are
10 specific to the endoplasmic reticulum and the Golgi apparatus, respectively, were constructed as described in Example 5. Isolation of the recombinant plasmid, preparation of protoplasts, and transformation of protoplasts were performed as in Example 2. The protoplast suspension was treated with brefeldin A at a concentration of 5 µg/ml. Then the protoplasts were transformed and incubated in the dark. Expression
15 of the chimeric protein was observed as in Example 3.

(b) Distribution of fluorescence from the two reporter proteins.

Brefeldin A (BFA) is known as an inhibitor of ADP-ribosylation factors (Arfs) in the animal cell (Morinaga et al., 1999). In the control protoplast, the green
20 fluorescence of BiP:GFP was observed in the endoplasmic reticulum (Figure 10(a)) and the red fluorescence of ST:RFP was observed in the Golgi apparatus (Figure 10(b)). When the protoplast was treated with BFA, ST:GFP was targeted to the endoplasmic reticulum instead of the Golgi apparatus that was destroyed by BFA (Figure 10(e)). As observed, the fluorescent images of the chemical-treated protoplasts are distinctively
25 different from those of the control protoplasts (Figure 10(c) and 10(f)). This suggests that the intracellular trafficking of two or more proteins can be simultaneously detected by using two or more fluorescent proteins with different colors. The result observed for ST:GFP shows the destruction process of the Golgi apparatus.

30 **Example 12.** The effect of brefeldin A (BFA) on the biogenesis and structure conservation of the endoplasmic reticulum.

(a) Plasmid construction, transformation, and expression of the chimeric protein.

A recombinant plasmid for BiP:RFP was constructed as described in Example 5. Isolation of the recombinant plasmid, preparation of protoplasts, and transformation of
35 protoplasts were performed as in Example 2. The protoplast suspension was treated with brefeldin A at a concentration of 5 µg/ml. Then the protoplasts were transformed and

incubated in the dark. Expression of the chimeric protein was observed as in Example 3.

(b) The effect of chemicals on the trafficking of the chimeric protein.

Localization of the red fluorescence in the BFA-treated protoplast was compared with that of the control protoplast that was not treated with BFA. As shown in Figure 11(a), while the red fluorescence of BiP:RFP was present along the structure of the endoplasmic reticulum in the control protoplast, the structure of the ER was observed to be destroyed (Figure 11(c)) when treated with BFA. Effect of using a reporter protein that specifically binds to a subcellular organelle can be clarified by comparing the fluorescence images with the images taken by an optical microscope under bright field (Figure 11(a) and 11(d)).

Example 13. Inhibition of trafficking of proteins that are specific to phospholipids.

(a) Plasmid construction, transformation, and expression of the chimeric protein.

Recombinant plasmids for the reporter proteins, GFP:EBD and GFP:EBDC1358S were constructed as described in Example 7. These recombinant plasmids were used to transform the protoplasts according to the method of Example 2. The transformed protoplasts were treated with wortmannin at a concentration of 1.0 $\mu\text{g/ml}$ or with 2-(4-morpholinyl)-8-phenyl-4H-1-benzopyran-4-one, a specific inhibitor of phosphatidylinositol 3-kinase, at a concentration of 10 $\mu\text{g/ml}$ (LY294002, Vlahos et al., 1994) and incubated at 22°C. Fluorescence images were monitored at various time points.

(b) The effect of chemicals on the localization of the chimeric proteins.

In the control protoplast that was not treated with the chemicals, the green fluorescence was localized in the endosome along the distribution of the EBD-binding phospholipid PI(3)P (Figure 12(a)). In the protoplasts treated with wortmannin or LY294002, however, the green fluorescence was distributed uniformly throughout the cytosol (Figure 12(a)). On the other hand, when the same experiment was performed for GFP:EBDC1358S in which the amino acid residue 1358 is mutated, the chemicals did not affect the distribution of the fluorescence.

Example 14. Change in the expression level induced by cycloheximide, an inhibitor of the protein expression.

(a) Plasmid construction, transformation, and expression of the chimeric protein.

A recombinant plasmid for RA:GFP was constructed as described in Example 1. Isolation of the recombinant plasmid, preparation of protoplasts, and transformation of protoplasts were performed as in Example 2. The protoplast suspension was treated with cycloheximide a concentration of 5 µg/ml. Then the protoplasts were transformed and
5 incubated at 22°C in the dark. Expression of the chimeric protein was observed as in Example 3.

(b) The effect of the chemical on the expression of the chimeric protein.

The intensity of the green fluorescence was decreased in the cycloheximide
10 treated-protoplast (Figure 13(a)) compared to that in the control protoplast that was not treated with cycloheximide (Figure 13(b)). This result shows that the expression of the reporter protein is inhibited, suggesting that the effect of a chemical on the transcription and/or translation of protein can be visualized. The red fluorescence images in Figure 13(c) and 13(d) shows the auto-fluorescence of the chloroplast, which are measured to
15 relatively compare the growth and metabolism of the cells.

The present invention described above is not limited by the aforementioned examples and the attached drawings. The present invention can be substituted, changed, and modified without departing from the technical thoughts described in the
20 specification and the claims, and such substitutions, changes, and modifications fall within the spirit and scope of the present invention.

All references disclosed in this application are incorporated by reference. The following references are specifically incorporated by reference.

25

Patents:

Harpold, M. M., and Brust, P. US 5,401,629. ASSAY METHOD AND
SOMPOSITIONS USEFUL FOR MEASURING THE TRANSDUCTION OF AN
INTRACELLULAR SIGNAL.

30 Thastrup, O., Tullin, S., Poulsen, L. K., and Bjorn, S. P. WO96/23898. A
METHOD OF DETECTING BIOLOGICALLY ACTIVE SUBSTANCES

Ward, W., and Chalfie, M. WO95/21191. BIOLUMINESCENT INDICATOR
BASED UPON THE EXPRESSION OF A GENE FOR A MODIFIED GREEN-
FLUORESCENT PROTEIN.

35 Chalfie, M. and Prasher, D. US 5,491,084. USES OF GREEN-
FLUORESCENT PROTEIN

- Evans, K. WO98/21355. MUTANTS OF GREEN FLUORESCENT PROTEIN.
- Davis, S. J., and Viestra, R. D. (1998). Soluble, highly fluorescent variants of green fluorescent protein (GFP) for use in higher plants. *Plant Mol. Biol.* 36, 521-528.
- Dowler, S., Currie, R. A., Campbell, D. G., Deak, M., Kular, G., Downes, C. P.,
5 and Alessi, D. R. (2000). Identification of pleckstrin-homology-domain-containing proteins with novel phosphoinositide-binding specificities. *Biochem J.* 351(Pt 1), 19-31.
- Gillooly, D. J., Morrow, I. C., Lindsay, M., Gould, R., Bryant, N. J., Gaullier, J-M., Parton, R. G., and Stenmark, H. (2000). Localization of phosphatidylinositol 3-phosphate in yeast and mammalian cells. *EMBO J.* 19, 4577-4588.
- 10 Goldfarb, D. S., Gariepy, J., Schoolnik, G., and Konberg, R. D. (1986). Synthetic peptides as nuclear localization signals. *Nature* 322, 641-644.
- Hu and Cheng (1995). Expression of aequorea green fluorescent protein in plant cells. *FEBS Lett.* 369, 331-334.
- Jiang, L., and Rogers, J. C. (1998). Integral membrane protein sorting to
15 vacuoles in plant cells: evidence for two pathways. *J. Cell Biol.* 143, 1183-1199.
- Jin, J. B., Kim, T. A., Kim, S. J., Lee, S. H., Kim, D. H., Cheong, G.-W., and Hwang, I. (2001). A new Dynamin-like protein, ADL6, is involved in trafficking from the *trans*-Golgi network to the central vacuole in *Arabidopsis*. *Plant Cell* (in press)
- Kaether and Gerdes (1995) Visualization of protein transport along the
20 secretory pathway using green fluorescent protein. *FEBS Lett.* 369, 267-271.
- Kim, D. H., Eu, Y. J., Yoo, C. M., Kim, Y. W., Pih, K. T., Jin, J. B., Kim, S. J., Stenmark, H., and Hwang, I. (2001). Trafficking of Phosphatidylinositol 3-Phosphate from the *trans*-Golgi Network to the Lumen of the Central Vacuole in Plant Cells. *Plant Cell* 13, 287-301.
- 25 Kost, B., Spielhofer, P., and Chua, N. H. (1998). A GFP-mouse talin fusion protein labels plant actin filaments *in vivo* and visualizes the actin cytoskeleton in growing pollen tubes. *Plant J.* 16, 393-401.
- Morinaga, N., Adamik, R., Moss, J., and Vaughan, M. (1999). Brefeldin A inhibited activity of the sec7 domain of p200, a mammalian guanine nucleotide-exchange protein for ADP-ribosylation factors. *J. Biol. Chem.* 274, 17417-17423.
- 30 Morise et al. (1974). Intermolecular energy transfer in the bioluminescent system of aequorea. *Biochem.* 13, 2656-2662.
- Niwa, Y., Hirano, T., Yoshimoto, K., Shimizu, M., and Kobayashi, H. (1999). Non-invasive quantitative detection and applications of non-toxic, S65T-type green
35 fluorescent protein in living plants. *Plant J.* 18, 455-463
- Pih, K. T., Yi, M. J., Liang, Y. S., Shin, B. J., Cho, M. J., Hwang, I., and Son, D.

- (2000). Molecular cloning and targeting of a fibrillarin homolog from *Arabidopsis*. *Plant Physiol.* 123, 51-58.
- Stauffer, T. P., Ahn, S. and Meyer, T. (1988). Receptor-induced transient reduction in plasma membrane PtdIns(4,5)P₂ concentration monitored in living cells. *Curr. Biol.* 8, 343-346.
- 5 Ui, M., Okada, T., Hazeki, K., and Haxeki, O. (1995). Wormannin as a unique probe for an intracellular signaling protein, phosphoinositide 3-kinase. *TIBS* 20, 303-307.
- Vlahos, C., Matter, W. F., Jui, K., and Brown, R. F. (1994). A specific inhibitor
10 of phosphatidylinositol 3-kinase, 2-(4-morpholinyl)-8-phenyl-4H-1-benzopyran-4-one (LY294002). *J. Biol. Chem.* 269, 5241-5248.
- Weiss, E. L., Bishop, A. C., Shokat, K. M., and Drubin, D. G. (2000). Chemical genetic analysis of the budding-yeast p21-activated kinase cla4p. *Nat. Cell Biol.* 2, 677-685.
- 15 Zheng, H., von Mollard, G. F., Kovaleva, V., Stevens, T. H., and Raikhel, N. V. (1999). The plant vesicle-associated SNARE AtVTI1a likely mediates vesicle transport from the trans-Golgi network to the prevacuolar compartment. *Mol. Biol. Cell* 10, 2251-2264.

What is claimed is:

1. A method for detecting trafficking and localization of a selected protein in a cell,
5 which method comprises:
 - (a) preparing a recombinant gene comprising a gene encoding a signal protein which includes a trafficking signal targeting to a specific subcellular organelle and a gene encoding a fluorescent protein linked thereto;
 - (b) preparing a recombinant plasmid including the recombinant gene with a
10 promoter and a terminator operably linked thereto so that the recombinant gene can be expressed in the cell;
 - (c) transforming the cell with at least one recombinant plasmid prepared in step (b);
 - (d) expressing the reporter protein comprising the signal protein and the
15 fluorescent protein in the transformed cell; and
 - (e) detecting the cellular distribution of the reporter protein by monitoring the fluorescence image of the transformed cell during the expression or trafficking, or thereafter.
- 20 2. A method for screening chemicals that affect specific characteristics related to trafficking and localization of a selected protein in a cell, which method comprises:
 - (a) preparing a recombinant gene comprising a gene encoding a signal protein which includes a trafficking signal targeting to a specific subcellular organelle and a gene encoding a fluorescent protein linked thereto;
 - 25 (b) preparing a recombinant plasmid comprising the recombinant gene with a promoter and a terminator operably linked thereto so that the recombinant gene can be expressed in the cell
 - (c) transforming the cell with at least one recombinant plasmid prepared in step (b),
 - 30 (d) expressing the reporter protein comprising the signal protein and the fluorescent protein in the transformed cell, while treating the transformed cell with a chemical before, after, or at the same time as the expression;
 - (e) detecting the cellular distribution of the reporter protein by monitoring the fluorescence image of the transformed cell treated with the chemical
35 during the expression or trafficking, or thereafter; and
 - (f) determining the effect of the chemical by comparing the fluorescence

image obtained in step (e) with that of a control transformed cell which is not treated with the chemical.

3. A method for detecting specific characteristics related to trafficking and localization of a selected protein in a plant cell, which method comprises:
- (a) preparing a recombinant gene comprising a gene encoding a signal protein which includes a trafficking signal targeting to a specific subcellular organelle and a gene encoding a fluorescent protein linked thereto;
 - (b) preparing a recombinant plasmid including the recombinant gene with a promoter and a terminator operably linked thereto so that the recombinant gene can be expressed in the cell;
 - (c) transforming the protoplast of the plant cell whose cell wall is removed with at least one recombinant plasmid prepared in step (b);
 - (d) expressing the reporter protein comprising the signal protein and the fluorescent protein in the transformed cell; and
 - (e) detecting the cellular distribution of the reporter protein by monitoring the fluorescence image of the transformed cell during the expression or trafficking, or thereafter.
4. A method for screening chemicals that affect specific characteristics related to trafficking and localization of a selected protein in a plant cell, which method comprises:
- (a) preparing a recombinant gene comprising a gene encoding a signal protein which includes a trafficking signal targeting to a specific subcellular organelle and a gene encoding a fluorescent protein linked thereto;
 - (b) preparing a recombinant plasmid comprising the recombinant gene with a promoter and a terminator operably linked thereto so that the recombinant gene can be expressed in the cell;
 - (c) transforming the protoplast of the plant cell whose cell wall is removed with at least one recombinant plasmid prepared in step (b);
 - (d) expressing the reporter protein comprising the signal protein and the fluorescent protein in the transformed cell, while treating the transformed cell with a chemical before, after, or at the same time as the expression;
 - (e) detecting the cellular distribution of the reporter protein by monitoring the fluorescence image of the transformed cell treated with the chemical during the expression or trafficking, or thereafter; and

- (f) determining the effect of the chemical by comparing the fluorescence image obtained in step (e) with that of a control transformed cell which is not treated with the chemical.

- 5 5. The method according to any of claims 1 and 2, wherein the cell is a eukaryotic cell.
6. The method according to any of claims 1 to 4, wherein the fluorescent protein is one selected from the group consisting of green fluorescent protein (GFP), red fluorescent protein (RFP), mutants thereof, and derivatives thereof.
- 10 7. The method according to any of claims 1 to 4, wherein the specific subcellular organelle targeted by the reporter protein is one selected from the group consisting of nucleus, mitochondria, chloroplast, and peroxisome.
- 15 8. The method according to any of claims 1 to 4, wherein the coding region of the signal protein included in the recombinant gene is a coding sequence of a full length protein selected from the group consisting of NLS (nuclear localization signal), AtOEP7, Cab (chlorophyll a/b binding protein), SKL (peroxisome targeting motif), RbcS (Rubisco Small Subunit), RA (Rubisco Activase), and F1-H⁺-ATPase, or a portion thereof that
- 20 includes the trafficking signal.
9. The method according to any of claims 1 to 4, wherein the specific subcellular organelle targeted by the reporter protein is one selected from the group consisting of plasma membrane, endoplasmic reticulum, Golgi apparatus, ribosome, lysosome,
- 25 cytoskeleton, centriole, storage vacuole, lytic vacuole, and prevacuolar compartment.
10. The method according to any of claims 1 to 4, wherein the coding region of the signal protein included in the recombinant gene is a coding sequence of a full length protein selected from the group consisting of H⁺-ATPase, BiP (chaperon binding protein), ST (sialyltransferase), Chi (chitinase), clone 526, clone 491, clone 500, AtVTI1a, and SPO (sporamin), or a portion thereof that includes the trafficking signal.
- 30 11. The method according to any of claims 1 to 4, wherein the coding region of the signal protein included in the recombinant gene encodes a protein that binds specifically
- 35 to a phospholipid.

12. The method according to any of claims 1 to 4, wherein the coding region of the signal protein included in the recombinant gene is a coding sequence of a full length protein selected from the group consisting of EBD, AtPH, FAPP1, and PH which bind specifically to phospholipids, or a portion thereof that includes the trafficking signal.
- 5 13. The method according to any of claims 2 and 4, wherein the chemical is one selected from the group consisting of chemical compounds, polypeptides, mixtures of chemical compounds and polypeptides, and extracts of natural products.
- 10 14. The method according to any of claims 1 to 4, wherein the fluorescence image is monitored at one or more fluorescence wavelengths using a fluorescence microscope equipped with optical filters, each of which transmit fluorescence at a specific wavelength.
- 15 15. The method according to any of claims 2 and 4, wherein the chemical inhibits or enhances the trafficking of the reporter protein, so that the trafficking speed or the distribution of the reporter protein in the transformed cell treated with the chemical is altered compared to those in the control transformed cell.
- 20 16. The method according to any of claims 2 and 4, wherein the chemical inhibits or enhances the transcription in which mRNA is produced from the recombinant plasmid or the translation in which the reporter gene is produced from mRNA, so that the intensity and the distribution of the reporter protein in the transformed cell treated with the chemical is altered compared to those in the control transformed cell.
- 25 17. The method according to claims 2 and 4, wherein the chemical causes the cytotoxicity such as deformation, damage, or disruption of the cell or the subcellular organelles, so that the intensity and the distribution of the reporter protein in the transformed cell treated with the chemical is altered compared to those in the control
- 30 transformed cell.
18. The method according to any of claims 1 to 4, wherein the method of introducing the recombinant plasmid into the cell is one selected from the group consisting of the methods of using a chemical such as PEG (polyethylene glycol), calcium phosphate, and DEAE-dextran, lifofection using cationic lipid, microinjection, electroporation, and
- 35 electrofusion.

19. The method according to any of claims 1 to 4, wherein introduction of the recombinant plasmid into the cell is conducted by using 5 to 40% PEG.
- 5 20. The method according to any of claims 1 to 4, wherein observation of the expression, the trafficking, or the localization of the reporter protein is conducted in the time range of 5 min to 80 hrs during or after the incubation of the transformed cell.
- 10 21. A recombinant gene comprising a gene encoding a signal protein which includes a trafficking signal targeting to a specific subcellular organelle and a gene encoding a fluorescent protein linked thereto.
- 15 22. The recombinant gene according to claim 21, wherein the coding region of the signal protein is a coding sequence of a full length protein selected from the group consisting of AtOEP7 (*Arabidopsis* outer membrane envelope), RbcS (Rubisco small subunit), RA (Rubisco activase), Cab (Chlorophyll a/b binding protein), H⁺-ATPase, BiP (Chaperon binding protein), Chi (Chitinase), clone 526, clone 491, clone 500, AtVTI1a, SPO (sporamin), AtPH, and FAPP1, or a portion thereof that includes the trafficking signal.
- 20 23. The recombinant gene according to claim 21, wherein the specific subcellular organelle targeted by the reporter gene expressed from the recombinant gene is one selected from the group consisting of chloroplast, plasma membrane, endoplasmic reticulum, Golgi apparatus, storage vacuole, lytic vacuole, and prevacuolar
- 25 compartment.
- 30 24. The recombinant gene according to claim 21, wherein the fluorescent protein is one selected from the group consisting of green fluorescent protein (GFP), red fluorescent protein (RFP), mutants thereof, and derivatives thereof.
- 35 25. A recombinant plasmid comprising the recombinant gene of any of claims 21 to 24 with a promoter and a terminator operably linked thereto so that the recombinant gene can be expressed in a cell.
- 26 A chimeric protein expressed from the recombinant gene of any of claims 21 to 24.

27. A method of using the chimeric protein of claim 26 as a carrier for targeting chemicals to a subcellular organelle.

28. The recombinant plasmid of claim 25, wherein the promoter is 35S promoter.

5

29. A method for detecting trafficking of a selected protein in a cell, which method comprises expressing in the cell at least one reporter protein comprising at least one signal protein and at least one fluorescent tag and detecting the reporter protein to detect trafficking of the selected protein in the cell.

10

30. A method for detecting a compound that modulates trafficking of a selected protein in a cell, the method comprising expressing in the cell at least one reporter protein comprising at least one signal protein and at least one fluorescent tag, contacting the cell with the compound under conditions conducive to modulating protein trafficking in the cell and detecting the reporter protein to detect the compound.

15

31. The method of any one of claims 1-4, 29 and 30, wherein the reporter protein does not include a Group III signal protein.

20 32. The method of claim 31, wherein the Group III signal protein excluded from the reporter protein is an endosome binding domain (EBD) of human early endosome antigen 1.

25 33. The method of claim 32, wherein the EBD domain consists of amino acids 1257 to 1411 of the human early endosome antigen 1.

34. The method of claims 1-4 and 29-33, wherein the reporter protein does not include a Group II signal protein.

30 35. The method of claim 34, wherein the Group II signal protein excluded from the reporter protein is one of ST; BiP; Sporamin; and H⁺-ATPase.

Names	Schematic presentation	Localization
F1-H ⁺ -ATPase:GFP		Mitochondria
RbcS:GFP		Chloroplast (stroma)
RbcS:RFP		Chloroplast (stroma)
Cab:GFP		Chloroplast (stroma)
RA:GFP		Chloroplast (stroma)
AtOEP7:GFP		Chloroplast (outer envelop membrane)
AtOEP7:RFP		Chloroplast (outer envelop membrane)
NLS:GFP		Nucleus
NLS:RFP		Nucleus
GFP:SKL		Peroxisome

Fig. 1

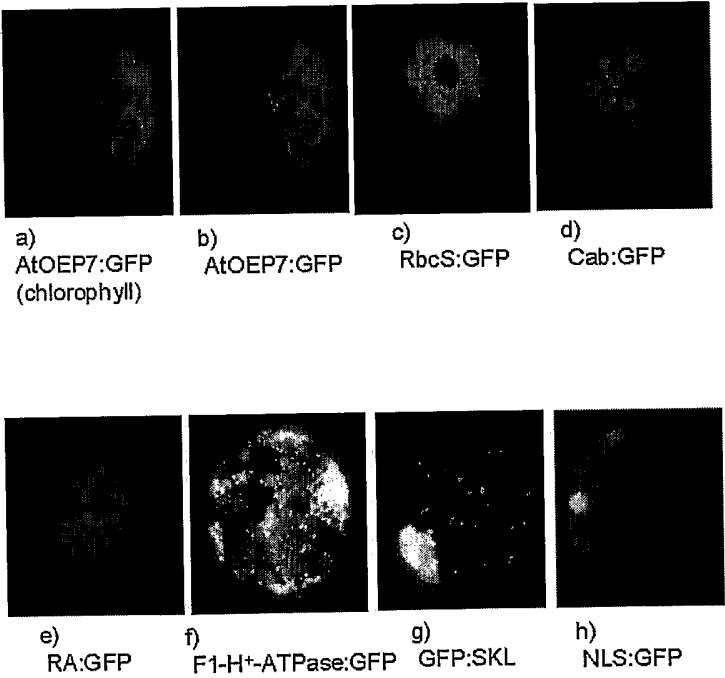


Fig. 2

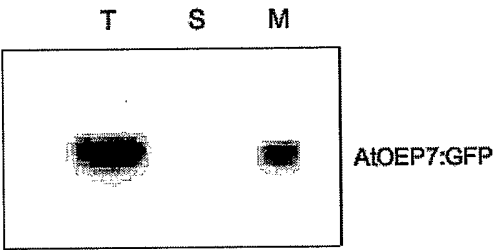


Fig. 3

Name	Schematic presentation	Localization
H ⁺ -ATPase:GFP		Plasma membrane
BiP:GFP		ER (lumen)
BiP:RFP		ER (lumen)
ST:GFP		Golgi apparatus
ST:RFP		Golgi apparatus
Chi-n:RFP:Chi-c		Storage vacuole
Chi-n:RFP		Unknown
500:GFP		Plasma membrane
526:GFP		Storage vacuole membrane
491:GFP		Prevacuolar compartment
500:GFP:KKXX		ER membrane
RFP:AtVTI1a		trans-Golgi network
AtVTI1a:GFP		trans-Golgi network
SPO:GFP		Lytic (central vacuole)

Fig. 4

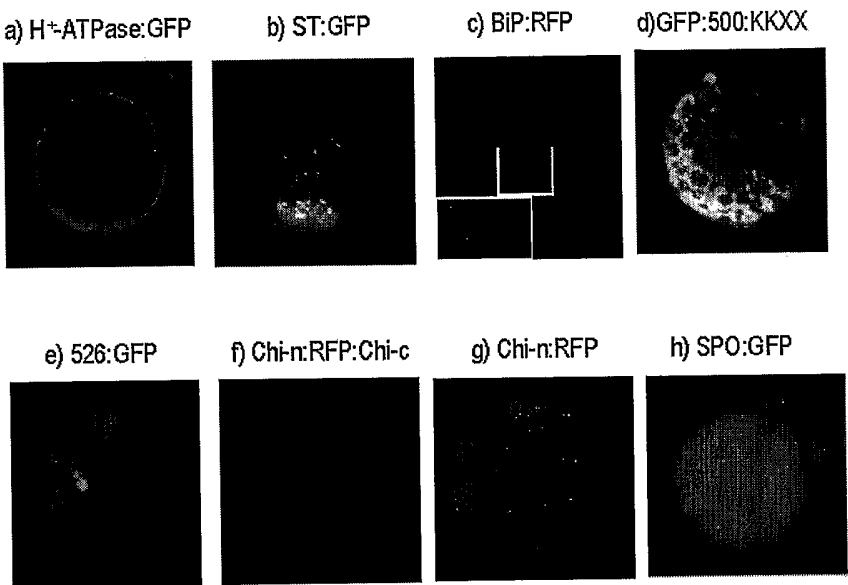


Fig. 5

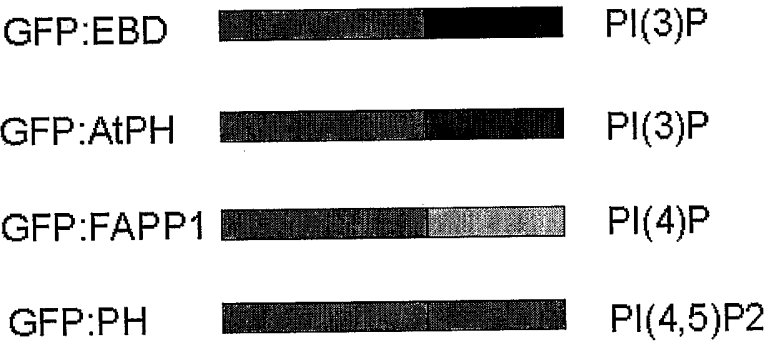


Fig. 6

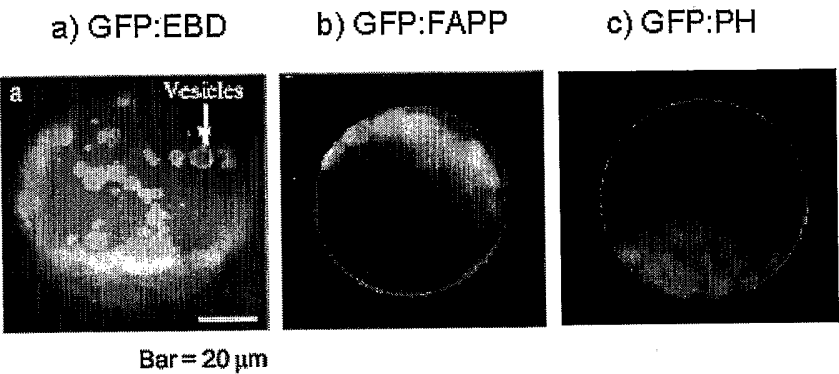


Fig. 7

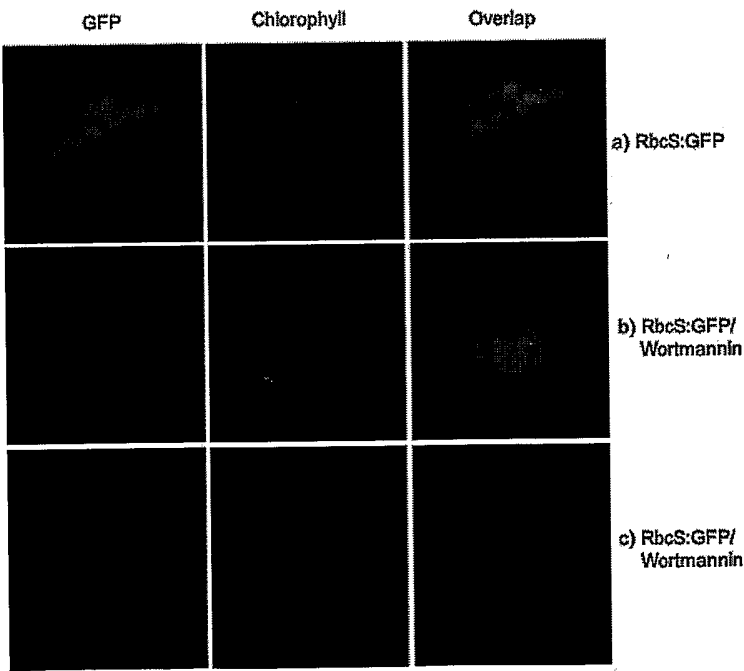


Fig. 8

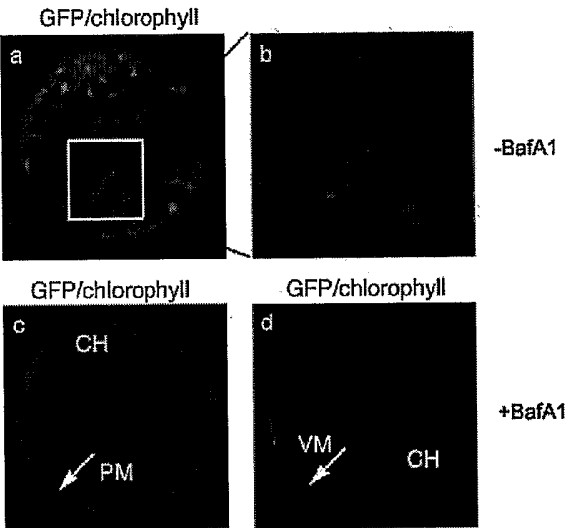


Fig. 9

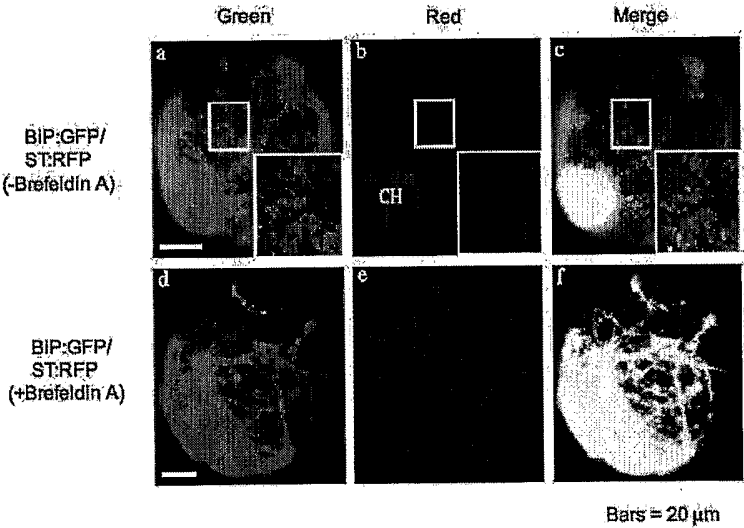


Fig. 10

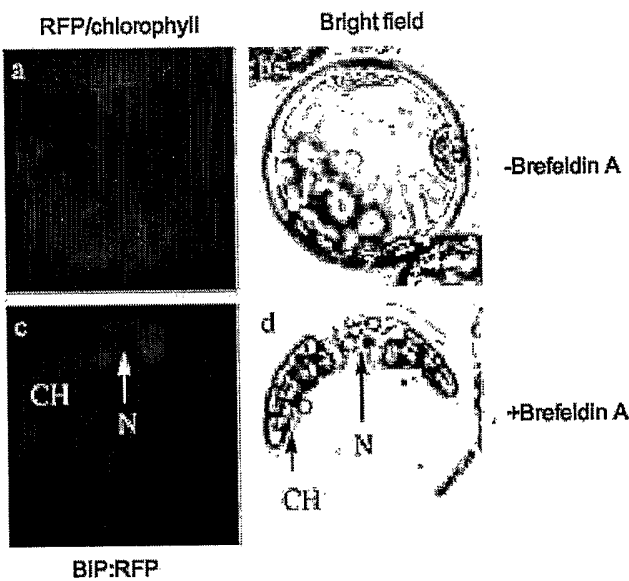


Fig. 11

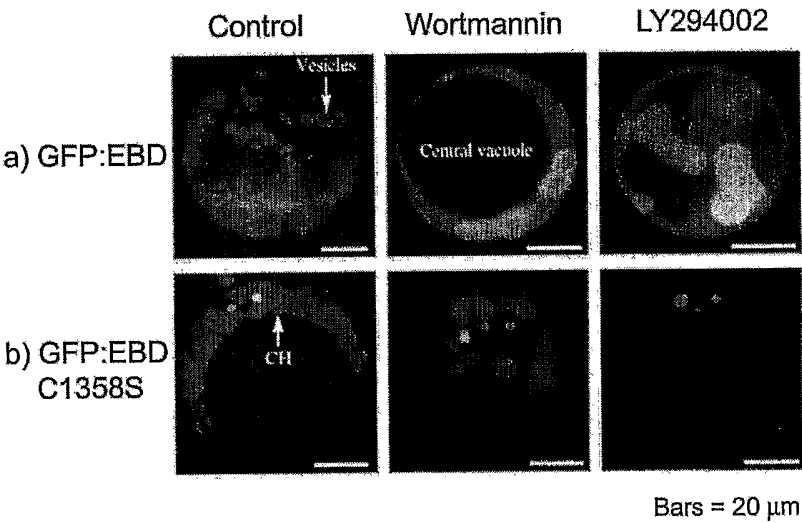


Fig. 12

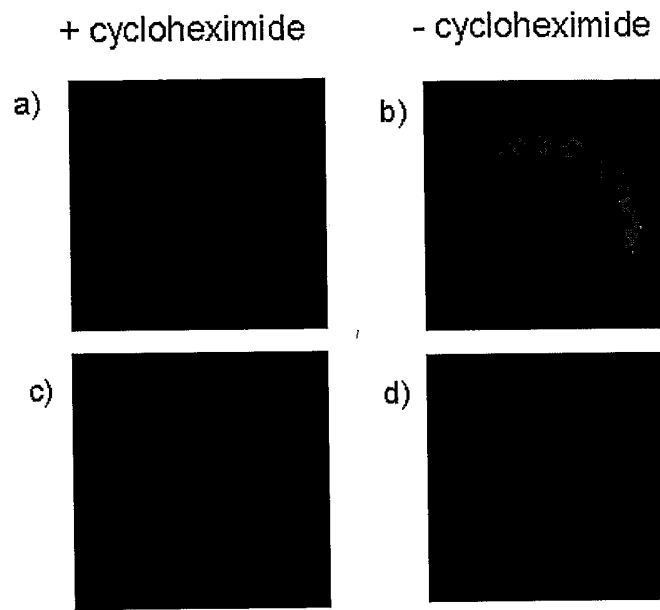


Fig. 13

SEQUENCE LISTING

<110> HWANG, Inhwan
KIM, Dae Heon
LEE, Yong Jik
JIN, Jin Bo

<120> SYSTEM FOR INTRACELLULAR PROCESS MONITORING AND IN VIVO DRUG
SCREENING

<130> APB01/US/PRO

<160> 62

<170> PatentIn version 3.1

<210> 1
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Sequence

<220>
<221> misc_feature
<222> (1)..(18)
<223> 5' PCR primer for partial F1-ATPase coding sequence

<400> 1
ctttaatcaa tggcaatg
18

<210> 2
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Sequence

<220>
<221> misc_feature
<222> (1)..(25)
<223> 3' PCR primer for partial F1-ATPase coding sequence

<400> 2
ccatggcctg aactgctcta agctt
25

<210> 3
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Sequence

<220>
<221> misc_feature
<222> (1)..(18)
<223> 5' PCR primer for partial rubisco small subunit coding

sequence

<400> 3
cctcagtcac acaaagag
18

<210> 4
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Sequence

<220>
<221> misc_feature
<222> (1)..(24)
<223> 3' PCR primer for partial rubisco small subunit coding
sequence

<400> 4
actcgagggga atcggtaagg tcag
24

<210> 5
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Sequence

<220>
<221> misc_feature
<222> (1)..(18)
<223> 5' PCR primer for partial chlorophyll a/b binding protein
coding
sequence

<400> 5
tagagagaaa cgatggcg
18

<210> 6
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Sequence

<220>
<221> misc_feature
<222> (1)..(25)
<223> 3' PCR primer for partial chlorophyll a/b binding protein
coding
sequence

<400> 6

ggatcccgtt tgggagtgga actcc
25

<210> 7
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Sequence

<220>
<221> misc_feature
<222> (1)..(24)
<223> 5' PCR primer for partial rubisco acitvase coding sequence

<400> 7
tctagaatgg ccgccgcagt ttcc
24

<210> 8
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Sequence

<220>
<221> misc_feature
<222> (1)..(25)
<223> 3' PCR primer for partial rubisco acitvase coding sequence

<400> 8
ggatccatct gtctccatcg gtttg
25

<210> 9
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Sequence

<220>
<221> misc_feature
<222> (1)..(18)
<223> 5' PCR primer for partial AtoEP7 coding sequence

<400> 9
gacgacgacg cagcgatg
18

<210> 10
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Sequence

<220>
<221> misc_feature
<222> (1)..(25)
<223> 3' PCR primer for partial AtOEP7 coding sequence

<400> 10
ggatccccaa accctctttg gatgt
25

<210> 11
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Sequence

<220>
<221> misc_feature
<222> (1)..(17)
<223> 5' PCR primer for peroxisome targeting sequence

<400> 11
ccgtatgtta catcacc
17

<210> 12
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Sequence

<220>
<221> misc_feature
<222> (1)..(30)
<223> 3' PCR primer for peroxisome targeting sequence

<400> 12
ttatagcttt gatttgtata gttcatccat
30

<210> 13
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Sequence

<220>
<221> misc_feature
<222> (1)..(18)
<223> 5' PCR primer for H⁺-ATPase coding sequence

<400> 13
gagatgtcga gtctcgaa
18

<210> 14
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Sequence

<220>
<221> misc_feature
<222> (1)..(23)
<223> 3' PCR primer for H⁺-ATPase coding sequence

<400> 14
ctcgagcaca gtgtagtgac tgg
23

<210> 15
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Sequence

<220>
<221> misc_feature
<222> (1)..(18)
<223> 5' PCR primer for luminal binding protein (BiP) coding
sequence

<400> 15
tacgcaaaag tttccgat
18

<210> 16
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Sequence

<220>
<221> misc_feature
<222> (1)..(18)
<223> 3' PCR primer for luminal binding protein (BiP) coding
sequence

<400> 16
ctagagctca tcgtgaga
18

<210> 17
<211> 21

<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Sequence

<220>
<221> misc_feature
<222> (1)..(21)
<223> 5' PCR primer for beta-galactoside alpha 2,6-sialtransferase
codi
ng sequence

<400> 17
atgattcata ccaacttgaa g
21

<210> 18
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Sequence

<220>
<221> misc_feature
<222> (1)..(24)
<223> 3' PCR primer for beta-galactoside alpha 2,6-sialtransferase
codi
ng sequence

<400> 18
ggatccacaa cgaatgttcc ggaa
24

<210> 19
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Sequence

<220>
<221> misc_feature
<222> (1)..(24)
<223> 5' PCR primer for 500:GFP:KKXX coding sequence comprising an
endo
plamic reticulum targeting sequence

<400> 19
ggatcctcta gaggatcgat ccgg
24

<210> 20
<211> 51
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Sequence

<220>
<221> misc_feature
<222> (1)..(51)
<223> 3' PCR primer for 500:GFP:KKXX coding sequence comprising an
endo
plasmic reticulum targeting sequence

<400> 20
ttagatgagt ttcttttttct caaagaaagt ttctaaaagg aatccccctc c
51

<210> 21
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Sequence

<220>
<221> misc_feature
<222> (1)..(25)
<223> 5' PCR primer for partial EBD coding sequence

<400> 21
gaattcgtgg caatctagtc aacgg
25

<210> 22
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Sequence

<220>
<221> misc_feature
<222> (1)..(21)
<223> 3' PCR primer for partial EBD coding sequence

<400> 22
ctaattgtag tgtaatatta c
21

<210> 23
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Sequence

<220>
<221> misc_feature
<222> (1)..(26)
<223> 5' PCR primer for coding sequence of pleckstrin homology

domain o
f Arabidopsis thaliana (AtPH)

<400> 23
cccggaat ggagagtatg tggcga
26

<210> 24
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Sequence

<220>
<221> misc_feature
<222> (1)..(21)
<223> 3' PCR primer for coding sequence of pleckstrin homology
domain o
f Arabidopsis thaliana (AtPH)

<400> 24
taatcaccgc ctgtgatcat a
21

<210> 25
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Sequence

<220>
<221> misc_feature
<222> (1)..(27)
<223> 5' PCR primer for coding sequence of partial pleckstrin
homology
domain of FAPP

<400> 25
ctcgagatgg agggggttct gtacaag
27

<210> 26
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Sequence

<220>
<221> misc_feature
<222> (1)..(24)
<223> 3' PCR primer for coding sequence of partial pleckstrin
homology
domain of FAPP

<400> 26
tcacgctttg gagctcccaa gggc
24

<210> 27
<211> 234
<212> DNA
<213> Arabidopsis thaliana

<220>
<221> CDS
<222> (1)..(234)
<223> Coding sequence for partial gamma subunit of mitochondria F1-ATP
ase comprising a mitochondria targeting sequence

<300>
<308> GenBank/D88374
<309> 1999-02-07

<400> 27
atg gca atg gct gtt ttc cgt cgc gaa ggg agg cgt ctc ctc cct tca
48
Met Ala Met Ala Val Phe Arg Arg Glu Gly Arg Arg Leu Leu Pro Ser
1 5 10 15

atc gcc gct cgc cca atc gct gct atc cga tct cct ctc tct tct gac
96
Ile Ala Ala Arg Pro Ile Ala Ala Ile Arg Ser Pro Leu Ser Ser Asp
20 25 30

cag gag gaa gga ctt ctt gga gtt cga tct atc tca act caa gtg gtg
144
Gln Glu Glu Gly Leu Leu Gly Val Arg Ser Ile Ser Thr Gln Val Val
35 40 45

cgt aac cgc atg aag agt gtt aag aac atc caa aag atc aca aag gca
192
Arg Asn Arg Met Lys Ser Val Lys Asn Ile Gln Lys Ile Thr Lys Ala
50 55 60

atg aag atg gtt gct gct tcc aag ctt aga gca gtt cag gcc
234
Met Lys Met Val Ala Ala Ser Lys Leu Arg Ala Val Gln Ala
65 70 75

<210> 28
<211> 78
<212> PRT
<213> Arabidopsis thaliana

<400> 28

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

Ile Ala Ala Arg Pro Ile Ala Ala Ile Arg Ser Pro Leu Ser Ser Asp
20 25 30

Gln Glu Glu Gly Leu Leu Gly Val Arg Ser Ile Ser Thr Gln Val Val

35

40

45

Arg Asn Arg Met Lys Ser Val Lys Asn Ile Gln Lys Ile Thr Lys Ala
 50 55 60

Met Lys Met Val Ala Ala Ser Lys Leu Arg Ala Val Gln Ala
 65 70 75

<210> 29

<211> 237

<212> DNA

<213> Arabidopsis thaliana

<220>

<221> CDS

<222> (1)..(237)

<223> Coding sequence for partial rubisco small subunit comprising a
ch

loroplast (stroma) targeting sequence

<300>

<308> GenBank/NM_105379

<309> 2002-01-10

<400> 29

atg gct tcc tct atg ctc tct tcc gct act atg gtt gcc tct ccg gct
 48

Met Ala Ser Ser Met Leu Ser Ser Ala Thr Met Val Ala Ser Pro Ala
 1 5 10 15

cag gcc act atg gtc gct cct ttc aac gga ctt aag tcc tcc gct gcc
 96

Gln Ala Thr Met Val Ala Pro Phe Asn Gly Leu Lys Ser Ser Ala Ala
 20 25 30

ttc cca gcc acc cgc aag gct aac aac gac att act tcc atc aca agc
 144

Phe Pro Ala Thr Arg Lys Ala Asn Asn Asp Ile Thr Ser Ile Thr Ser
 35 40 45

aac ggc gga aga gtt aac tgc atg cag gtg tgg cct ccg att gga aag
 192

Asn Gly Gly Arg Val Asn Cys Met Gln Val Trp Pro Pro Ile Gly Lys
 50 55 60

aag aag ttt gag act ctc tct tac ctt cct gac ctt acc gat tcc
 237

Lys Lys Phe Glu Thr Leu Ser Tyr Leu Pro Asp Leu Thr Asp Ser
 65 70 75

<210> 30

<211> 79

<212> PRT

<213> Arabidopsis thaliana

<400> 30

Met Ala Ser Ser Met Leu Ser Ser Ala Thr Met Val Ala Ser Pro Ala
 1 5 10 15

Gln Ala Thr Met Val Ala Pro Phe Asn Gly Leu Lys Ser Ser Ala Ala
 20 25 30

Phe Pro Ala Thr Arg Lys Ala Asn Asn Asp Ile Thr Ser Ile Thr Ser
 35 40 45

Asn Gly Gly Arg Val Asn Cys Met Gln Val Trp Pro Pro Ile Gly Lys
 50 55 60

Lys Lys Phe Glu Thr Leu Ser Tyr Leu Pro Asp Leu Thr Asp Ser
 65 70 75

<210> 31
 <211> 105
 <212> DNA
 <213> Arabidopsis thaliana

 <220>
 <221> CDS
 <222> (1)..(105)
 <223> Coding sequence for partial chlorophyll a/b binding protein
 compr
 ising a chloroplast (stroma) targeting sequence

<300>
 <308> GenBank/X56062
 <309> 1992-10-12

<400> 31
 atg gcg tcg aac tcg ctt atg agc tgt ggc ata gcc gcc gtg tac cct
 48
 Met Ala Ser Asn Ser Leu Met Ser Cys Gly Ile Ala Ala Val Tyr Pro
 1 5 10 15

 tcg ctt ctc tct tct tcc aag tct aaa ttc gta tcc gcc gga gtt cca
 96
 Ser Leu Leu Ser Ser Ser Lys Ser Lys Phe Val Ser Ala Gly Val Pro
 20 25 30

 ctc cca aac
 105
 Leu Pro Asn
 35

<210> 32
 <211> 35
 <212> PRT
 <213> Arabidopsis thaliana

<400> 32

 Met Ala Ser Asn Ser Leu Met Ser Cys Gly Ile Ala Ala Val Tyr Pro
 1 5 10 15

 Ser Leu Leu Ser Ser Ser Lys Ser Lys Phe Val Ser Ala Gly Val Pro
 20 25 30

Leu Pro Asn

35

<210> 33
 <211> 204
 <212> DNA
 <213> Arabidopsis thaliana

 <220>
 <221> CDS
 <222> (1)..(204)
 <223> Coding sequence for partial rubisco activase comprising a
 chlorop
 last (stroma) targeting sequence

<300>
 <308> GenBank/X14212
 <309> 1993-09-12

<400> 33
 atg gcc gcc gca gtt tcc acc gtc ggt gcc atc aac aga gct ccg ttg
 48
 Met Ala Ala Ala Val Ser Thr Val Gly Ala Ile Asn Arg Ala Pro Leu
 1 5 10 15

 agc ttg aac ggg tca gga tca gga gct gta tca gcc cca gct tca acc
 96
 Ser Leu Asn Gly Ser Gly Ser Gly Ala Val Ser Ala Pro Ala Ser Thr
 20 25 30

 ttc ttg gga aag aaa gtt gta act gtg tcg aga ttc gca cag agc aac
 144
 Phe Leu Gly Lys Lys Val Val Thr Val Ser Arg Phe Ala Gln Ser Asn
 35 40 45

 aag aag agc aac gga tca ttc aag gtg ttg gct gtg aaa gaa gac aaa
 192
 Lys Lys Ser Asn Gly Ser Phe Lys Val Leu Ala Val Lys Glu Asp Lys
 50 55 60

 caa acc gat gga
 204
 Gln Thr Asp Gly
 65

<210> 34
 <211> 68
 <212> PRT
 <213> Arabidopsis thaliana

<400> 34

 Met Ala Ala Ala Val Ser Thr Val Gly Ala Ile Asn Arg Ala Pro Leu
 1 5 10 15

 Ser Leu Asn Gly Ser Gly Ser Gly Ala Val Ser Ala Pro Ala Ser Thr
 20 25 30

 Phe Leu Gly Lys Lys Val Val Thr Val Ser Arg Phe Ala Gln Ser Asn
 35 40 45

Lys Lys Ser Asn Gly Ser Phe Lys Val Leu Ala Val Lys Glu Asp Lys
50 55 60

Gln Thr Asp Gly
65

<210> 35
<211> 192
<212> DNA
<213> Arabidopsis thaliana

<220>
<221> CDS
<222> (1)..(192)
<223> Coding sequence for partial AtOEP7, a homologue of pea OEP14,
com
prising a chloroplast (outer envelop membrane) targeting
sequence

<300>
<308> GenBank/NM_115102
<309> 2002-01-10

<400> 35
atg gga aaa act tcg gga gcg aaa cag gcg act gtg gtg gtc gca gcg
48
Met Gly Lys Thr Ser Gly Ala Lys Gln Ala Thr Val Val Val Ala Ala
1 5 10 15

atg gcg tta gga tgg tta gcc ata gag atc gct ttc aag cct ttc ctc
96
Met Ala Leu Gly Trp Leu Ala Ile Glu Ile Ala Phe Lys Pro Phe Leu
20 25 30

gat aaa ttc cgc tcc tca atc gac aaa tct gac cca acc aaa gac ccc
144
Asp Lys Phe Arg Ser Ser Ile Asp Lys Ser Asp Pro Thr Lys Asp Pro
35 40 45

gat gac ttc gac acc gcc gct act gca acc aca tcc aaa gag ggt ttg
192
Asp Asp Phe Asp Thr Ala Ala Thr Ala Thr Thr Ser Lys Glu Gly Leu
50 55 60

<210> 36
<211> 64
<212> PRT
<213> Arabidopsis thaliana

<400> 36

Met Gly Lys Thr Ser Gly Ala Lys Gln Ala Thr Val Val Val Ala Ala
1 5 10 15

Met Ala Leu Gly Trp Leu Ala Ile Glu Ile Ala Phe Lys Pro Phe Leu
20 25 30

Asp Lys Phe Arg Ser Ser Ile Asp Lys Ser Asp Pro Thr Lys Asp Pro
35 40 45

Asp Asp Phe Asp Thr Ala Ala Thr Ala Thr Thr Ser Lys Glu Gly Leu
50 55 60

<210> 37
<211> 54
<212> DNA
<213> SV40

<220>
<221> CDS
<222> (1)..(54)
<223> Coding sequence for SV40 nuclear localization signal (NLS)

<400> 37
atg ggt gct cct cca aaa aag aag aga aag gta gct agt aaa gga gaa
48
Met Gly Ala Pro Pro Lys Lys Lys Arg Lys Val Ala Ser Lys Gly Glu
1 5 10 15
gaa ctt
54
Glu Leu

<210> 38
<211> 18
<212> PRT
<213> SV40

<400> 38

Met Gly Ala Pro Pro Lys Lys Lys Arg Lys Val Ala Ser Lys Gly Glu
1 5 10 15

Glu Leu

<210> 39
<211> 2847
<212> DNA
<213> Arabidopsis thaliana

<220>
<221> gene
<222> (1)..(2847)
<223> Gene for H⁺-ATPase comprising a plasma membrane targeting
sequenc
e

<220>
<221> CDS
<222> (1)..(2844)
<223> H⁺-ATPase comprising a plasma membrane targeting sequence

<300>
<308> GenBank/J05570
<309> 1993-04-27

```

<400> 39
atg tcg agt ctc gaa gat atc aag aac gag act gtt gat ctg gaa aaa
48
Met Ser Ser Leu Glu Asp Ile Lys Asn Glu Thr Val Asp Leu Glu Lys
1 5 10 15

att ccg att gag gaa gtt ttc cag cag cta aaa tgt tca agg gaa gga
96
Ile Pro Ile Glu Glu Val Phe Gln Gln Leu Lys Cys Ser Arg Glu Gly
20 25 30

ttg aca acg cag gaa ggg gag gac agg att cag atc ttt ggc ccc aac
144
Leu Thr Thr Gln Glu Gly Glu Asp Arg Ile Gln Ile Phe Gly Pro Asn
35 40 45

aag ctc gaa gag aaa aag gaa agc aaa ctt ctg aag ttt ttg ggg ttt
192
Lys Leu Glu Glu Lys Lys Glu Ser Lys Leu Leu Lys Phe Leu Gly Phe
50 55 60

atg tgg aat cca ctt tca tgg gtc atg gaa atg gct gca atc atg gcc
240
Met Trp Asn Pro Leu Ser Trp Val Met Glu Met Ala Ala Ile Met Ala
65 70 75 80

att gct ttg gcc aac ggt gat ggt agg cct ccg gat tgg cag gat ttt
288
Ile Ala Leu Ala Asn Gly Asp Gly Arg Pro Pro Asp Trp Gln Asp Phe
85 90 95

gtt ggt att atc tgt ctg ttg gtt atc aac tct acc atc agt ttt atc
336
Val Gly Ile Ile Cys Leu Leu Val Ile Asn Ser Thr Ile Ser Phe Ile
100 105 110

gaa gaa aac aat gct ggt aat gct gct gct gct ctt atg gct ggt ctt
384
Glu Glu Asn Asn Ala Gly Asn Ala Ala Ala Ala Leu Met Ala Gly Leu
115 120 125

gct cct aaa acc aag gtt ctt agg gat gga aag tgg agt gaa caa gaa
432
Ala Pro Lys Thr Lys Val Leu Arg Asp Gly Lys Trp Ser Glu Gln Glu
130 135 140

gct gct att ctt gtc cca gga gat att gtt agc att aaa tta gga gac
480
Ala Ala Ile Leu Val Pro Gly Asp Ile Val Ser Ile Lys Leu Gly Asp
145 150 155 160

att atc cca gct gat gcc cgt cta ctt gaa ggt gat cct tta aag gtt
528
Ile Ile Pro Ala Asp Ala Arg Leu Leu Glu Gly Asp Pro Leu Lys Val
165 170 175

gac caa tct gct cta act gga gag tcc ctt cct gta acc aag cac ccg
576
Asp Gln Ser Ala Leu Thr Gly Glu Ser Leu Pro Val Thr Lys His Pro
180 185 190

ggg caa gaa gtt ttc tct ggt tca acc tgc aaa caa gga gaa atc gag
624
Gly Gln Glu Val Phe Ser Gly Ser Thr Cys Lys Gln Gly Glu Ile Glu
195 200 205

```

gcg gtt gtt att gcc act ggg gtt cat acc ttc ttc ggt aaa gct gct
 672
 Ala Val Val Ile Ala Thr Gly Val His Thr Phe Phe Gly Lys Ala Ala
 210 215 220

cac ctt gtg gac agc act aac caa gtt gga cat ttc cag aag gtt ctt
 720
 His Leu Val Asp Ser Thr Asn Gln Val Gly His Phe Gln Lys Val Leu
 225 230 235 240

aca gcc att ggg aac ttc tgt atc tgt tcc att gct atc ggt atg gtg
 768
 Thr Ala Ile Gly Asn Phe Cys Ile Cys Ser Ile Ala Ile Gly Met Val
 245 250 255

att gag atc atc gtc atg tat ccg atc caa cgc cga aag tac aga gat
 816
 Ile Glu Ile Ile Val Met Tyr Pro Ile Gln Arg Arg Lys Tyr Arg Asp
 260 265 270

gga att gac aac ctt ttg gtc ctc ttg atc ggt ggt atc ccc att gct
 864
 Gly Ile Asp Asn Leu Leu Val Leu Leu Ile Gly Gly Ile Pro Ile Ala
 275 280 285

atg cct aca gtc ttg tcc gtg acc atg gct att ggg tct cac agg ttg
 912
 Met Pro Thr Val Leu Ser Val Thr Met Ala Ile Gly Ser His Arg Leu
 290 295 300

tct cag caa ggt gcc atc acc aag cgt atg act gcc att gaa gag atg
 960
 Ser Gln Gln Gly Ala Ile Thr Lys Arg Met Thr Ala Ile Glu Glu Met
 305 310 315 320

gca gga atg gat gtc ctg tgc agt gac aaa acc ggg aca cta acc ctc
 1008
 Ala Gly Met Asp Val Leu Cys Ser Asp Lys Thr Gly Thr Leu Thr Leu
 325 330 335

aac aaa ttg agt gtg gac aaa aac ttg gtc gag gtt ttc tgc aag ggt
 1056
 Asn Lys Leu Ser Val Asp Lys Asn Leu Val Glu Val Phe Cys Lys Gly
 340 345 350

gtg gag aaa gat caa gtc cta tta ttt gca gct atg gct tcc agg gtt
 1104
 Val Glu Lys Asp Gln Val Leu Leu Phe Ala Ala Met Ala Ser Arg Val
 355 360 365

gag aac cag gat gcc att gat gca gcc atg gtt ggg atg ctt gct gat
 1152
 Glu Asn Gln Asp Ala Ile Asp Ala Ala Met Val Gly Met Leu Ala Asp
 370 375 380

cca aag gag gct aga gct gga atc agg gaa gtt cac ttc ctt cca ttc
 1200
 Pro Lys Glu Ala Arg Ala Gly Ile Arg Glu Val His Phe Leu Pro Phe
 385 390 395 400

aac cct gtg gat aag aga act gct ttg act tac att gac ggc agt ggt
 1248
 Asn Pro Val Asp Lys Arg Thr Ala Leu Thr Tyr Ile Asp Gly Ser Gly
 405 410 415

aac tgg cac aga gtc agt aaa ggt gct cct gag cag atc ctc gaa ctt

1296

Asn Trp His Arg Val Ser Lys Gly Ala Pro Glu Gln Ile Leu Glu Leu
 420 425 430

gcc aaa gcc agc aat gat ctt agc aag aag gtg ctc tcc att att gac
 1344

Ala Lys Ala Ser Asn Asp Leu Ser Lys Lys Val Leu Ser Ile Ile Asp
 435 440 445

aag tat gct gag cgt ggt ctt agg tcg ttg gct gtt gct cgc cag gtg
 1392

Lys Tyr Ala Glu Arg Gly Leu Arg Ser Leu Ala Val Ala Arg Gln Val
 450 455 460

gtg cca gag aaa aca aag gaa agc cca ggt gcg cca tgg gaa ttt gtt
 1440

Val Pro Glu Lys Thr Lys Glu Ser Pro Gly Ala Pro Trp Glu Phe Val
 465 470 475 480

ggc ttg ttg cca ctt ttt gat ccc cca aga cat gac agt gct gaa aca
 1488

Gly Leu Leu Pro Leu Phe Asp Pro Pro Arg His Asp Ser Ala Glu Thr
 485 490 495

att cga cgg gct ttg aat ctt ggt gtt aac gtc aag atg atc act ggt
 1536

Ile Arg Arg Ala Leu Asn Leu Gly Val Asn Val Lys Met Ile Thr Gly
 500 505 510

gac caa ctt gct att ggt aag gaa act ggt cgc aga ctt gga atg gga
 1584

Asp Gln Leu Ala Ile Gly Lys Glu Thr Gly Arg Arg Leu Gly Met Gly
 515 520 525

aca aac atg tat cca tct tcg gct ctt ctt ggt aca cac aaa gac gca
 1632

Thr Asn Met Tyr Pro Ser Ser Ala Leu Leu Gly Thr His Lys Asp Ala
 530 535 540

aac ctc gca tcc att cct gtt gag gag ttg att gaa aag gct gat gga
 1680

Asn Leu Ala Ser Ile Pro Val Glu Glu Leu Ile Glu Lys Ala Asp Gly
 545 550 555 560

ttt gcc gga gtc ttc cca gag cac aaa tac gaa att gtg aaa aag ttg
 1728

Phe Ala Gly Val Phe Pro Glu His Lys Tyr Glu Ile Val Lys Lys Leu
 565 570 575

cag gag agg aag cat att gtt gga atg act ggt gat ggt gtc aat gat
 1776

Gln Glu Arg Lys His Ile Val Gly Met Thr Gly Asp Gly Val Asn Asp
 580 585 590

gcc cct gct cta aag aaa gct gat atc ggt att gct gtt gct gat gct
 1824

Ala Pro Ala Leu Lys Lys Ala Asp Ile Gly Ile Ala Val Ala Asp Ala
 595 600 605

aca gat gct gct cgt ggt gct tca gat atc gtg ctc act gag cct gga
 1872

Thr Asp Ala Ala Arg Gly Ala Ser Asp Ile Val Leu Thr Glu Pro Gly
 610 615 620

ctc agc gtt att atc agt gct gtt ctc acc agc aga gct att ttc cag
 1920

Leu Ser Val Ile Ile Ser Ala Val Leu Thr Ser Arg Ala Ile Phe Gln
 625 630 635 640
 aga atg aag aac tat act atc tat gca gtc tca atc acc atc cgt att
 1968
 Arg Met Lys Asn Tyr Thr Ile Tyr Ala Val Ser Ile Thr Ile Arg Ile
 645 650 655
 gtg ttt ggt ttc atg ctt att gct ttg ata tgg gaa ttt gac ttc tca
 2016
 Val Phe Gly Phe Met Leu Ile Ala Leu Ile Trp Glu Phe Asp Phe Ser
 660 665 670
 gcc ttc atg gtt ctg atc att gcc att ctt aac gac ggt acc atc atg
 2064
 Ala Phe Met Val Leu Ile Ile Ala Ile Leu Asn Asp Gly Thr Ile Met
 675 680 685
 aca atc tca aag gac aga gtt aag cca tct ccc aca cct gat agc tgg
 2112
 Thr Ile Ser Lys Asp Arg Val Lys Pro Ser Pro Thr Pro Asp Ser Trp
 690 695 700
 aaa ctt aaa gaa att ttt gct act gga gtc gtt cta gga ggc tac cag
 2160
 Lys Leu Lys Glu Ile Phe Ala Thr Gly Val Val Leu Gly Gly Tyr Gln
 705 710 715 720
 gcc atc atg act gtt att ttc ttc tgg gcg gcg cac aag act gac ttt
 2208
 Ala Ile Met Thr Val Ile Phe Phe Trp Ala Ala His Lys Thr Asp Phe
 725 730 735
 ttc tcg gac aca ttc ggt gtg agg tcc att agg gac aat aac cac gag
 2256
 Phe Ser Asp Thr Phe Gly Val Arg Ser Ile Arg Asp Asn Asn His Glu
 740 745 750
 cta atg ggt gcg gtg tac tta caa gtt agt atc att agt caa gct ctg
 2304
 Leu Met Gly Ala Val Tyr Leu Gln Val Ser Ile Ile Ser Gln Ala Leu
 755 760 765
 atc ttc gtc aca aga tca agg agt tgg tct ttt gtt gaa cgt cct gga
 2352
 Ile Phe Val Thr Arg Ser Arg Ser Trp Ser Phe Val Glu Arg Pro Gly
 770 775 780
 gca ttg ctg atg att gct ttc ctc att gca caa ctg att gct act ttg
 2400
 Ala Leu Leu Met Ile Ala Phe Leu Ile Ala Gln Leu Ile Ala Thr Leu
 785 790 795 800
 att gcg gtt tac gcc aac tgg gaa ttt gca aag att agg ggt att gga
 2448
 Ile Ala Val Tyr Ala Asn Trp Glu Phe Ala Lys Ile Arg Gly Ile Gly
 805 810 815
 tgg gga tgg gct ggt gtg atc tgg cta tac agt att gtc aca tac ttc
 2496
 Trp Gly Trp Ala Gly Val Ile Trp Leu Tyr Ser Ile Val Thr Tyr Phe
 820 825 830
 cca ttg gac gtt ttc aag ttt gcc att cga tac atc ttg agc gga aag
 2544
 Pro Leu Asp Val Phe Lys Phe Ala Ile Arg Tyr Ile Leu Ser Gly Lys

835 840 845
 gcg tgg ctc aac ttg ttt gag aac aag acg gct ttc acg atg aag aaa
 2592
 Ala Trp Leu Asn Leu Phe Glu Asn Lys Thr Ala Phe Thr Met Lys Lys
 850 855 860
 gat tac gga aaa gaa gag aga gag gct caa tgg gca ctt gct caa agg
 2640
 Asp Tyr Gly Lys Glu Glu Arg Glu Ala Gln Trp Ala Leu Ala Gln Arg
 865 870 875 880
 aca ctt cac ggt tta cag cca aaa gaa gct gtt aac atc ttc cct gag
 2688
 Thr Leu His Gly Leu Gln Pro Lys Glu Ala Val Asn Ile Phe Pro Glu
 885 890 895
 aaa gga agt tac aga gaa ttg tct gag atc gct gag caa gct aag aga
 2736
 Lys Gly Ser Tyr Arg Glu Leu Ser Glu Ile Ala Glu Gln Ala Lys Arg
 900 905 910
 aga gct gag atc gct agg ctt agg gag ctg cac aca ctc aag gga cat
 2784
 Arg Ala Glu Ile Ala Arg Leu Arg Glu Leu His Thr Leu Lys Gly His
 915 920 925
 gtg gaa tca gtc gtg aag cta aag ggc ttg gac att gaa act ccc agt
 2832
 Val Glu Ser Val Val Lys Leu Lys Gly Leu Asp Ile Glu Thr Pro Ser
 930 935 940
 cac tac act gtg tag
 2847
 His Tyr Thr Val
 945

 <210> 40
 <211> 948
 <212> PRT
 <213> Arabidopsis thaliana

 <400> 40

 Met Ser Ser Leu Glu Asp Ile Lys Asn Glu Thr Val Asp Leu Glu Lys
 1 5 10 15

 Ile Pro Ile Glu Glu Val Phe Gln Gln Leu Lys Cys Ser Arg Glu Gly
 20 25 30

 Leu Thr Thr Gln Glu Gly Glu Asp Arg Ile Gln Ile Phe Gly Pro Asn
 35 40 45

 Lys Leu Glu Glu Lys Lys Glu Ser Lys Leu Leu Lys Phe Leu Gly Phe
 50 55 60

 Met Trp Asn Pro Leu Ser Trp Val Met Glu Met Ala Ala Ile Met Ala
 65 70 75 80

 Ile Ala Leu Ala Asn Gly Asp Gly Arg Pro Pro Asp Trp Gln Asp Phe

Val	Gly	Ile	Ile	Cys	Leu	Leu	Val	Ile	Asn	Ser	Thr	Ile	Ser	Phe	Ile
			100					105					110		
Glu	Glu	Asn	Asn	Ala	Gly	Asn	Ala	Ala	Ala	Ala	Leu	Met	Ala	Gly	Leu
		115					120					125			
Ala	Pro	Lys	Thr	Lys	Val	Leu	Arg	Asp	Gly	Lys	Trp	Ser	Glu	Gln	Glu
	130					135					140				
Ala	Ala	Ile	Leu	Val	Pro	Gly	Asp	Ile	Val	Ser	Ile	Lys	Leu	Gly	Asp
	145				150					155					160
Ile	Ile	Pro	Ala	Asp	Ala	Arg	Leu	Leu	Glu	Gly	Asp	Pro	Leu	Lys	Val
				165					170					175	
Asp	Gln	Ser	Ala	Leu	Thr	Gly	Glu	Ser	Leu	Pro	Val	Thr	Lys	His	Pro
			180					185					190		
Gly	Gln	Glu	Val	Phe	Ser	Gly	Ser	Thr	Cys	Lys	Gln	Gly	Glu	Ile	Glu
		195					200					205			
Ala	Val	Val	Ile	Ala	Thr	Gly	Val	His	Thr	Phe	Phe	Gly	Lys	Ala	Ala
	210					215					220				
His	Leu	Val	Asp	Ser	Thr	Asn	Gln	Val	Gly	His	Phe	Gln	Lys	Val	Leu
225					230					235					240
Thr	Ala	Ile	Gly	Asn	Phe	Cys	Ile	Cys	Ser	Ile	Ala	Ile	Gly	Met	Val
				245					250					255	
Ile	Glu	Ile	Ile	Val	Met	Tyr	Pro	Ile	Gln	Arg	Arg	Lys	Tyr	Arg	Asp
			260					265					270		
Gly	Ile	Asp	Asn	Leu	Leu	Val	Leu	Leu	Ile	Gly	Gly	Ile	Pro	Ile	Ala
		275					280					285			
Met	Pro	Thr	Val	Leu	Ser	Val	Thr	Met	Ala	Ile	Gly	Ser	His	Arg	Leu
	290					295					300				
Ser	Gln	Gln	Gly	Ala	Ile	Thr	Lys	Arg	Met	Thr	Ala	Ile	Glu	Glu	Met
305					310					315					320
Ala	Gly	Met	Asp	Val	Leu	Cys	Ser	Asp	Lys	Thr	Gly	Thr	Leu	Thr	Leu
				325					330					335	
Asn	Lys	Leu	Ser	Val	Asp	Lys	Asn	Leu	Val	Glu	Val	Phe	Cys	Lys	Gly
				340				345					350		

Val Glu Lys Asp Gln Val Leu Leu Phe Ala Ala Met Ala Ser Arg Val
 355 360 365

Glu Asn Gln Asp Ala Ile Asp Ala Ala Met Val Gly Met Leu Ala Asp
 370 375 380

Pro Lys Glu Ala Arg Ala Gly Ile Arg Glu Val His Phe Leu Pro Phe
 385 390 395 400

Asn Pro Val Asp Lys Arg Thr Ala Leu Thr Tyr Ile Asp Gly Ser Gly
 405 410 415

Asn Trp His Arg Val Ser Lys Gly Ala Pro Glu Gln Ile Leu Glu Leu
 420 425 430

Ala Lys Ala Ser Asn Asp Leu Ser Lys Lys Val Leu Ser Ile Ile Asp
 435 440 445

Lys Tyr Ala Glu Arg Gly Leu Arg Ser Leu Ala Val Ala Arg Gln Val
 450 455 460

Val Pro Glu Lys Thr Lys Glu Ser Pro Gly Ala Pro Trp Glu Phe Val
 465 470 475 480

Gly Leu Leu Pro Leu Phe Asp Pro Pro Arg His Asp Ser Ala Glu Thr
 485 490 495

Ile Arg Arg Ala Leu Asn Leu Gly Val Asn Val Lys Met Ile Thr Gly
 500 505 510

Asp Gln Leu Ala Ile Gly Lys Glu Thr Gly Arg Arg Leu Gly Met Gly
 515 520 525

Thr Asn Met Tyr Pro Ser Ser Ala Leu Leu Gly Thr His Lys Asp Ala
 530 535 540

Asn Leu Ala Ser Ile Pro Val Glu Glu Leu Ile Glu Lys Ala Asp Gly
 545 550 555 560

Phe Ala Gly Val Phe Pro Glu His Lys Tyr Glu Ile Val Lys Lys Leu
 565 570 575

Gln Glu Arg Lys His Ile Val Gly Met Thr Gly Asp Gly Val Asn Asp
 580 585 590

Ala Pro Ala Leu Lys Lys Ala Asp Ile Gly Ile Ala Val Ala Asp Ala
 595 600 605

Thr Asp Ala Ala Arg Gly Ala Ser Asp Ile Val Leu Thr Glu Pro Gly

610		615		620
Leu Ser Val Ile Ile Ser Ala Val Leu Thr Ser Arg Ala Ile Phe Gln				
625		630		635
Arg Met Lys Asn Tyr Thr Ile Tyr Ala Val Ser Ile Thr Ile Arg Ile				
		645		650
Val Phe Gly Phe Met Leu Ile Ala Leu Ile Trp Glu Phe Asp Phe Ser				
		660		665
Ala Phe Met Val Leu Ile Ile Ala Ile Leu Asn Asp Gly Thr Ile Met				
		675		680
Thr Ile Ser Lys Asp Arg Val Lys Pro Ser Pro Thr Pro Asp Ser Trp				
		690		695
Lys Leu Lys Glu Ile Phe Ala Thr Gly Val Val Leu Gly Gly Tyr Gln				
		705		710
Ala Ile Met Thr Val Ile Phe Phe Trp Ala Ala His Lys Thr Asp Phe				
		725		730
Phe Ser Asp Thr Phe Gly Val Arg Ser Ile Arg Asp Asn Asn His Glu				
		740		745
Leu Met Gly Ala Val Tyr Leu Gln Val Ser Ile Ile Ser Gln Ala Leu				
		755		760
Ile Phe Val Thr Arg Ser Arg Ser Trp Ser Phe Val Glu Arg Pro Gly				
		770		775
Ala Leu Leu Met Ile Ala Phe Leu Ile Ala Gln Leu Ile Ala Thr Leu				
		785		790
Ile Ala Val Tyr Ala Asn Trp Glu Phe Ala Lys Ile Arg Gly Ile Gly				
		805		810
Trp Gly Trp Ala Gly Val Ile Trp Leu Tyr Ser Ile Val Thr Tyr Phe				
		820		825
Pro Leu Asp Val Phe Lys Phe Ala Ile Arg Tyr Ile Leu Ser Gly Lys				
		835		840
Ala Trp Leu Asn Leu Phe Glu Asn Lys Thr Ala Phe Thr Met Lys Lys				
		850		855
Asp Tyr Gly Lys Glu Glu Arg Glu Ala Gln Trp Ala Leu Ala Gln Arg				
		865		870
				875
				880

Thr Leu His Gly Leu Gln Pro Lys Glu Ala Val Asn Ile Phe Pro Glu
885 890 895

Lys Gly Ser Tyr Arg Glu Leu Ser Glu Ile Ala Glu Gln Ala Lys Arg
900 905 910

Arg Ala Glu Ile Ala Arg Leu Arg Glu Leu His Thr Leu Lys Gly His
915 920 925

Val Glu Ser Val Val Lys Leu Lys Gly Leu Asp Ile Glu Thr Pro Ser
930 935 940

His Tyr Thr Val
945

<210> 41

<211> 132

<212> DNA

<213> Arabidopsis thaliana

<220>

<221> CDS

<222> (1)..(132)

<223> Coding sequence for amino terminal region of luminal binding
prot

ein (BiP) comprising an endoplasmic reticulum (lumen)
targeting s
equence or its portion

<300>

<308> GenBank/D84414, D82817

<309> 1999-02-06

<400> 41

atg gct cgc tcg ttt gga gca aac agc act gtt gtg ctt gcg atc atc
48

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

ttc ttc gga tgt tta ttt gcg ttt tcc act gca aaa gaa gag gct acg
96

Phe Phe Gly Cys Leu Phe Ala Phe Ser Thr Ala Lys Glu Glu Ala Thr
20 25 30

aag tta gga tca gtt att ggg ata gat ctt ggt acc
132

Lys Leu Gly Ser Val Ile Gly Ile Asp Leu Gly Thr
35 40

<210> 42

<211> 44

<212> PRT

<213> Arabidopsis thaliana

<400> 42

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

Phe Phe Gly Cys Leu Phe Ala Phe Ser Thr Ala Lys Glu Glu Ala Thr
20 25 30

Lys Leu Gly Ser Val Ile Gly Ile Asp Leu Gly Thr
35 40

```
<210> 43
<211> 243
<212> DNA
<213> Arabidopsis thaliana
```

```

<220>
<221>  gene
<222>  (1)..(243)
<223>  Gene for carboxyl terminal region of luminal binding protein
(BiP
      ) comprising an endoplasmic reticulum (lumen) targeting
sequence
      or its portion

```

```
<220>
<221> CDS
<222> (1)..(240)
<223> Carboxyl terminal region of luminal binding protein (BiP)
compris
ing an endoplasmic reticulum (lumen) targeting sequence or its
po
rtion
```

```
<300>
<308> GenBank/D84414, D82817
<309> 1999-02-06
```

<400> 43
aag ctt gca gac aaa ttg gaa gga gat gag aag gag aag ata gaa gca
48
Lys Leu Ala Asp Lys Leu Glu Gly Asp Glu Lys Glu Lys Ile Glu Ala
1 5 10 15

gcg acg aaa gag gcc ttg gag tgg ctc gac gag aac caa aac tca gag
96
Ala Thr Lys Glu Ala Leu Glu Trp Leu Asp Glu Asn Gln Asn Ser Glu
20 25 30

aaa gaa gag tac gac gag aag ctc aag gag gta gag gca gtg tgt aac
144
Lys Glu Glu Tyr Asp Glu Lys Leu Lys Glu Val Glu Ala Val Cys Asn
35 40 45

cca atc atc aca gca gtt tac cag aga tca ggc gga gca cca ggt gca
192
Pro Ile Ile Thr Ala Val Tyr Gln Arg Ser Gly Gly Ala Pro Gly Ala
50 55 60

gga gga gaa tca tcg act gag gag gaa gat gag tct cac gat gag ctc
240
Gly Gly Glu Ser Ser Thr Glu Glu Glu Asp Glu Ser His Asp Glu Leu
65 70 75 80

tag

243

<210> 44
<211> 80
<212> PRT
<213> Arabidopsis thaliana

<400> 44

Lys Leu Ala Asp Lys Leu Glu Gly Asp Glu Lys Glu Lys Ile Glu Ala
1 5 10 15

Ala Thr Lys Glu Ala Leu Glu Trp Leu Asp Glu Asn Gln Asn Ser Glu
20 25 30

Lys Glu Glu Tyr Asp Glu Lys Leu Lys Glu Val Glu Ala Val Cys Asn
35 40 45

Pro Ile Ile Thr Ala Val Tyr Gln Arg Ser Gly Gly Ala Pro Gly Ala
50 55 60

Gly Gly Glu Ser Ser Thr Glu Glu Glu Asp Glu Ser His Asp Glu Leu
65 70 75 80

<210> 45
<211> 1212
<212> DNA
<213> Rattus norvegicus

<220>
<221> gene
<222> (1)..(1212)
<223> Gene for beta-galactoside alpha 2,6-sialyltransferase
comprising
a Golgi apparatus targeting sequence

<220>
<221> CDS
<222> (1)..(1209)
<223> Beta-galactoside alpha 2,6-sialyltransferase comprising a
Golgi a
pparatus targeting sequence

<300>
<308> GenBank/M18769
<309> 1993-04-27

<400> 45
atg att cat acc aac ttg aag aaa aag ttc agc ctc ttc atc ctg gtc
48
Met Ile His Thr Asn Leu Lys Lys Lys Phe Ser Leu Phe Ile Leu Val
1 5 10 15
ttt ctc ctg ttc gca gtc atc tgt gtt tgg aag aaa ggg agc gac tat
96
Phe Leu Leu Phe Ala Val Ile Cys Val Trp Lys Lys Gly Ser Asp Tyr
20 25 30

gag gcc ctt aca ctg caa gcc aag gaa ttc cag atg ccc aag agc cag
 144
 Glu Ala Leu Thr Leu Gln Ala Lys Glu Phe Gln Met Pro Lys Ser Gln
 35 40 45

gag aaa gtg gcc atg ggg tct gct tcc cag gtt gtg ttc tca aac agc
 192
 Glu Lys Val Ala Met Gly Ser Ala Ser Gln Val Val Phe Ser Asn Ser
 50 55 60

aag caa gac cct aag gaa gac att cca atc ctc agt tac cac agg gtc
 240
 Lys Gln Asp Pro Lys Glu Asp Ile Pro Ile Leu Ser Tyr His Arg Val
 65 70 75 80

aca gcc aag gtc aaa cca cag cct tcc ttc cag gtg tgg gac aag gac
 288
 Thr Ala Lys Val Lys Pro Gln Pro Ser Phe Gln Val Trp Asp Lys Asp
 85 90 95

tcc aca tac tca aaa ctt aac ccc agg ctg ctg aag atc tgg aga aac
 336
 Ser Thr Tyr Ser Lys Leu Asn Pro Arg Leu Leu Lys Ile Trp Arg Asn
 100 105 110

tat ctg aac atg aac aaa tat aaa gta tcc tac aag gga ccg ggg cca
 384
 Tyr Leu Asn Met Asn Lys Tyr Lys Val Ser Tyr Lys Gly Pro Gly Pro
 115 120 125

gga gtc aag ttc agc gta gaa gca ctg cgt tgc cac ctt cga gac cat
 432
 Gly Val Lys Phe Ser Val Glu Ala Leu Arg Cys His Leu Arg Asp His
 130 135 140

gtg aac gtg tct atg ata gag gcc aca gat ttt ccc ttc aac acc act
 480
 Val Asn Val Ser Met Ile Glu Ala Thr Asp Phe Pro Phe Asn Thr Thr
 145 150 155 160

gag tgg gag ggt tac ctg ccc aag gag aac ttt aga acc aag gtt ggg
 528
 Glu Trp Glu Gly Tyr Leu Pro Lys Glu Asn Phe Arg Thr Lys Val Gly
 165 170 175

cct tgg caa agg tgt gcc gtc gtc tct tct gca gga tct ctg aaa aac
 576
 Pro Trp Gln Arg Cys Ala Val Val Ser Ser Ala Gly Ser Leu Lys Asn
 180 185 190

tcc cag ctt ggt cga gag att gat aat cat gat gca gtt ctg agg ttt
 624
 Ser Gln Leu Gly Arg Glu Ile Asp Asn His Asp Ala Val Leu Arg Phe
 195 200 205

aat ggg gcc cct acc gac aac ttc caa cag gat gtg ggc tca aaa act
 672
 Asn Gly Ala Pro Thr Asp Asn Phe Gln Gln Asp Val Gly Ser Lys Thr
 210 215 220

acc att cgc cta atg aac tct cag tta gtc acc aca gaa aag cgc ttc
 720
 Thr Ile Arg Leu Met Asn Ser Gln Leu Val Thr Thr Glu Lys Arg Phe
 225 230 235 240

ctc aag gac agt ttg tac acc gaa gga atc cta att gta tgg gac cca

tcc gtg tat cat gca gat atc cca aag tgg tat cag aaa cca gac tac
816

aat ttc ttc gaa acc tat aag agt tac cga agg ctg aac ccc agc cag
864

cca ttt tat atc ctc aag ccc cag atg cca tgg gaa ctg tgg gac atc
912

att cag gaa atc tct gca gat ctg att cag cca aat ccc cca tcc tcc
960

ggc atg ctg ggt atc atc atc atg atg acg ctg tgt gac cag gta gat
1008

att tac gag ttc ctc cca tcc aag cgc aag acg gac gtg tgc tat tat
1056

cac caa aag ttc ttt gac agc gct tgc acg atg ggt gcc tac gac ccg
1104

ctc ctc ttc gag aag aat atg gtg aag cat ctc aat gag gga aca gat
1152

Leu Leu Phe Glu Lys Asn Met Val Lys His Leu Asn Glu Gly Thr Asp
370 375 380

gaa gac att tat ttg ttt ggg aaa gcc acc ctt tct ggc ttc cgg aac
1200

Glu Asp Ile Tyr Leu Phe Gly Lys Ala Thr Leu Ser Gly Phe Arg Asn
 385 390 395 400

att cgt tgt tga

1212

Ile Arg Cys

<210> 46

<211> 403

<212> PRT

<213> Rattus norvegicus

<400> 46

Met Ile His Thr Asn Leu Lys Lys Lys Phe Ser Leu Phe Ile Leu Val
1 5 10 15

Phe Leu Leu Phe Ala Val Ile Cys Val Trp Lys Lys Gly Ser Asp Tyr
 20 25 30

Glu Ala Leu Thr Leu Gln Ala Lys Glu Phe Gln Met Pro Lys Ser Gln
 35 40 45

Glu Lys Val Ala Met Gly Ser Ala Ser Gln Val Val Phe Ser Asn Ser
 50 55 60

Lys Gln Asp Pro Lys Glu Asp Ile Pro Ile Leu Ser Tyr His Arg Val
 65 70 75 80

Thr Ala Lys Val Lys Pro Gln Pro Ser Phe Gln Val Trp Asp Lys Asp
 85 90 95

Ser Thr Tyr Ser Lys Leu Asn Pro Arg Leu Leu Lys Ile Trp Arg Asn
 100 105 110

Tyr Leu Asn Met Asn Lys Tyr Lys Val Ser Tyr Lys Gly Pro Gly Pro
 115 120 125

Gly Val Lys Phe Ser Val Glu Ala Leu Arg Cys His Leu Arg Asp His
 130 135 140

Val Asn Val Ser Met Ile Glu Ala Thr Asp Phe Pro Phe Asn Thr Thr
 145 150 155 160

Glu Trp Glu Gly Tyr Leu Pro Lys Glu Asn Phe Arg Thr Lys Val Gly
 165 170 175

Pro Trp Gln Arg Cys Ala Val Val Ser Ser Ala Gly Ser Leu Lys Asn
 180 185 190

Ser Gln Leu Gly Arg Glu Ile Asp Asn His Asp Ala Val Leu Arg Phe
 195 200 205

Asn Gly Ala Pro Thr Asp Asn Phe Gln Gln Asp Val Gly Ser Lys Thr
 210 215 220

Thr Ile Arg Leu Met Asn Ser Gln Leu Val Thr Thr Glu Lys Arg Phe
 225 230 235 240

Leu Lys Asp Ser Leu Tyr Thr Glu Gly Ile Leu Ile Val Trp Asp Pro
 245 250 255

Ser Val Tyr His Ala Asp Ile Pro Lys Trp Tyr Gln Lys Pro Asp Tyr
 260 265 270

Asn Phe Phe Glu Thr Tyr Lys Ser Tyr Arg Arg Leu Asn Pro Ser Gln
 275 280 285

Pro Phe Tyr Ile Leu Lys Pro Gln Met Pro Trp Glu Leu Trp Asp Ile
 290 295 300

Ile Gln Glu Ile Ser Ala Asp Leu Ile Gln Pro Asn Pro Pro Ser Ser
 305 310 315 320

Gly Met Leu Gly Ile Ile Ile Met Met Thr Leu Cys Asp Gln Val Asp
 325 330 335

Ile Tyr Glu Phe Leu Pro Ser Lys Arg Lys Thr Asp Val Cys Tyr Tyr
 340 345 350

His Gln Lys Phe Phe Asp Ser Ala Cys Thr Met Gly Ala Tyr Asp Pro
 355 360 365

Leu Leu Phe Glu Lys Asn Met Val Lys His Leu Asn Glu Gly Thr Asp
 370 375 380

Glu Asp Ile Tyr Leu Phe Gly Lys Ala Thr Leu Ser Gly Phe Arg Asn
 385 390 395 400

Ile Arg Cys

<210> 47
 <211> 180
 <212> DNA
 <213> Phaseolus vulgaris

<220>
 <221> CDS
 <222> (1)..(180)
 <223> Coding sequence for amino terminal region of chitinase
 comprising
 a part of a storage vacuole targeting sequence

<300>
 <308> GenBank/M13968
 <309> 1993-04-27

<400> 47
 atg aag aag aat agg atg atg atg atg ata tgg agc gta gga gtg gtg
 48
 Met Lys Lys Asn Arg Met Met Met Met Ile Trp Ser Val Gly Val Val
 1 5 10 15
 tgg atg ctg ttg ttg gtt gga gga agc tac gga gag cag tgt gga agg
 96
 Trp Met Leu Leu Leu Val Gly Gly Ser Tyr Gly Glu Gln Cys Gly Arg
 20 25 30
 caa gca gga ggt gca ctc tgc cca ggg ggc aac tgt tgc agc cag ttc
 144
 Gln Ala Gly Gly Ala Leu Cys Pro Gly Gly Asn Cys Cys Ser Gln Phe
 35 40 45

ggg tgg tgc ggc tcc acc acc gac tac tgc ggc ccg

180

Gly Trp Cys Gly Ser Thr Thr Asp Tyr Cys Gly Pro
50 55 60

<210> 48

<211> 60

<212> PRT

<213> Phaseolus vulgaris

<400> 48

Met Lys Lys Asn Arg Met Met Met Met Ile Trp Ser Val Gly Val Val
1 5 10 15

Trp Met Leu Leu Leu Val Gly Gly Ser Tyr Gly Glu Gln Cys Gly Arg
20 25 30

Gln Ala Gly Gly Ala Leu Cys Pro Gly Gly Asn Cys Cys Ser Gln Phe
35 40 45

Gly Trp Cys Gly Ser Thr Thr Asp Tyr Cys Gly Pro
50 55 60

<210> 49

<211> 411

<212> DNA

<213> Phaseolus vulgaris

<220>

<221> CDS

<222> (1)..(411)

<223> Coding sequence for carboxyl terminal region of chitinase
compris

ing a part of a storage vacuole targeting sequence

<400> 49

ata tcc tgg aac tac aac tac ggt cag tgc gga agg gcc att ggg gtt
48

Ile Ser Trp Asn Tyr Asn Tyr Gly Gln Cys Gly Arg Ala Ile Gly Val
1 5 10 15

gac ttg ctc aac aaa cct gat cta gtc gcc act gac tct gtc atc tcc
96

Asp Leu Leu Asn Lys Pro Asp Leu Val Ala Thr Asp Ser Val Ile Ser
20 25 30

ttc aag tcc gcc ctc tgg ttc tgg atg acc gca cag tcc ccc aag cct
144

Phe Lys Ser Ala Leu Trp Phe Trp Met Thr Ala Gln Ser Pro Lys Pro
35 40 45

tcc tcc cac gac gtc atc acc tct cga tgg acc ccc tcc tct gcc gac
192

Ser Ser His Asp Val Ile Thr Ser Arg Trp Thr Pro Ser Ser Ala Asp
50 55 60

gtc gcc gcc cgc cgg ctt ccc ggc tac ggc act gtg acg aac atc atc
240

Val Ala Ala Arg Arg Leu Pro Gly Tyr Gly Thr Val Thr Asn Ile Ile
65 70 75 80

aac gga ggc ctg gag tgc ggg cga gga cag gac agc agg gtt caa gac
288

Asn Gly Gly Leu Glu Cys Gly Arg Gly Gln Asp Ser Arg Val Gln Asp
85 90 95

cgc atc gga ttc ttc aag aga tac tgt gat ctg ctt gga gtc ggt tat
336

Arg Ile Gly Phe Phe Lys Arg Tyr Cys Asp Leu Leu Gly Val Gly Tyr
100 105 110

ggc aac aac ctt gac tgc tac tct cag act cca ttt gga aat tca ctc
384

Gly Asn Asn Leu Asp Cys Tyr Ser Gln Thr Pro Phe Gly Asn Ser Leu
115 120 125

tta ctc tct gac ctt gtc acc tct cag
411

Leu Leu Ser Asp Leu Val Thr Ser Gln
130 135

<210> 50

<211> 137

<212> PRT

<213> Phaseolus vulgaris

<400> 50

Ile Ser Trp Asn Tyr Asn Tyr Gly Gln Cys Gly Arg Ala Ile Gly Val
1 5 10 15

Asp Leu Leu Asn Lys Pro Asp Leu Val Ala Thr Asp Ser Val Ile Ser
20 25 30

Phe Lys Ser Ala Leu Trp Phe Trp Met Thr Ala Gln Ser Pro Lys Pro
35 40 45

Ser Ser His Asp Val Ile Thr Ser Arg Trp Thr Pro Ser Ser Ala Asp
50 55 60

Val Ala Ala Arg Arg Leu Pro Gly Tyr Gly Thr Val Thr Asn Ile Ile
65 70 75 80

Asn Gly Gly Leu Glu Cys Gly Arg Gly Gln Asp Ser Arg Val Gln Asp
85 90 95

Arg Ile Gly Phe Phe Lys Arg Tyr Cys Asp Leu Leu Gly Val Gly Tyr
100 105 110

Gly Asn Asn Leu Asp Cys Tyr Ser Gln Thr Pro Phe Gly Asn Ser Leu
115 120 125

Leu Leu Ser Asp Leu Val Thr Ser Gln
130 135

<210> 51
 <211> 666
 <212> DNA
 <213> Arabidopsis thaliana

<220>
 <221> gene
 <222> (1)..(666)
 <223> VTI1a gene encoding v-SNARE AtVTI1a protein comprising a
 trans-Golgi network targeting sequence

<220>
 <221> CDS
 <222> (1)..(663)
 <223> v-SNARE AtVTI1a comprising a trans-Golgi network targeting
 sequence

<300>
 <308> GenBank/NM_123313
 <309> 2002-01-10

<400> 51
 atg agt gac gtg ttt gat gga tat gag cga caa tac tgt gag ctc tcg
 48
 Met Ser Asp Val Phe Asp Gly Tyr Glu Arg Gln Tyr Cys Glu Leu Ser
 1 5 10 15
 gcg agt ctt tca aag aag tgt tcc tct gct att tcg ctt gat gga gaa
 96
 Ala Ser Leu Ser Lys Lys Cys Ser Ser Ala Ile Ser Leu Asp Gly Glu
 20 25 30
 caa aag aag cag aag ttg tct gaa ata aaa tcc ggg ctg gaa aat gca
 144
 Gln Lys Lys Gln Lys Leu Ser Glu Ile Lys Ser Gly Leu Glu Asn Ala
 35 40 45
 gaa gta ttg att agg aaa atg gat ctc gag gca agg acc ctt cca cct
 192
 Glu Val Leu Ile Arg Lys Met Asp Leu Glu Ala Arg Thr Leu Pro Pro
 50 55 60
 aat ctt aag tcc agc ctc ctt gtc aaa tta aga gaa ttc aag tct gat
 240
 Asn Leu Lys Ser Ser Leu Leu Val Lys Leu Arg Glu Phe Lys Ser Asp
 65 70 75 80
 ctg aac aat ttc aaa act gaa gtg aag aga atc aca tct gga cag ttg
 288
 Leu Asn Asn Phe Lys Thr Glu Val Lys Arg Ile Thr Ser Gly Gln Leu
 85 90 95
 aac gct gct gct cgt gat gag ttg cta gaa gct ggt atg gct gac aca
 336
 Asn Ala Ala Ala Arg Asp Glu Leu Leu Glu Ala Gly Met Ala Asp Thr
 100 105 110
 aag acg gct tca gct gat cag aga gca aga ttg atg atg tca act gag
 384
 Lys Thr Ala Ser Ala Asp Gln Arg Ala Arg Leu Met Met Ser Thr Glu
 115 120 125

cgt ctg ggc cga acc act gac aga gtc aag gac agt cgt aga aca atg
432
Arg Leu Gly Arg Thr Thr Asp Arg Val Lys Asp Ser Arg Arg Thr Met
130 135 140

atg gag act gaa gaa atc ggt gtt tca atc ctc caa gac ttg cac ggt
480
Met Glu Thr Glu Glu Ile Gly Val Ser Ile Leu Gln Asp Leu His Gly
145 150 155 160

cag aga cag tct ctc tta cgt gcc cac gaa acg ctt cat gga gta gac
528
Gln Arg Gln Ser Leu Leu Arg Ala His Glu Thr Leu His Gly Val Asp
165 170 175

gat aac att gga aag agc aag aag atc ttg aca gac atg acg agg agg
576
Asp Asn Ile Gly Lys Ser Lys Lys Ile Leu Thr Asp Met Thr Arg Arg
180 185 190

atg aac aag aac aaa tgg acc att gga gcc atc atc atc gca ttg att
624
Met Asn Lys Asn Lys Trp Thr Ile Gly Ala Ile Ile Ile Ala Leu Ile
195 200 205

gcc gcc att ttc atc atc ttg tac ttc aaa ctc acc aag taa
666
Ala Ala Ile Phe Ile Ile Leu Tyr Phe Lys Leu Thr Lys
210 215 220

<210> 52
<211> 221
<212> PRT
<213> Arabidopsis thaliana

<400> 52

Met Ser Asp Val Phe Asp Gly Tyr Glu Arg Gln Tyr Cys Glu Leu Ser
1 5 10 15

Ala Ser Leu Ser Lys Lys Cys Ser Ser Ala Ile Ser Leu Asp Gly Glu
20 25 30

Gln Lys Lys Gln Lys Leu Ser Glu Ile Lys Ser Gly Leu Glu Asn Ala
35 40 45

Glu Val Leu Ile Arg Lys Met Asp Leu Glu Ala Arg Thr Leu Pro Pro
50 55 60

Asn Leu Lys Ser Ser Leu Leu Val Lys Leu Arg Glu Phe Lys Ser Asp
65 70 75 80

Leu Asn Asn Phe Lys Thr Glu Val Lys Arg Ile Thr Ser Gly Gln Leu
85 90 95

Asn Ala Ala Ala Arg Asp Glu Leu Leu Glu Ala Gly Met Ala Asp Thr
100 105 110

Lys Thr Ala Ser Ala Asp Gln Arg Ala Arg Leu Met Met Ser Thr Glu
 115 120 125

Arg Leu Gly Arg Thr Thr Asp Arg Val Lys Asp Ser Arg Arg Thr Met
 130 135 140

Met Glu Thr Glu Glu Ile Gly Val Ser Ile Leu Gln Asp Leu His Gly
 145 150 155 160

Gln Arg Gln Ser Leu Leu Arg Ala His Glu Thr Leu His Gly Val Asp
 165 170 175

Asp Asn Ile Gly Lys Ser Lys Lys Ile Leu Thr Asp Met Thr Arg Arg
 180 185 190

Met Asn Lys Asn Lys Trp Thr Ile Gly Ala Ile Ile Ile Ala Leu Ile
 195 200 205

Ala Ala Ile Phe Ile Ile Leu Tyr Phe Lys Leu Thr Lys
 210 215 220

<210> 53
 <211> 207
 <212> DNA
 <213> Ipomoea batatas

<220>
 <221> CDS
 <222> (1)..(207)
 <223> Coding sequence for partial sporamin comprising a lytic
 vacuole t
 argeting sequence

<300>
 <308> GenBank/M16883
 <309> 1993-04-27

<400> 53
 atg aaa gcc ctc gca ctc ttc ttc cta ctt tcc ctc tat ctc ctc ccc
 48
 Met Lys Ala Leu Ala Leu Phe Phe Leu Leu Ser Leu Tyr Leu Leu Pro
 1 5 10 15
 aac cca gct cat tcc aag ttc aat ccc atc cgc ctc cgc ccc gcc cac
 96
 Asn Pro Ala His Ser Lys Phe Asn Pro Ile Arg Leu Arg Pro Ala His
 20 25 30
 gaa acg gcg tcg tcc gaa act ccg gtg ctc gac atc aac ggc gac gaa
 144
 Glu Thr Ala Ser Ser Glu Thr Pro Val Leu Asp Ile Asn Gly Asp Glu
 35 40 45
 gtc cgg gcc ggc gaa aat tac tac att gtc tcc gcc ata tgg ggc gcc
 192
 Val Arg Ala Gly Glu Asn Tyr Tyr Ile Val Ser Ala Ile Trp Gly Ala
 50 55 60

ggc gga gga ggc ctg
207
Gly Gly Gly Gly Leu
65

<210> 54
<211> 69
<212> PRT
<213> Ipomoea batatas

<400> 54

Met Lys Ala Leu Ala Leu Phe Phe Leu Leu Ser Leu Tyr Leu Leu Pro
1 5 10 15

Asn Pro Ala His Ser Lys Phe Asn Pro Ile Arg Leu Arg Pro Ala His
20 25 30

Glu Thr Ala Ser Ser Glu Thr Pro Val Leu Asp Ile Asn Gly Asp Glu
35 40 45

Val Arg Ala Gly Glu Asn Tyr Tyr Ile Val Ser Ala Ile Trp Gly Ala
50 55 60

Gly Gly Gly Gly Leu
65

<210> 55
<211> 465
<212> DNA
<213> Homo sapiens

<220>
<221> gene
<222> (1)..(465)
<223> Gene for partial EBD (C-terminal domain of human EEA1)
comprising
a phosphatidylinositol 3-phosphate targeting sequence

<220>
<221> CDS
<222> (1)..(462)
<223> Partial EBD (C-terminal domain of human EEA1) comprising a
phosph
'atidylinositol 3-phosphate targeting sequence

<300>
<308> GenBank/NM_003566
<309> 2002-05-14

<400> 55
tgg caa tct agt caa cgg aga gtt agt gag ctt gag aaa caa acg gat
48
Trp Gln Ser Ser Gln Arg Arg Val Ser Glu Leu Glu Lys Gln Thr Asp
1 5 10 15

gac tta cgg ggt gaa att gca gta tta gaa gca acg gtt cag aat aat

96
 Asp Leu Arg Gly Glu Ile Ala Val Leu Glu Ala Thr Val Gln Asn Asn
 20 25 30

 caa gat gaa agg aga gca cta ctg gaa aga tgt ctt aaa gga gaa ggt
 144
 Gln Asp Glu Arg Arg Ala Leu Leu Glu Arg Cys Leu Lys Gly Glu Gly
 35 40 45

 gaa ata gaa aag ctt caa acc aaa gta tta gaa ttg caa aga aag ctg
 192
 Glu Ile Glu Lys Leu Gln Thr Lys Val Leu Glu Leu Gln Arg Lys Leu
 50 55 60

 gat aat aca act gca gtg cag gag ctg ggc aga gaa aac caa tca ctt
 240
 Asp Asn Thr Thr Ala Val Gln Glu Leu Gly Arg Glu Asn Gln Ser Leu
 65 70 75 80

 cag atc aaa cat aca caa gcg ttg aat aga aag tgg gcc gaa gac aat
 288
 Gln Ile Lys His Thr Gln Ala Leu Asn Arg Lys Trp Ala Glu Asp Asn
 85 90 95

 gaa gta caa aac tgt atg gcc tgt ggg aaa ggc ttt tca gta aca gtg
 336
 Glu Val Gln Asn Cys Met Ala Cys Gly Lys Gly Phe Ser Val Thr Val
 100 105 110

 aga cgg cat cac tgc cga cag tgt gga aat atc ttc tgt gct gaa tgt
 384
 Arg Arg His His Cys Arg Gln Cys Gly Asn Ile Phe Cys Ala Glu Cys
 115 120 125

 tca gcc aaa aat gcc tta act cct tcc tcc aag aag cct gtt cgt gtc
 432
 Ser Ala Lys Asn Ala Leu Thr Pro Ser Ser Lys Lys Pro Val Arg Val
 130 135 140

 tgt gat gca tgt ttc aat gac ttg caa gga taa
 465
 Cys Asp Ala Cys Phe Asn Asp Leu Gln Gly
 145 150

<210> 56
 <211> 154
 <212> PRT
 <213> Homo sapiens

<400> 56

Trp Gln Ser Ser Gln Arg Arg Val Ser Glu Leu Glu Lys Gln Thr Asp
 1 5 10 15

Asp Leu Arg Gly Glu Ile Ala Val Leu Glu Ala Thr Val Gln Asn Asn
 20 25 30

Gln Asp Glu Arg Arg Ala Leu Leu Glu Arg Cys Leu Lys Gly Glu Gly
 35 40 45

Glu Ile Glu Lys Leu Gln Thr Lys Val Leu Glu Leu Gln Arg Lys Leu
 50 55 60

Asp Asn Thr Thr Ala Val Gln Glu Leu Gly Arg Glu Asn Gln Ser Leu
65 70 75 80

Gln Ile Lys His Thr Gln Ala Leu Asn Arg Lys Trp Ala Glu Asp Asn
85 90 95

Glu Val Gln Asn Cys Met Ala Cys Gly Lys Gly Phe Ser Val Thr Val
100 105 110

Arg Arg His His Cys Arg Gln Cys Gly Asn Ile Phe Cys Ala Glu Cys
115 120 125

Ser Ala Lys Asn Ala Leu Thr Pro Ser Ser Lys Lys Pro Val Arg Val
130 135 140

Cys Asp Ala Cys Phe Asn Asp Leu Gln Gly
145 150

<210> 57
<211> 438
<212> DNA
<213> Arabidopsis thaliana

<220>
<221> gene
<222> (1)..(438)
<223> Gene for pleckstrin homology domain of Arabidopsis thaliana
(AtPH), a putative homologue of human pleckstrin domain, comprising
a phosphatidylinositol 3-phosphate targeting sequence

<220>
<221> CDS
<222> (1)..(435)
<223> pleckstrin homology domain of Arabidopsis thaliana (AtPH), a
putative homologue of human pleckstrin domain, comprising a
phosphatidylinositol 3-phosphate targeting sequence

<300>
<308> GenBank/AB005903
<309> 1999-10-05

<400> 57
atg gag agt atg tgg cga atc gcg acg gga caa gat ccg agc cgt gaa
48
Met Glu Ser Met Trp Arg Ile Ala Thr Gly Gln Asp Pro Ser Arg Glu
1 5 10 15
gat tac gaa ggg atc gag ttc tgg tca aac cct gag cgt tct ggt tgg
96
Asp Tyr Glu Gly Ile Glu Phe Trp Ser Asn Pro Glu Arg Ser Gly Trp
20 25 30

gtt ctc aaa cga ggg aag ctt ctc tgg ttc aaa gat caa gcc gct gct
192

gga att cgt gga tct acg ccg cgt ggt gtg atc tcc gtt ggt gat tgt
240

ctc acc gtg aaa gga gct gag gat gtt gtg aat aag cct ttt gct ttt
288

gag cta tct agt ggt agc tat acc atg ttc ttc att gct gat aat gag
336

aag gag aaa gaa gag tgg att aat tcg att gga aga tcg att gtg caa
384

cac tcg agg tct gtc acg gat tct gag gtt ctc gat tat gat cac agg
432

cgg tga
438
Arg
145

```
<210> 58
<211> 145
<212> PRT
<213> Arabidopsis thaliana
```

<400> 58

Met Glu Ser Met Trp Arg Ile Ala Thr Gly Gln Asp Pro Ser Arg Glu
1 5 10 15

Asp Tyr Glu Gly Ile Glu Phe Trp Ser Asn Pro Glu Arg Ser Gly Trp
20 25 30

Leu Thr Lys Gln Gly Asp Tyr Ile Lys Thr Trp Arg Arg Arg Trp Phe
35 40 45

Val Leu Lys Arg Gly Lys Leu Leu Trp Phe Lys Asp Gln Ala Ala Ala
50 55 60

Gly Ile Arg Gly Ser Thr Pro Arg Gly Val Ile Ser Val Gly Asp Cys
65 70 75 80

Leu Thr Val Lys Gly Ala Glu Asp Val Val Asn Lys Pro Phe Ala Phe
85 90 95

Glu Leu Ser Ser Gly Ser Tyr Thr Met Phe Phe Ile Ala Asp Asn Glu
100 105 110

Lys Glu Lys Glu Glu Trp Ile Asn Ser Ile Gly Arg Ser Ile Val Gln
115 120 125

His Ser Arg Ser Val Thr Asp Ser Glu Val Leu Asp Tyr Asp His Arg
130 135 140

Arg
145

<210> 59
<211> 282
<212> DNA
<213> Mus musculus

<220>
<221> CDS
<222> (1)..(282)
<223> Coding sequence for partial pleckstrin homology domain of FAPP
(family A (phosphoinositide binding specific) member 3),
comprising
a phosphatidylinositol 4-phosphate targeting sequence

<300>
<308> GenBank/BC031110
<309> 2002-06-13

<400> 59
atg gag ggg gtt ctg tac aag tgg acc aac tat ctc aca ggt tgg cag
48
Met Glu Gly Val Leu Tyr Lys Trp Thr Asn Tyr Leu Thr Gly Trp Gln
1 5 10 15
cct cga tgg ttt gtt ctg gat aat gga atc ctg tcc tac tat gac tca
96
Pro Arg Trp Phe Val Leu Asp Asn Gly Ile Leu Ser Tyr Tyr Asp Ser
20 25 30
cag gat gat gtc tgc aaa ggg agc aaa ggg agt ata aag atg gcg gtc
144
Gln Asp Asp Val Cys Lys Gly Ser Lys Gly Ser Ile Lys Met Ala Val
35 40 45
tgt gag att aaa gtc cat ccc gca gac aac aca aga atg gag tta atc
192
Cys Glu Ile Lys Val His Pro Ala Asp Asn Thr Arg Met Glu Leu Ile
50 55 60
att cca gga gag cag cat ttc tac atg aag gca gta aat gcc gcc gag
240
Ile Pro Gly Glu Gln His Phe Tyr Met Lys Ala Val Asn Ala Ala Glu
65 70 75 80

aga cag agg tgg ctg gtt gcc ctt ggg agc tcc aaa gcg tgc
 282
 Arg Gln Arg Trp Leu Val Ala Leu Gly Ser Ser Lys Ala Cys
 85 90

<210> 60
 <211> 94
 <212> PRT
 <213> Mus musculus

<400> 60

Met Glu Gly Val Leu Tyr Lys Trp Thr Asn Tyr Leu Thr Gly Trp Gln
 1 5 10 15

Pro Arg Trp Phe Val Leu Asp Asn Gly Ile Leu Ser Tyr Tyr Asp Ser
 20 25 30

Gln Asp Asp Val Cys Lys Gly Ser Lys Gly Ser Ile Lys Met Ala Val
 35 40 45

Cys Glu Ile Lys Val His Pro Ala Asp Asn Thr Arg Met Glu Leu Ile
 50 55 60

Ile Pro Gly Glu Gln His Phe Tyr Met Lys Ala Val Asn Ala Ala Glu
 65 70 75 80

Arg Gln Arg Trp Leu Val Ala Leu Gly Ser Ser Lys Ala Cys
 85 90

<210> 61
 <211> 525
 <212> DNA
 <213> Rattus norvegicus

<220>
 <221> gene
 <222> (1)..(525)
 <223> Gene for partial pleckstrin homology domain of phospholipase
 C-de lta 1 comprising phosphatidylinositol 4,5-diphosphate
 targeting s
 equence

<220>
 <221> CDS
 <222> (1)..(522)
 <223> Partial pleckstrin homology domain of phospholipase C-delta 1
 com prising phosphatidylinositol 4,5-diphosphate targeting
 sequence

<300>
 <308> GenBank/NM_017035
 <309> 2000-11-01

```

<400> 61
gac tcg ggt agg gac ttc ctg acc ctg cac ggg ctc cag gat gac ccg
48
Asp Ser Gly Arg Asp Phe Leu Thr Leu His Gly Leu Gln Asp Asp Pro
1 5 10 15

gac ctt cag gcc ctt ctg aag ggc agc cag ctt ctg aag gtg aag tcc
96
Asp Leu Gln Ala Leu Leu Lys Gly Ser Gln Leu Leu Lys Val Lys Ser
20 25 30

agc tcg tgg cgt agg gaa cgc ttc tac aag cta cag gag gac tgc aag
144
Ser Ser Trp Arg Arg Glu Arg Phe Tyr Lys Leu Gln Glu Asp Cys Lys
35 40 45

acc atc tgg cag gaa tct cga aag gtc atg agg tcc ccg gag tcg cag
192
Thr Ile Trp Gln Glu Ser Arg Lys Val Met Arg Ser Pro Glu Ser Gln
50 55 60

ctg ttc tcc atc gag gac att cag gag gta cgg atg gga cac cgc aca
240
Leu Phe Ser Ile Glu Asp Ile Gln Glu Val Arg Met Gly His Arg Thr
65 70 75 80

gaa ggc ctg gag aag ttt gcc cga gac atc ccc gag gat cga tgc ttc
288
Glu Gly Leu Glu Lys Phe Ala Arg Asp Ile Pro Glu Asp Arg Cys Phe
85 90 95

tcc att gtc ttc aag gac cag cgc aac acc cta gac ctc att gcc cca
336
Ser Ile Val Phe Lys Asp Gln Arg Asn Thr Leu Asp Leu Ile Ala Pro
100 105 110

tca cca gct gac gct cag cac tgg gtg cag ggc ctg cgc aag atc atc
384
Ser Pro Ala Asp Ala Gln His Trp Val Gln Gly Leu Arg Lys Ile Ile
115 120 125

cac cac tcc ggc tcc atg gac cag cgg cag aag ctg cag cac tgg att
432
His His Ser Gly Ser Met Asp Gln Arg Gln Lys Leu Gln His Trp Ile
130 135 140

cac tcc tgc ttg cga aag gct gat aaa aac aag gac aac aag atg aac
480
His Ser Cys Leu Arg Lys Ala Asp Lys Asn Lys Asp Asn Lys Met Asn
145 150 155 160

ttc aag gag ctg aag gac ttc ctg aag gag ctc aac atc cag tga
525
Phe Lys Glu Leu Lys Asp Phe Leu Lys Glu Leu Asn Ile Gln
165 170

```

```

<210> 62
<211> 174
<212> PRT
<213> Rattus norvegicus

```

```

<400> 62
Asp Ser Gly Arg Asp Phe Leu Thr Leu His Gly Leu Gln Asp Asp Pro
1 5 10 15

```

Asp Leu Gln Ala Leu Leu Lys Gly Ser Gln Leu Leu Lys Val Lys Ser
20 25 30

Ser Ser Trp Arg Arg Glu Arg Phe Tyr Lys Leu Gln Glu Asp Cys Lys
35 40 45

Thr Ile Trp Gln Glu Ser Arg Lys Val Met Arg Ser Pro Glu Ser Gln
50 55 60

Leu Phe Ser Ile Glu Asp Ile Gln Glu Val Arg Met Gly His Arg Thr
65 70 75 80

Glu Gly Leu Glu Lys Phe Ala Arg Asp Ile Pro Glu Asp Arg Cys Phe
85 90 95

Ser Ile Val Phe Lys Asp Gln Arg Asn Thr Leu Asp Leu Ile Ala Pro
100 105 110

Ser Pro Ala Asp Ala Gln His Trp Val Gln Gly Leu Arg Lys Ile Ile
115 120 125

His His Ser Gly Ser Met Asp Gln Arg Gln Lys Leu Gln His Trp Ile
130 135 140

His Ser Cys Leu Arg Lys Ala Asp Lys Asn Lys Asp Asn Lys Met Asn
145 150 155 160

Phe Lys Glu Leu Lys Asp Phe Leu Lys Glu Leu Asn Ile Gln
165 170